AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92 NEUILLY SUR SEINE FRANCE

Theory and Applications of Kalman Filtering

NORTH ATLANTIC TREATY ORGANIZATION

AGARDograph No. 139

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

THEORY AND APPLICATIONS OF KALMAN FILTERING

Edited by

C. T. Leondes

University of California
Los Angeles, USA

This AGARDograph was prepared at the request of the Guidance and Control Panel of AGARD – NATO

Published February 1970

629.7.058 : 519.61 : 629.78

Printed by Technical Editing and Reproduction Ltd
Harford House, 7-9 Charlotte Street, London W1P 1HD.

PREFACE

The theory and applications of Kalman filtering have now been under development for over ten years. This general area has achieved a level of maturity and importance to easily justify a thorough treatment of the subject. In view of the particular importance of Kalman filtering techniques to the field of guidance and control the Guidance and Control Panel of NATO-AGARD recommended that such a text be developed under the auspices of NATO-AGARD.

This textbook is the result of that recommendation. The text is organized into three principal parts. The first part examines the theory of Kalman filtering in depth. A number of significant new results of fundamental importance are included here. For instance, such questions as existence of Kalman filters under very general conditions, Kalman filtering for Gauss-Markov processes, suboptimal Kalman filtering techniques, and other areas are treated here.

The second part of the text deals with the general area of related topics. Questions of the comparison of Kalman filtering with other approaches such as Bayesian and maximum likelihood estimation, nonlinear filtering, linear and nonlinear smoothing (post-flight data analysis), and other topics are reviewed in depth in this part.

The third part is a very comprehensive review of many of the important applications of Kalman filtering. Although many very specific areas of application are treated in this part, many general principles and techniques for a very broad range of applications of Kalman filtering will be found here. As a result, the reader should also find this part quite valuable no matter what particular application he might have in mind.

It is a great pleasure to acknowledge the contributions of many individuals who made this text possible. First of all, Professor W.Wrigley's contributions as the first Panel Chairman of the Guidance and Control Panel cannot be praised too highly. His outstanding efforts in guiding this Panel during its formative stages cannot be praised enough. This book is one of the fruits of his efforts. All the members of the Guidance and Control Panel were solicited for their advice and suggestions, and their help is also greatly appreciated. Colonel L.Sugarman and Mr.Frank Sullivan provided much important support. Mr O.H.Schuck made many important suggestions. Colonel W.Studabaker and Major C.Mount provided invaluable assistance in their roles as Guidance and Control Panel Executive Officers. Mrs Gladys Flynn provided many invaluable services.

Cornelius T.Leondes

CONTENTS

CHAPTER I - LINEAR ESTIMATION THEORY

by

H. W. Sorenson* and A. R. Stubberud†

* University of California,
San Diego, La Jolla,
California, USA

† University of California,
Irvine, California, USA

NOTATION

t	the independent variable, usually referred to as time
t_0	initial time
$\underline{r}(t)$	the residual or innovations process and defined to be $\underline{r}(t/\tau) \triangleq \underline{z}(t) - H(t)\hat{\underline{x}}(t/\tau)$
$\underline{v}(t)$	additive white-noise process in the measurement data with mean zero and covariance matrix $R(t)\delta(t-\tau)$
$\underline{w}(t)$	additive white-noise process in the plant with mean zero and covariance matrix $Q(t)\delta(t-\tau)$
$\underline{x}(t)$	the n-dimensional state vector described by a linear differential equation $\dot{\underline{x}} = F(t)\underline{x} + \underline{w}$
$\hat{\underline{x}}(t/\tau)$	estimate of the state $\underline{x}(t)$ given the measurement data $\underline{Z}(\tau)$
$\tilde{\underline{x}}(t/\tau)$	error in the estimate, $\tilde{\underline{x}}(t/\tau) \triangleq \underline{x}(t) - \hat{\underline{x}}(t/\tau)$
$\underline{y}(t)$	the noise-free measurement data, $\underline{y}(t) \triangleq H(t)\underline{x}(t)$
$\underline{z}(t)$	the measurement data, $\underline{z}(t) \triangleq H(t)\underline{x} + \underline{v}$
$\underline{Z}(\tau)$	the measurement data, $\underline{Z}(\tau)$ for all $t_0 \leq \tau < t$
$C(t)$	cross-correlation matrix for plant and measurement noise, $E[\underline{w}(t)\underline{v}^T(\tau)] = C(t)\delta(t-\tau)$
$F(t)$	plant matrix, $\dot{\underline{x}} = F(t)\underline{x} + \underline{w}$
$H(t)$	observation matrix, $\underline{z} = H(t)\underline{x} + \underline{v}$
$K(t)$	optimal gain matrix for unbiased minimum variance filter
$M(t,\tau)$	observability matrix, $M(t,\tau) \triangleq \int_t^\tau \Phi^T(\sigma,t)H^T(\sigma)R^{-1}(\sigma)H(\sigma)\Phi(\sigma,t)\,d\sigma$
$P(t/\tau)$	covariance of the error in the estimate, $P(t/\tau) \triangleq E[\tilde{\underline{x}}(t/\tau)\tilde{\underline{x}}^T(b/\tau)]$
$Q(t)$	covariance matrix for the plant noise $\underline{w}(t)$
$R(t)$	covariance matrix for the measurement noise $\underline{v}(t)$
$\Phi(t,\tau)$	transition matrix associated with the plant equation, $\dot{\Phi}(t,\tau) = F(t)\Phi(t,\tau)$, $\Phi(\tau,\tau) = I$
$\Psi(t,\tau)$	transition matrix associated with the filter dynamics, $\dot{\Psi}(t,\tau) = [F(t) - K(t)H(t)]\Psi(t,\tau)$, $\Psi(\tau,\tau) = I$
I	identity matrix of appropriate dimension
$(\underline{\ })$	the quantity ( ) is a vector (or column matrix)
$(\ )^T$	the matrix transpose of ( )
$(\ )^{-1}$	the matrix inverse of ( )
$(\ )^{-T}$	the matrix inverse of $(\ )^T$
$E(a)$	expected value of the random variable a
$\delta(\cdot)$	the Dirac delta function
$(\ ) \triangleq [\ ]$	the quantity ( ) is defined to be equivalent to [ ]

CHAPTER 1 - LINEAR ESTIMATION THEORY

H.W. Sorenson and A.R. Stubberud

1. THE LINEAR ESTIMATION PROBLEM

1.1 Introduction to the Linear Estimation Problem

A large class of estimation problems is concerned with finding an optimal estimate of some quantity (an unknown parameter, a random variable, or a random signal) when a linear function of this quantity corrupted by an additive noise is available for generating the estimate. One of the first studies of this class of problems was performed by Gauss[1] in the early 1800's. In this work he studied least-squares estimates of unknown parameters.

In the early 1940's Wiener[2] and Kolmogorov[3] attacked the class of problems dealing with estimation of random signals. The key result from this work is an integral equation called the Wiener-Hopf equation. The solution of this equation is a weighting function which, when convolved with the corrupted linear measurement, produces an unbiased minimum variance estimate of the random signal. Since the Wiener-Hopf equation can only be solved explicitly for certain special cases of the general problem, this work has only limited practical application. Many generalizations of this work were presented in the 1940's and 1950's but none of these improved its practical applicability[4-7].

In the 1950's the idea of generating least-squares estimates recursively was introduced. This interest was stimulated by the increased usage of digital computers. These estimates are generated dynamically and the data processing algorithm is either a differential or a difference equation. Carlton[8] asserts that the first work on this subject was done by Follin[9] about 1955. In 1958 Swerling published a Rand Corporation report (which received wider distribution in 1959 in the Journal of the Astronautical Sciences[10]) that presented a recursive filtering procedure similar to that described shortly thereafter by Kalman[11]. The latter work is generally considered to have sparked the widespread interest in the subject and subsequent references to "Kalman filtering".

The paper by Kalman[11] in 1960 introduced a different approach to the problem of Wiener and Kolmogorov for random sequences. In 1961, Kalman and Bucy[12] generalized the results to random processes. Basically, this approach circumvents the problem of solving the Wiener-Hopf integral equation. By recognizing that digital computers are much more effective at solving differential equations than integral equations, Kalman and Bucy transformed the integral equation into an equivalent differential equation. Then, rather than demand an analytical solution for this equation, they recognized that, from a practical standpoint, it is better to put the computational burden on the computer. These results are closely related to those obtained for sequential least-squares estimation. The practicality of the Kalman approach to the estimation problem has made it immensely popular in aerospace applications, such as in navigation and guidance.

It is the intent in this chapter to discuss the fundamental aspects of the unbiased, minimum variance, linear estimation problem. The treatment of the problem as presented here has identifiable roots in the paper by Kalman and Bucy[12] and even more so in the remarkable treatise by Kalman that constitutes Reference 13. However, the presentation describes not only those early and basic results but attempts to present the developments in the years since the seminal papers appeared that either provide insights into the foundations or which add significantly to the theoretical structure. It may be an overstatement to suggest that there are as many derivations of the Kalman filter equations as there are workers in the field. On the other hand it cannot be denied that there are many approaches to the problem and each has its vociferous supporters. It is not practical, nor is it desirable, to attempt to satisfy all tastes. The development here is designed to appeal to intuition and to provide insight that explains those aspects which may not be obvious. In taking this approach, mathematical rigor is sacrificed or details are omitted when it appears that greater clarity can be achieved by ignoring what amounts to be a technical detail that does not affect the final results.

The basic mathematical model and a more precise definition of the linear estimation problem are given in Section 1.2. Some variations and generalizations of this model and problem are considered in Sections 2 and 3. A detailed summary of the more important results of Sections 2 and 3 is given in Section 4. It is suggested that this section be read after completing Section 1.2 and before going to the detailed discussions of Sections 2 and 3. In this way this summary can be used as a guide through the seeming morass of Kalman-Bucy filter theory.

1.2 The Mathematical Model and Problem Statement

Consider a dynamical system whose state evolution is described by a linear, stochastic, vector differential equation.

$$\dot{\underline{x}} = F(t)\underline{x} + \underline{w} . \tag{1.1}$$

where

$\underline{x}$ is an n-dimensional state vector

$F(t)$ is an $n \times n$ matrix whose elements are continuous functions of the independent variable t

$\underline{w}$ is an n-dimensional, gaussian white-noise process with the following statistics

$$E[\underline{w}(t)] = 0 \quad \text{for all } t$$

$$E[\underline{w}(t)\underline{w}^T(\tau)] = Q(t)\delta(t-\tau) ,$$

where $Q(t)$ is an $n \times n$ symmetric, non-negative-definite matrix and $\delta(t-\tau)$ is the Dirac delta function. The initial state $\underline{x}(t_0)$ is a random variable with known statistics.

$$E[\underline{x}(t_0)] = \underline{m}_0$$

$$E\{[\underline{x}(t_0) - \underline{m}_0][\underline{x}(t_0) - \underline{m}_0]^T\} \quad M_0 .$$

Also, $\underline{x}(t_0)$ is independent of $\underline{w}(t)$, that is,

$$E[\underline{w}(t)\underline{x}^T(t_0)] = 0 \quad \text{for all } t .$$

The solution of Equation (1.1) is given by[14]

$$\underline{x}(t) = \Phi(t,t_0)\underline{x}(t_0) + \int_{t_0}^{t} \Phi(t,\tau)\underline{w}(\tau)\, d\tau ,$$

where $\Phi(t,\tau)$ is the transition matrix and is the solution of the matrix differential equation

$$\frac{d\Phi(t,\tau)}{dt} = F(t)\Phi(t,\tau) ; \quad \Phi(\tau,\tau) = I \quad \text{for all } \tau .$$

The transition matrix has the property that

$$\Phi(t_k,t_j)\Phi(t_j,t_i) = \Phi(t_k,t_i) \quad \text{for all } t_k,t_j,t_i .$$

This implies that

$$\Phi^{-1}(t_k,t_j) = \Phi(t_j,t_k) ,$$

so that Φ is nonsingular for all the t_k,t_j. The solution and its properties will be used in the succeeding discussion.

The only information available about the state are m measurements $\underline{z}(t)$ that are related to $\underline{x}(t)$ according to

$$\underline{z}(t) = H(t)\underline{x}(t) + \underline{v}(t) . \tag{1.2}$$

The $\underline{v}(t)$ is an m-dimensional, gaussian white-noise process with the statistics

$$E[\underline{v}(t)] = 0 \quad \text{for all } t$$

$$E[\underline{v}(t)\underline{v}^T(\tau)] = R(t)\delta(t-\tau) .$$

The plant noise $\underline{w}(t)$ and the measurement noise $\underline{v}(t)$ are assumed to be independent for most of the discussion, although the generalization to correlated processes is discussed in Section 2.1.3. The symmetric $m \times m$ matrix $R(t)$ must be assumed to be positive-definite for the straightforward development of the Kalman filter equations. When $R(t)$ is not positive-definite, it is necessary to introduce the special considerations relating to the colored-noise problem discussed in Section 2.2. Also, by a corruption of the language, the matrices $Q(t)$ and $R(t)$ will often be referred to as the noise covariance matrices, although the covariance matrices actually are of infinite magnitude because of the delta function.

The estimation problem associated with the system Equations (1.1) - (1.2) can now be defined.

Unbiased, Minimum Variance, Linear Estimation Problem

Given the linear system described by Equations (1.1) and (1.2), determine an estimate $\hat{\underline{x}}(t|\tau)$ of the state $\underline{x}(t)$ that is a linear function of all measurement data* $\underline{z}(\sigma)$, $t_0 \leq \sigma \leq \tau$, and satisfies the following conditions:

(i) $\hat{\underline{x}}(t|\tau)$ is unbiased so that

$$E[\hat{\underline{x}}(t|\tau)] = E[\underline{x}(t)] .$$

(ii) $\hat{\underline{x}}(t|\tau)$ is "best" in the sense that the expected value of the square of the error magnitude is minimized. Thus, the estimate $\hat{\underline{x}}(t|\tau)$ is chosen so that

$$E\{[\underline{x}(t) - \hat{\underline{x}}(t|\tau)]^T [\underline{x}(t) - \hat{\underline{x}}(t|\tau)]\} = \text{minimum} .$$

The estimate that is obtained depends upon the amount of data that is available (i.e. as defined by τ) and can be described in terms of three special problems.

Prediction: Suppose that the state at a time t is to be estimated from data $\underline{Z}(\tau)$ where $\tau < t$. Thus, the state $\underline{x}(t)$ is to be predicted from data obtained at times prior to the time t. This shall be referred to here as the prediction problem.

Filtering: Consider the problem of estimating the $\underline{x}(t)$ from data $\underline{Z}(t)$ (i.e. $\tau = t$). This shall be referred to as the filtering problem.

Smoothing: Suppose that the state $\underline{x}(t)$ is to be estimated from data $\underline{Z}(\tau)$ where $\tau > t$. In this case $\underline{x}(t)$ is estimated from data obtained at times prior to, coincident with, and subsequent to, the time of interest, t. This shall be referred to as the smoothing problem.

Those three special cases of the estimation problem are also referred to as the extrapolation, smoothing and interpolation problems (see, for example, References 2 and 15). The problems have been stated in the order of increasing complexity of solution. The emphasis here is upon the prediction and filtering problems, although the smoothing problem is discussed briefly in Section 2.3. The so-called Kalman-Bucy filter is involved with the first two cases.

2. SOLUTION OF THE LINEAR ESTIMATION PROBLEM

2.1 Linear Prediction and Filtering

2.1.1 *Heuristic Derivation*

A derivation of the Kalman-Bucy equations is presented in this section that appeals primarily to intuitive reasoning rather than mathematical rigor. This approach is taken initially to provide insight into the character of the solution which is shown more rigorously in the following section to provide the unbiased, minimum variance estimate of the state $\underline{x}(t)$ based on the measurement data $\underline{Z}(t)$.

Suppose that an estimate of the state $\underline{x}(t_0)$ is available at some time t_0 that is based upon measurement data $\underline{Z}(t_0)$. Let this estimate be denoted as $\hat{\underline{x}}(t_0|t_0)$. In Equation (1.1), it is seen that the state changes in accordance with a linear differential equation with a white-noise forcing function. The process $\underline{z}(t)$ has zero mean and the value at different times is uncorrelated, regardless of the magnitude of the time difference. Thus, it is reasonable to expect, in the absence of additional data, that the behavior of the estimate would be described by

$$\dot{\hat{\underline{x}}} = \Gamma(t)\hat{\underline{x}} , \qquad t > t_0 , \tag{2.1}$$

where $\hat{\underline{x}}(t_0|t_0)$ is known. But when measurement data are available for time subsequent to t_0 one can consider the *residual* (i.e. the difference between the measurement data $\underline{z}(t)$ and that predicted by the estimate $H(t)\hat{\underline{x}}(t|t)$),

$$\underline{\gamma}(t) = \underline{z}(t) - H(t)\hat{\underline{x}}(t|t) . \tag{2.2}$$

This residual is referred to by Kailath[16] as the innovations process and will be examined in more detail in Section 2.1.2.

The residual can be considered to provide an indication of the error in the estimate $\hat{\underline{x}}(t|t)$. Let us assume that this error will be used to modify the estimate provided by Equation (2.1) by introducing an *unknown* weighting or gain matrix $K(t)$. Using this weighting matrix, assume that the estimate, including all new measurement data, is to have the form

$$\dot{\hat{\underline{x}}}(t|t) = F(t)\hat{\underline{x}}(t|t) + K(t)[\underline{z}(t) - H(t)\hat{\underline{x}}(t|t)] \qquad \text{for } t > t_0 . \tag{2.3}$$

* Henceforth, the collection of measurement data $\underline{z}(\sigma)$ on an interval $t_0 \leq \sigma \leq \tau$ will be denoted as $\underline{Z}(\tau)$.

where $\hat{\underline{x}}(t_0|t_0)$ is known. To simplify notation, the arguments of the estimate will be suppressed, so that $\hat{\underline{x}}$ will be synonymous with $\hat{\underline{x}}(t|t)$

Using the form assumed in Equation (2.3), the matrix $K(t)$ will be chosen so that the variance of the error is minimized as required in the statement of the filtering problem in Section 2.1. Note also that the solution of Equation (2.1) yields the solution of the prediction problem for $t > t_0$.

First note that the estimate provided by Equation (2.3) is unbiased if the initial conditions are selected in an appropriate manner.

Equation (2.3) is an unbiased estimator if the initial condition $\hat{\underline{x}}(t_0|t_0)$ satisfies the constraint that

$$E[\hat{\underline{x}}(t_0|t_0)] = E[\underline{x}(t_0)] . \tag{2.4}$$

To verify this, note from Equation (1.1) that

$$E[\dot{\underline{x}}] = F(t)E[\underline{x}] ,$$

and from Equation (2.3) that

$$E[\dot{\hat{\underline{x}}}] = F(t)E[\hat{\underline{x}}] + K(t)\{E[\underline{z}] - H(t)E[\hat{\underline{x}}]\} .$$

But from Equation (1.2) one sees that

$$E[\underline{z}] = H(t)E[\underline{x}] ,$$

so that

$$E[\dot{\hat{\underline{x}}}] = F(t)E[\hat{\underline{x}}] + K(t)H(t)\{E[\underline{x}] - E[\hat{\underline{x}}]\} .$$

Thus, it follows that

$$\frac{d}{dt}\{E[\underline{x} - \hat{\underline{x}}]\} = \{F(t) - K(t)H(t)\}E[\underline{x} - \hat{\underline{x}}] .$$

This is a homogeneous, linear differential equation in the variable $E[\underline{x} - \hat{\underline{x}}]$ so its solution has the general form

$$E[\underline{x}(t) - \hat{\underline{x}}(t)] = \Psi(t,t_0)E[\underline{x}(t_0) - \hat{\underline{x}}(t_0|t_0)] .$$

The matrix Ψ is the transition matrix associated with the filter dynamics and is the solution of the matrix differential equation

$$\dot{\Psi}(t,t_0) = [F(t) - K(t)H(t)]\Psi(t,t_0) , \qquad \Psi(t_0,t_0) = I .$$

Thus, if $$E[\underline{x}(t_0)] = E[\hat{\underline{x}}(t_0|t_0)] ,$$

it follows that $$E[\underline{x}(t)] = E[\hat{\underline{x}}(t|t)] ,$$

so that the estimate is unbiased.

Using the form for the estimate assumed in Equation (2.3), consider the *reformulated filtering problem*. Determine the time-varying matrix $K(t)$ so that $E[(\underline{x} - \hat{\underline{x}})^T(\underline{x} - \hat{\underline{x}})]$ is minimized.

Let $$\tilde{\underline{x}}(t|t) \triangleq \underline{x}(t) - \hat{\underline{x}}(t|t) , \tag{2.5}$$

so that
$$E[(\underline{x} - \hat{\underline{x}})^T(\underline{x} - \hat{\underline{x}})] = E[\tilde{\underline{x}}^T\tilde{\underline{x}}]$$
$$= \text{trace } E[\tilde{\underline{x}}\tilde{\underline{x}}^T] .$$

Define $$P(t|t) \triangleq E[\tilde{\underline{x}}(t|t)\tilde{\underline{x}}^T(t|t)] . \tag{2.6}$$

Thus, it is desired that the gain matrix $K(t)$ be chosen to minimize the trace of the error covariance matrix $P(t|t)$.

The rate of change of the error in the estimate is obtained by using Equations (1.1), (1.2), and (2.3).

$$\frac{d}{dt}(\tilde{\underline{x}}) = \frac{d}{dt}\underline{x} - \frac{d}{dt}\hat{\underline{x}}$$

$$= F(t)\underline{x} + \underline{w} - F(t)\hat{\underline{x}} - K(t)[\underline{z} - H(t)\hat{\underline{x}}]$$

$$= [F(t) - K(t)H(t)]\tilde{\underline{x}} - K(t)\underline{v} + \underline{w} . \qquad (2.7)$$

The rate of change of the error covariance matrix is given by

$$\frac{d}{dt}(P) = E[\dot{\tilde{\underline{x}}}\tilde{\underline{x}}^T] + E[\tilde{\underline{x}}\dot{\tilde{\underline{x}}}^T] \qquad (2.8)$$

where

$$E[\dot{\tilde{\underline{x}}}\tilde{\underline{x}}^T] = E\{([F(t) - K(t)H(t)]\tilde{\underline{x}}\tilde{\underline{x}}^T - K(t)\underline{v}\tilde{\underline{x}}^T + \underline{w}\tilde{\underline{x}}^T\} .$$

But it follows from the assumptions about the noise processes that

$$E[\underline{v}(t)\tilde{\underline{x}}^T(t|t)] = -\tfrac{1}{2}R(t)K^T(t)$$

and

$$E[\underline{w}(t)\tilde{\underline{x}}^T(t|t)] = \tfrac{1}{2}Q(t) .$$

For example, consider the first of these two relations. The solution of Equation (2.7) has the form

$$\tilde{\underline{x}}(t|t) = \Psi(t,t_0)\tilde{\underline{x}}(t_0|t_0) - \int_{t_0}^{t}\Psi(t,\tau)K(\tau)\underline{v}(\tau)\,d\tau + \int_{t_0}^{t}\Psi(t,\tau)\underline{w}(\tau)\,d\tau .$$

Using this, one sees that

$$E[\underline{v}(t)\tilde{\underline{x}}^T(t|t)] = E[\underline{v}(t)\tilde{\underline{x}}(t_0|t_0)]\Psi^T(t,t_0) - \int_{t_0}^{t}E[\underline{v}(t)\underline{v}^T(\tau)]K^T(\tau)\Psi^T(t,\tau)\,d\tau + \int_{t_0}^{t}E[\underline{v}(t)\underline{w}^T(\tau)]\Psi^T(t,\tau)\,d\tau .$$

But

$$E[\underline{v}(t)\tilde{\underline{x}}(t_0|t_0)] = 0$$

and

$$E[\underline{v}(t)\underline{w}^T(\tau)] = 0 ,$$

from the basic assumptions, so that

$$E[\underline{v}(t)\tilde{\underline{x}}^T(t|t)] = -\int_{t_0}^{t}R(\tau)\delta(t-\tau)K^T(\tau)\Psi^T(t,\tau)\,d\tau$$

$$= -\tfrac{1}{2}R(t)K^T(t) .$$

As a result, the error covariance matrix is found to satisfy the differential equation.

$$\frac{d}{dt}P = [F(t) - K(t)H(t)]P + P[F(t) - K(t)H(t)]^T + K(t)R(t)K^T(t) + Q(t) . \qquad (2.9)$$

It is important to realize that Equation (2.9) describes the behavior of the error covariance matrix for *any* gain matrix $K(t)$. This relation is next used to derive the $K(t)$ that yields the minimum error variance. Also, note from the definition that $P(t|t)$ is symmetric.

Combine all terms that contain the unknown gain matrix $K(t)$. Then

$$\dot{P} = F(t)P + PF^T(t) + Q(t) + \{K(t)R(t)K^T(t) - K(t)H(t)P - PH^T(t)K^T(t)\} .$$

The terms outside the brackets cannot be affected directly by the choice of the gain matrix $K(t)$, so they will be ignored for the moment. The matrix $R(t)$ has been assumed to be symmetric and positive-definite so it can be factored into the product of a nonsingular matrix $S(t)$ and its transpose $S^T(t)$,

$$R(t) = S(t)S^T(t) .$$

Assume the existence of a matrix $A(t)$ such that

$$KRK^T - KHP - PH^TK^T = [KS - A][KS - A]^T - AA^T .$$

For this to be valid, choose

$$A = PH^TS^{-T}$$

where $[\,]^{-T}$ denotes the inverse of $[\,]^T$. With this choice for A, the differential equation for P becomes

$$\dot{P} = FP + PF^T + Q - PH^TR^{-1}HP + [K - PH^TR^{-1}]R[K - PH^TR^{-1}]^T . \tag{2.10}$$

Thus the gain matrix enters the quadratic term alone. For a prescribed initial value $P(t_0)$ and recognition of the fact that

$$\frac{d}{dt}(\text{trace } P) = \text{trace } \frac{d}{dt}(P) ,$$

it follows that (trace P) is minimized at each time by choosing $K(t)$ so that $\dot{P}$ is as small as possible. This is accomplished by choosing K so that the quadratic term in Equation (2.10) is eliminated. Thus, the optimal gain is

$$K(t) = P(t|t)H^T(t)R^{-1}(t) \tag{2.11}$$

and the error covariance matrix for the optimal gain is described by

$$\dot{P} = F(t)P + PF^T(t) - PH^T(t)R^{-1}(t)H(t)P + Q(t) , \tag{2.12}$$

with prescribed initial condition $P(t_0)$. The arguments leading to Equations (2.11) and (2.12) can be made more precise by using variational arguments, as discussed by Athans[17]. The error covariance matrix $P(t|t)$ is determined by solving Equation (2.12). This equation is a matrix Ricatti equation[18] and is discussed in more detail in Section 3.2.

This completes the heuristic derivation of the equations of the Kalman-Bucy filter.

Summary of Principal Results

The unbiased, minimum variance estimate of the linear system described by Equations (1.1) – (1.2) is given as the solution of the system

$$\dot{\hat{x}}(t|t) = F(t)\hat{x}(t|t) + K(t)[z(t) - H(t)\hat{x}(t|t)] , \tag{2.3}$$

where $\hat{x}(t_0|t_0)$ is selected so that

$$E[\hat{x}(t_0|t_0)] = E[x(t_0)] .$$

The optimal gain matrix is given by

$$K(t) = P(t|t)H^T(t)R^{-1}(t) \tag{2.11}$$

and the error covariance is obtained as the solution of the matrix Ricatti equation

$$\frac{d}{dt}P(t|t) = F(t)P(t|t) + P(t|t)F^T(t) + Q(t) - P(t|t)H^T(t)R^{-1}(t)H(t)P(t|t) . \tag{2.12}$$

2.1.2 *The Innovations Approach*

A more rigorous derivation of the solution of the linear, unbiased, minimum variance filtering problem is presented in this section. The approach follows that of Kailath[16] and has a strong similarity to the derivation presented by Kalman and Bucy in their original paper[12].

To begin, the properties of the residual $r(t)$ defined in Equation (2.2) will be examined in more detail. Kailath refers to this process as the *innovations process* because the residual contains in essence the "new" information contained in the measurement data $z(t)$.

The process

$$r(t) = z(t) - H(t)\hat{x}(t|t)$$

is a white-noise process with the same statistics as the measurement noise process $v(t)$ *when* $\hat{x}(t|t)$ *is the minimum variance estimate of* $x(t)$ *using* $Z(t)$.

The plausibility of this statement can be seen by direct calculation. First, verify that $\underline{r}(t)$ is a white-noise process.

Let $$\underline{y}(t) \triangleq H(t)\underline{z}(t) , \tag{2.13}$$

Then $$\hat{\underline{y}}(t|t) = H(t)\hat{\underline{z}}(t|t) \tag{2.14}$$

and $$\tilde{\underline{y}}(t|t) = H(t)\tilde{\underline{z}}(t|t) ; \tag{2.15}$$

so $$\begin{aligned} \underline{r}(t) &= \underline{z}(t) - \hat{\underline{y}}(t|t) \\ &= H(t)\tilde{\underline{z}}(t|t) + \underline{v}(t) . \end{aligned} \tag{2.16}$$

Suppose $t > s$. Using Equation (2.16), form

$$E[\underline{r}(t)\underline{r}^T(s)] = E[\tilde{\underline{y}}(t|t)\tilde{\underline{y}}^T(s|s)] + E[\underline{v}(t)\tilde{\underline{y}}^T(s|s)] + E[\tilde{\underline{y}}(t|t)\underline{v}^T(s)] + E[\underline{v}(t)\underline{v}^T(s)] .$$

Since $\underline{v}(t)$ is a white-noise process,

$$E[\underline{v}(t)\tilde{\underline{y}}^T(s)] = 0$$

and $$E[\underline{v}(t)\underline{y}^T(s)] = 0 .$$

Consequently

$$\begin{aligned} E[\underline{r}(t)\underline{r}^T(s)] &= E(\tilde{\underline{y}}(t|t)[\tilde{\underline{y}}(s|s) + \underline{v}(s)]^T) \\ &= E(\tilde{\underline{y}}(t|t)[\underline{z}(s) - \hat{\underline{y}}(s)]^T) . \end{aligned}$$

But it is well-known that the error $\tilde{\underline{y}}(t|t)$ for the minimum variance estimate must be orthogonal to the measurements $\underline{Z}(s)$ in the sense that

$$E[\tilde{\underline{y}}^T(t|t)\underline{z}(s)] = 0 , \qquad s < t . \tag{2.17}$$

This property is proven in Section 2.1.3 and is essentially the Wiener-Hopf equation.

Assuming that $\hat{\underline{y}}(t|t)$ is a linear function of the measurement data $\underline{Z}(t)$ and is the minimum variance estimate, it follows from Equation (2.17) that

$$E[\underline{r}(t)\underline{r}^T(s)] = 0 \qquad \text{for} \quad t > s .$$

A similar argument for $t < s$ shows that

$$E[\underline{r}(t)\underline{r}^T(s)] = 0 \qquad \text{for} \quad t < s ,$$

so it only remains to consider $t = s$. It is necessary to demonstrate that $E[\underline{r}(t)\underline{r}^T(t)]$ is infinite in order to conclude that $\underline{r}(t)$ is a white-noise process.

$$E[\underline{r}(t)\underline{r}^T(t)] = E[\underline{v}(t)\underline{r}^T(t)] + E[\tilde{\underline{y}}(t|t)\tilde{\underline{y}}^T(t|t)] + E[\underline{v}(t)\tilde{\underline{y}}^T(t|t)] + E[\tilde{\underline{y}}(t|t)\underline{v}^T(t)] .$$

But $\tilde{\underline{y}}(t|t)$ is defined as the error in the estimate of $\underline{y}(t)$ for the data $\underline{z}(s)$, $t_0 \leq s < t$ and $\underline{v}(t)$ is white-noise, so

$$E[\tilde{\underline{y}}(t|t)\underline{y}^T(t)] = 0 = E[\underline{v}(t)\tilde{\underline{y}}^T(t|t)]$$

and $$E[\underline{r}(t)\underline{r}^T(t)] = R(t)\delta(t-t) + E[\tilde{\underline{y}}(t|t)\tilde{\underline{y}}^T(t|t)] .$$

But the covariance $E[\tilde{\underline{y}}(t|t)\tilde{\underline{y}}^T(t|t)]$ is finite so it is negligible by comparison with $R(t)\delta(t-t)$ and one concludes that

$$E[\underline{r}(t)\underline{r}^T(s)] = R(t)\delta(t-s) .$$

The residual $\underline{r}(t)$ represents the difference between the new measurement $\underline{z}(t)$ and the prediction based on all previous data $\hat{\underline{y}}(t|t)$. Thus $\underline{r}(t)$ contains the information being contributed by the $\underline{z}(t)$, so $\underline{r}(t)$ and $\underline{z}(t)$ can be used interchangeably. Suppose that the optimal estimate is a linear function of the innovations process and has the form

$$\hat{\underline{x}}(t|t) = \int_{t_0}^{t} W(t,s)\underline{r}(s)\,ds \,, \tag{2.18}$$

where $W(t,s)$ is to be chosen so that Equation (2.17) is satisfied. Form

$$E[\hat{\underline{x}}(t|t)\underline{r}^T(\sigma)] = \int_{t_0}^{t} W(t,s)E[\underline{r}(s)\underline{r}^T(\sigma)]\,ds \,.$$

Using the orthogonality of the error in the estimate and the measurement data, one obtains

$$E[\hat{\underline{x}}(t|t)\underline{r}^T(\sigma)] = E[\underline{x}(t)\underline{r}^T(\sigma)] \,.$$

From the white-noise property of $\underline{r}(s)$, it follows that

$$E[\underline{x}(t)\underline{r}^T(\sigma)] = W(t,\sigma)R(\sigma) \,, \qquad t_0 \leq \sigma < t \,. \tag{2.19}$$

Since $R(\sigma)$ is positive-definite, the weighting matrix can be determined and the estimate is given by

$$\hat{\underline{x}}(t|t) = \int_{t_0}^{t} E[\underline{x}(t)\underline{r}^T(s)]R^{-1}(s)\underline{r}(s)\,ds \,. \tag{2.20}$$

The Kalman-Bucy filter equations presented in the preceding section are obtained from Equation (2.20) by differentiating with respect to t and by using Equation (1.1).

Differentiate Equation (2.20) to obtain

$$\frac{d}{dt}\hat{\underline{x}}(t|t) = E[\underline{x}(t)\underline{r}^T(t)]R^{-1}(t)\underline{r}(t) + \int_{t_0}^{t}\frac{d}{dt}E[\underline{x}(t)\underline{r}^T(s)]R^{-1}(s)\underline{r}(s)\,ds$$

$$= E[\underline{x}(t)\underline{r}^T(t)]R^{-1}(t)\underline{r}(t) + F(t)\int_{t_0}^{t} E[\underline{x}(t)\underline{r}^T(s)]R^{-1}(s)\underline{r}(s)\,ds + \int_{t_0}^{t} E[\underline{w}(t)\underline{r}^T(s)]R^{-1}(s)\underline{r}(s)\,ds \,.$$

But the second term is equal to $F(t)\hat{\underline{x}}(t|t)$ and the plant noise and residual are independent, so this reduces to

$$\dot{\hat{\underline{x}}}(t|t) = F(t)\hat{\underline{x}}(t|t) + E[\underline{x}(t)\underline{r}^T(t)]R^{-1}(t)\underline{r}(t) \,.$$

Letting

$$K(t) \triangleq E[\underline{x}(t)\underline{r}^T(t)]R^{-1}(t) \,, \tag{2.21}$$

the differential equation assumed in Section 2.1.1 is obtained:

$$\dot{\hat{\underline{x}}}(t|t) = F(t)\hat{\underline{x}}(t|t) + K(t)[\underline{z}(t) - H(t)\hat{\underline{x}}(t|t)] \,. \tag{2.3}$$

Consider $E[\underline{x}(t)\underline{r}^T(t)]$ and observe that

$$E[\underline{x}(t)\underline{r}^T(t)] \quad E\{\underline{x}(t)[H(t)\hat{\underline{x}}(t|t) + \underline{v}(t)]^T\}$$

$$= E\{[\underline{x}(t|t) + \hat{\underline{x}}(t|t)]\underline{x}^T(t|t))H^T(t)$$

$$= P(t|t)H^T(t) \,, \tag{2.22}$$

since the estimate and error are orthogonal and $P(t|t) \triangleq E[\underline{x}(t|t)\underline{x}^T(t|t)]$.

The error covariance matrix is specified by deriving the matrix differential equation that it must satisfy. This is accomplished by differentiating $P(t|t)$ and using the gain established by Equations (2.21) and (2.22). These operations are similar to the discussion of Section 2.1.1 and, as expected, yield the result obtained there. More specifically, $P(t|t)$ is the solution of the matrix Riccati equation

$$\dot{P} = F(t)P + PF^T(t) - PH^T(t)R^{-1}(t)H(t)P + Q(t) \,, \tag{2.12}$$

where it has been shown, using Equations (2.21) and (2.22), that

$$K(t) = P(t|t)H^T(t)R^{-1}(t) \,. \tag{2.11}$$

This completes the derivation of the Kalman-Bucy filter by the innovations approach.

2.1.3 *Other Considerations*

There are many other aspects that could be considered in conjunction with the problem discussed in the preceding sections. Some of these will be treated in this section. In Section 2.1.2 the orthogonality of the optimal linear estimate and the associated error was used without proof and this will now be supplied. Also, the generalization to the case in which the plant and measurement noise processes are correlated is discussed, as is the effect of a deterministic forcing function in the plant model (e.g., arising from control system considerations).

2.1.3.1 *Orthogonal projections and the Wiener-Hopf equation*

The orthogonality condition (Eqn (2.17)) used in Section 2.1.2 is a special case of the geometric property of orthogonal projection. Since the estimate is a linear function of the measurement data, it is contained in the linear subspace spanned by those data. If the state vector is not contained in this subspace, it is clear that the error in the estimate will not vanish. In fact, the error will have its smallest magnitude when the estimate is taken as the orthogonal projection of the state on the subspace spanned by the measurement data. This idea will be made more precise below and then applied to derive the well-known Wiener-Hopf equation and the orthogonality conditions used in the preceding section.

Consider a linear space X such that an inner product $\langle \underline{x}, \underline{y} \rangle$ is defined for any two elements $\underline{x}, \underline{y}$ in X. Define a norm by

$$\|\underline{x}\| = \langle \underline{x}, \underline{x} \rangle^{\frac{1}{2}} .$$

Let M be a subspace of X and consider the problem of finding a vector $\hat{\underline{y}}$ in M which minimizes $\|\underline{x} - \underline{y}\|$ with respect to any $\underline{y} \in M$. The solution of this problem, if it exists, is given by the following result.

Orthogonal Projection Lemma: $\|\underline{x} - \hat{\underline{y}}\|$ is a minimum for all $\underline{y} \in M$,

$$\|\underline{x} - \underline{y}\| \geq \|\underline{x} - \hat{\underline{y}}\| \quad \text{for all } \underline{y} \in M ,$$

if and only if $(\underline{x} - \hat{\underline{y}})$ is orthogonal to all $\underline{y} \in M$

$$\langle \underline{x} - \hat{\underline{y}}, \underline{y} \rangle = 0 \quad \text{for all } \underline{y} \in M . \tag{2.23}$$

Thus, the $\underline{y} \in M$ can be regarded as the linear combination of the elements spanning M (e.g. the measurement data). This estimate yields an error that is orthogonal to all elements $\underline{y} \in M$.

The proof of this result is straightforward. Assume that Equation (2.23) is valid. Then, for any $\underline{y} \in M$,

$$\begin{aligned} \|\underline{x} - \underline{y}\|^2 &= \|(\underline{x} - \hat{\underline{y}}) + (\hat{\underline{y}} - \underline{y})\|^2 \\ &= \|\underline{x} - \hat{\underline{y}}\|^2 + 2\langle \underline{x} - \hat{\underline{y}}, \hat{\underline{y}} - \underline{y} \rangle + \|\hat{\underline{y}} - \underline{y}\|^2 . \end{aligned}$$

But $(\hat{\underline{y}} - \underline{y}) \in M$; so, by Equation (2.23), the middle term vanishes and

$$\begin{aligned} \|\underline{x} - \underline{y}\|^2 &= \|\underline{x} - \hat{\underline{y}}\|^2 + \|\hat{\underline{y}} - \underline{y}\|^2 \\ &\geq \|\underline{x} - \hat{\underline{y}}\|^2 \end{aligned}$$

with equality if and only if $\underline{y} = \hat{\underline{y}}$.

To complete the proof, assume that $\hat{\underline{y}}$ minimizes $\|\underline{x} - \underline{y}\|$ for all $\underline{y} \in M$ and assume that there exists some $\underline{y}_1 \in M$ such that

$$\langle \underline{x} - \hat{\underline{y}}, \underline{y}_1 \rangle = \alpha \neq 0 .$$

Then $$\|\underline{x} - \hat{\underline{y}} - \beta \underline{y}_1\|^2 = \|\underline{x} - \hat{\underline{y}}\|^2 - 2\alpha\beta + \beta^2 \|\underline{y}_1\|^2 .$$

But, by appropriate choice of β, it is possible to make the last two terms negative, thereby contradicting the minimality of $\hat{\underline{y}}$.

The orthogonal projection lemma can be used to derive the Wiener-Hopf equation and associated orthogonality conditions for the estimation problem described in Section 1.2. Let X be the n-dimensional state space and define the inner product of two elements in X to be

$$\langle \underline{x}, \underline{y} \rangle \triangleq E[\underline{x}^T \underline{y}] , \quad \text{for all } \underline{x}, \underline{y} \in X .$$

For convenience, it will be assumed that the variables all have zero mean. The norm of $\underline{x} \in X$ is

$$\|\underline{x}\|^2 = \langle \underline{x}, \underline{x} \rangle = E[\underline{x}^T \underline{x}] = \text{trace } E[\underline{x}\underline{x}^T] .$$

From the orthogonal projection lemma, the trace of the error covariance matrix is minimized if

$$(\underline{x}(t) - \hat{\underline{x}}(t|t),\ \underline{z}(s)) \;=\; E[\tilde{\underline{x}}^T(t|t)\underline{z}(s)] \;=\; 0\ , \qquad \text{for}\quad t_0 \le s < t\ . \tag{2.24}$$

Similarly, this lemma is used to establish Equation (2.17) and the other orthogonality conditions used in Section 2.1.2.

To derive the Wiener-Hopf equation for the nonstationary linear system of this presentation, suppose that an estimate of the state $\underline{x}(t)$ is to be a linear function of the measurement data $\underline{Z}(t)$ having the form

$$\hat{\underline{x}}(t|t) \;=\; \int_{t_0}^{t} W(t,s)\underline{z}(s)\,ds\ . \tag{2.25}$$

Since the components of $\hat{\underline{x}}(t|t)$ can be chosen independently of each other, it suffices to chose $\hat{\underline{x}}(t|t)$ so that

$$E\{[\underline{x}(t) - \hat{\underline{x}}(t|t)]\underline{z}^T(s)\} \;=\; 0\ , \qquad t_0 \le s < t\ .$$

Using Equation (2.25), one obtains the Wiener-Hopf equation

$$E[\underline{x}(t)\underline{z}^T(s)] \;=\; \int_{t_0}^{t} W(t,\sigma)E[\underline{z}(\sigma)\underline{z}^T(s)]\,d\sigma \qquad \text{for}\quad t_0 \le s < t\ . \tag{2.26}$$

2.1.3.2 Correlation between plant and measurement noise processes

It has been assumed heretofore that

$$E[\underline{w}(t)\underline{v}^T(s)] \;=\; 0 \qquad \text{for all}\quad t\ ,\ s\ .$$

In this section, consider the more general case in which

$$E[\underline{w}(t)\underline{v}^T(s)] \;=\; C(t)\delta(t-s) \qquad \text{for all}\quad t\ ,\ s\ . \tag{2.27}$$

Relative to the development of Section 2.1.1, note that the effect of the correlation appears first in the derivation of Equation (2.9). One sees, in this case, that

$$E[\underline{v}(t)\underline{z}^T(t|t)] \;=\; \tfrac{1}{2}R(t)K^T(t) + \tfrac{1}{2}C^T(t)$$

and

$$E[\underline{w}(t)\underline{z}^T(t|t)] \;=\; \tfrac{1}{2}Q(t) - \tfrac{1}{2}C(t)K^T(t)\ .$$

Consequently, Equation (2.9) is found to be modified and the differential equation describing the error covariance matrix becomes

$$\begin{aligned}\frac{d}{dt}P \;&=\; [F(t) - K(t)H(t)]P + P[F(t) - K(t)H(t)]^T + Q(t) + K(t)R(t)K^T(t) - K(t)C^T(t) - C(t)K^T(t)\\ &=\; F(t)P + PF^T(t) + Q(t) + (K(t)R(t)K^T(t) - K(t)[H(t)P + C^T(t)] - [PH^T(t) + C(t)]K^T(t))\ .\end{aligned} \tag{2.28}$$

The optimal gain matrix is found to be

$$K(t) \;=\; [P(t|t)H^T(t) + C(t)]R^{-1}(t) \tag{2.29}$$

and the optimal error covariance is described by

$$\dot{P} \;=\; F(t)P + PF^T(t) - [PH^T(t) + C(t)]R^{-1}[C^T(t) + H(t)P] + Q(t)\ . \tag{2.30}$$

The same result is obtained using the innovations approach where Equation (2.22) is modified to account for the correlation.

2.1.3.3 Deterministic plant forcing function

Suppose that the plant model is altered to

$$\dot{\underline{x}} \;=\; F(t)\underline{x} + \underline{g}(t) + \underline{w}\ , \tag{2.31}$$

where $\underline{g}$ is a known function. The only modification required to the filter equations in this case is to note that the prediction must account for $\underline{g}(t)$. Thus, the estimate is given by

$$\dot{\hat{x}} = F(t)\hat{x} + d(t) + K(t)[z(t) - H(t)\hat{x}] , \tag{2.32}$$

where $K(t)$ and $P(t|t)$ are unchanged from Equations (2.11) and (2.12).

2.1.3.4 Conditional mean and minimum variance estimates

The minimum variance estimate has an interesting interpretation in terms of the conditional density of the state $x(t)$ given the measurement data $Z(\tau)$. This relation is described in the following lemma.

Lemma: Suppose that a random variable x is to be estimated from measurement data Z and suppose that x and Z have the joint probability density function $p(x,Z)$. An estimate $\hat{x}$ is to be determined from the data Z so that

$$E[(x - \hat{x})^T(x - \hat{x})] = \text{minimum} .$$

Then, the minimum variance estimate $\hat{x}$ is

$$\hat{x} = E[x|Z] . \tag{2.33}$$

Proof: Write $E[(x - \hat{x})^T(x - \hat{x})]$ in terms of the conditional density, using the identity

$$E[(x - \hat{x})^T(x - \hat{x})] = E(E[(x - \hat{x})^T(x - \hat{x})|Z]) .$$

But

$$E[(x - \hat{x})^T(x - \hat{x})|Z] = \hat{x}^T\hat{x} - 2\hat{x}^T E[x|Z] + E[x^T x|Z]$$

$$= (\hat{x} - E[x|Z])^T(\hat{x} - E[x|Z]) + E[x^T x|Z] - E^T[x|Z]E[x|Z] . \tag{2.34}$$

By definition this quantity is positive; so, to minimize $E[(x - \hat{x})^T(x - \hat{x})]$, it is sufficient to minimize Equation (2.34). Only the first term involves $\hat{x}$ and it is quadratic so that the smallest value it can assume is zero. This obtains when

$$\hat{x} = E[x|Z] ;$$

so Equation (2.33) is proven.

Note also that the conditional mean provides an unbiased estimate of x. This follows easily by recognizing that

$$E[\hat{x}] = E(E[x|Z]) = E[x] .$$

For the linear system (Eqns (1.1) - (1.2)) and the gaussian character of the initial state and the noise processes, it follows that $x(t)$ is gaussian for any time t. Further, it can be shown that the conditional density of the state $x(t)$, given the data $Z(t)$, is gaussian with mean value and covariance described by Equations (2.3), (2.11), and (2.12). Thus, the Kalman-Bucy estimate describes the behavior of the mean and covariance of the conditional density function. As a result of this fact and the lemma proved above, the linear Kalman-Bucy equations provide the best estimate, even when nonlinear estimators are considered. Probabilistic aspects of the estimation problem are considered in depth in Chapter 6.

2.1.4 The Time-Discrete Problem

The time-discrete linear filtering and prediction problem[11] is included for the sake of completeness and results are stated without proof. More details can be found in References 19 and 20 or Chapter 6 of this book. One can use arguments that are very similar to those used in the preceding sections and the results contain few surprises. The principal difference resides in the fact that the measurement noise covariance matrix does not have to be positive-definite, although a related matrix must have this property. No attempt will be made here to treat the time-discrete problem exhaustively. Most results for the time-continuous case can be modified without difficulty to apply to this problem.

Consider a system whose state is described by a linear difference equation

$$x_k = \Phi_{k,k-1}x_{k-1} + w_{k-1} , \qquad k = 1, 2, \ldots \tag{2.35}$$

and which is observed through measurement data z_k obtained at discrete instants of time t_k. These data are assumed to be linearly related to the state according to

$$z_k = H_k x_k + v_k . \tag{2.36}$$

The noise sequences are assumed to be gaussian and uncorrelated between sampling instants (i.e. white-noise sequences) with zero means and covariances:

$$E[\underline{w}_k \underline{w}_j^T] = Q_k \delta_{kj}$$

$$E[\underline{v}_k \underline{v}_j^T] = R_k \delta_{kj}$$

$$E[\underline{w}_k \underline{v}_j^T] = C_k \delta_{kj} .$$

The $\underline{w}_k$ and $\underline{v}_k$ are assumed to be uncorrelated with the initial state $\underline{x}_0$, which is a gaussian random variable with

$$E[\underline{x}_0] = \underline{a}_0$$

$$E[(\underline{x}_0 - \underline{a}_0)(\underline{x}_0 - \underline{a}_0)^T] = M_0 .$$

For this system one can prove that the unbiased, min. variance estimate of $\underline{x}_k$, given the data $\underline{Z}(k)$, is described by the following system. The estimate of the state $\underline{x}_k$ is

$$\hat{\underline{x}}_k' = \Phi_{k,k-1} \hat{\underline{x}}_{k-1} + A_{k-1}[\underline{z}_{k-1} - H_{k-1} \hat{\underline{x}}_{k-1}'] \tag{2.37}$$

before the measurement $\underline{z}_k$ is processed. The estimate when $\underline{z}_k$ is included is obtained by modifying the predicted estimate $\hat{\underline{x}}_k'$ according to

$$\hat{\underline{x}}_k = \hat{\underline{x}}_k' + K_k[\underline{z}_k - H_k \hat{\underline{x}}_k'] . \tag{2.38}$$

The optimal gains A_{k-1} and K_k are given by

$$A_{k-1} = C_{k-1}[H_{k-1} P_{k-1}' H_{k-1}^T + R_{k-1}]^{-1} \tag{2.39a}$$

$$K_k = P_k' H_k^T [H_k P_k' H_k^T + R_k]^{-1} , \tag{2.39b}$$

where the predicted error covariance matrix is described by

$$P_k' = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + Q_{k-1} - \Phi_{k,k-1} K_{k-1} C_{k-1}^T -$$

$$- C_{k-1} K_{k-1}^T \Phi_{k,k-1}^T - A_{k-1}[H_{k-1} P_{k-1}' H_{k-1}^T + R_{k-1}]^{-1} A_{k-1}^T \tag{2.40}$$

and the current error covariance matrix is

$$P_k = P_k' - K_k H_k P_k' . \tag{2.41}$$

When the plant and measurement noise are uncorrelated, the matrix C_k vanishes and causes A_k to vanish and the covariance matrices reduce to

$$P_k' = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + Q_{k-1} \tag{2.42}$$

and

$$P_k = P_k' - K_k H_k P_k' . \tag{2.43}$$

Note that the gain matrix requires the inversion of $[H_k P_k' H_k^T + R_k]$ instead of the measurement covariance matrix R_k, as occurs in the time-continuous case. However, one can derive a form for K_k that is more similar to Section 2.1.1 when P_k' and R_k are positive-definite by making use of the following lemma.

Matrix Inversion Lemma: Suppose the $n \times n$ matrix B and the $m \times m$ matrix R are positive-definite and let H be an arbitrary $m \times n$ matrix. If A is given by

$$A = B - BH^T[HBH^T + R]^{-1} HB , \tag{2.44}$$

then the inverse of A is

$$A^{-1} = B^{-1} + H^T R^{-1} H . \tag{2.45}$$

The proof follows by multiplication of A and A^{-1}.

The error covariance matrix can be rewritten, by substituting Equation (2.42) into Equation (2.43),

$$P_k = P_k' - P_k' H_k^T [H_k P_k' H_k^T + R_k]^{-1} H_k P_k' . \tag{2.46}$$

Using the lemma, one obtains

$$P_k^{-1} = (P_k')^{-1} + H_k^T R_k^{-1} H_k \quad . \tag{2.47}$$

The gain matrix can also be modified:

$$K_k = (P_k P_k^{-1}) P_k' H_k^T R_k^{-1} [H_k P_k' H_k^T R_k^{-1} + I]^{-1} \quad .$$

But, from Equation (2.43),

$$P_k^{-1} P_k' = I + H_k^T R_k^{-1} H_k P_k' \quad ,$$

so

$$K_k = P_k [I + H_k^T R_k^{-1} H_k P_k'] H_k^T R_k^{-1} [I + H_k P_k' H_k^T R_k^{-1}]^{-1}$$

$$= P_k H_k^T R_k^{-1} \quad . \tag{2.48}$$

Equations (2.47) and (2.48) provide an alternative form when P_k' and R_k are positive-definite. Equation (2.47) has the disadvantage that an $n \times n$ matrix must be inverted, rather than the $m \times m$ matrix that must be inverted in Equation (2.42) (usually, $m < n$).

2.2 The Colored-Noise Problem

2.2.1 *Colored-Noise, Shaping Filters, and State Vector Augmentation*

It has been assumed that the plant and measurement noise processes are gaussian and white. This assumption is not always satisfied in practice, so it is desirable to consider linear systems in which the noise exhibits correlation between different instants of time (i.e. the noise is "colored"). This problem is successfully approached when the noise can be described by a shaping filter[21].

Definition: Consider a gaussian random process $\underline{n}(t)$ that has zero mean and whose second-order correlations are given by

$$E[\underline{n}(t)\underline{n}^T(\tau)] = D(t,\tau) \quad . \tag{2.49}$$

A linear dynamical system driven by a gaussian white-noise process whose output has the same statistical characteristics as $\underline{n}(t)$ is called a shaping filter.

Thus, by introducing a shaping filter, many random processes $\underline{n}(t)$ can be described by

$$\dot{\underline{n}} = A(t)\underline{n} + \underline{w} \quad , \tag{2.50}$$

where $A(t)$ and the statistics of the white-noise $\underline{w}(t)$ are chosen so that $\underline{n}(t)$ has the prescribed statistics. The problem of determining shaping filters for random processes will not be treated here, but the filter development for those processes describable by Equation (2.50) will be given below.

Consider the following linear system (analogous to Equations (1.1) and (1.2)), in which colored noise exists in the plant and measurement systems. To distinguish between white and colored-noise components, consider the following partitioned system.

$$\begin{bmatrix} \dot{\underline{x}}_1 \\ \dot{\underline{x}}_2 \end{bmatrix} = \begin{bmatrix} F_1(t) & F_2(t) \\ F_3(t) & F_4(t) \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} + \begin{bmatrix} \underline{w}_1 \\ \underline{\xi}_1 \end{bmatrix} \quad , \tag{2.51}$$

where $\underline{w}_1$ is a white-noise process

$\underline{\xi}_1$ is a colored-noise process

where

$$\dot{\underline{\xi}}_1 = A(t)\underline{\xi}_1 + \underline{w}_2 \tag{2.52}$$

and $\underline{w}_2$ is a white-noise process.

The measurement data are described by

$$\begin{bmatrix} \underline{z}_1 \\ \underline{z}_2 \end{bmatrix} = \begin{bmatrix} H_1(t) & H_2(t) \\ H_3(t) & H_4(t) \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} + \begin{bmatrix} \underline{v}_1 \\ \underline{\xi}_2 \end{bmatrix} \quad , \tag{2.53}$$

where $\underline{x}_1$ is a white-noise process

$\underline{\xi}_2$ is a colored-noise process

where
$$\dot{\underline{\xi}}_2 = B(t)\underline{\xi}_2 + \underline{w}_3 \tag{2.54}$$

and $\underline{w}_3$ is a white-noise process.

The problem described by Equations (2.51) - (2.54) can be reformulated by state vector augmentation to obtain a system in which the state variables $\underline{x}_1$ and $\underline{x}_2$ are combined with the colored-noise variables $\underline{\xi}_1$ and $\underline{\xi}_2$ to define a system in which only white-noise appears explicitly. This system is

$$\begin{bmatrix} \dot{\underline{x}}_1 \\ \dot{\underline{x}}_2 \\ \dot{\underline{\xi}}_1 \\ \dot{\underline{\xi}}_2 \end{bmatrix} = \begin{bmatrix} F_1(t) & F_2(t) & 0 & 0 \\ F_3(t) & F_4(t) & I & 0 \\ 0 & 0 & A(t) & 0 \\ 0 & 0 & 0 & B(t) \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \underline{\xi}_1 \\ \underline{\xi}_2 \end{bmatrix} + \begin{bmatrix} \underline{w}_1 \\ 0 \\ \underline{w}_2 \\ \underline{w}_3 \end{bmatrix}, \tag{2.55a}$$

with measurement data

$$\begin{bmatrix} \underline{z}_1 \\ \underline{z}_2 \end{bmatrix} \quad \begin{bmatrix} H_1(t) & H_2(t) & 0 & 0 \\ H_3(t) & H_4(t) & 0 & I \end{bmatrix} \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \\ \underline{\xi}_1 \\ \underline{\xi}_2 \end{bmatrix} \quad \begin{bmatrix} \underline{v}_1 \\ 0 \end{bmatrix}, \tag{2.56a}$$

or, more succinctly,

$$\dot{\underline{X}} = F(t)\underline{X} + \underline{W} \tag{2.55b}$$

$$\underline{Z} = H(t)\underline{X} + \underline{V}, \tag{2.56b}$$

where
$$\underline{X}^T \triangleq [\underline{x}_1^T, \underline{x}_2^T, \underline{\xi}_1^T, \underline{\xi}_2^T]$$

with similar definitions for $F(t)$, $H(t)$, $\underline{W}$, and $\underline{V}$. Thus, it appears that the problem has been reformulated to the model given initially by Equations (1.1) and (1.2), so that the solutions already obtained can be applied. Unfortunately, this is not true because the zero components appearing in the measurement noise vector $\underline{V}$ prevent the covariance matrix $R(t)$ from being positive-definite. This property is required in order to form the optimal gain matrix $K(t)$, since $R^{-1}(t)$ is required. Thus, shaping filters and state vector augmentation permit the plant equation to be rewritten in a manner that is compatible with our previous results (i.e. the plant covariance matrix Q does *not* have to be inverted) but additional consideration must be given to colored measurement noise.

2.2.2 *Solution for Colored Measurement Noise*

The components of the measurement vector which contain only colored-noise provide the source of the difficulty in applying the filter equations developed earlier. This problem was first discussed by Cox[22] and Bryson and Johansen[23] and has more recently been considered by Bucy[24], Stear and Stubberud[25], and Sarachik[26]. To develop a solution for this case, assume that the model has the form

$$\dot{\underline{X}} = F(t)\underline{X} + \underline{W} \tag{2.57}$$

and
$$\begin{bmatrix} \underline{z}_1 \\ \underline{z}_2 \end{bmatrix} = \begin{bmatrix} H_1(t) \\ H_2(t) \end{bmatrix} \underline{X} + \begin{bmatrix} \underline{v} \\ 0 \end{bmatrix}, \tag{2.58}$$

where $\underline{X}$ is an $(n+p)$ dimensional state vector

$\underline{z}_1$ is an m-dimensional measurement vector containing additive white-noise

$\underline{z}_2$ is a p-dimensional noise-free measurement vector.

This model is motivated by Equations (2.55) and (2.56), although $F(t)$, $H_1(t)$, and $H_2(t)$ will not be required to have the form given there. The white-noise $\underline{v}$ is assumed to have a positive-definite covariance matrix $R(t)$ and the H_1 and H_2 are assumed to have maximal rank.

There are several aspects of this problem that require mention. First observe that one might solve the problem formally by differentiating the measurement data $\underline{z}_2$:

$$\dot{\underline{z}}_2 = H_2\dot{\underline{x}} + \dot{H}_2\underline{x}$$

$$= (H_2F + \dot{H}_2)\underline{x} + H_2\underline{w} . \tag{2.59}$$

If $H_2QH_2^T$ (i.e. the covariance matrix of the noise vector $H_2\underline{w}$) is positive-definite, then $\dot{\underline{z}}_2$ can be treated as the measurement data in place of $\underline{z}_2$ since they are equivalent if the initial conditions on $\underline{x}$ are selected appropriately. Now the Kalman-Bucy results can be applied using the measurement data $\underline{z}_1$ and $\dot{\underline{z}}_2$. If $H_2QH_2^T$ has rank $r < p$, then one can transform the $\dot{\underline{z}}_2$ into a vector that is separated into components either with or without white-noise. The noise-free components can be differentiated again in an attempt to introduce white-noise. The procedure of separation and differentiation can be repeated until a set of p measurements containing white-noise are generated to replace $\underline{z}_2$. If it proves to be impossible to introduce white-noise into every component, then these "perfect" measurements can be dealt with as described by Bryson and Johansen[23].

Note, however, that the noise-free measurements in Equation (2.58) provide perfect knowledge of p variables, so that one would expect that the filter equations would only involve n variables instead of the $(n+p)$ appearing in the state vector of Equation (2.57). This permits a desirable relaxation of computational requirements. Furthermore, it is undesirable in practice to differentiate data, so these two disadvantages motivate one to develop other means for dealing with colored measurement noise.

Using the definition of $\underline{z}_2$ and the assumption that H_2 has maximal rank, it is possible to define an $(n-p)\times n$ matrix H_3 such that a nonsingular matrix $T(t)$, with $t \geq t_0$, can be formed.

$$T(t) \triangleq \begin{bmatrix} H_2(t) \\ H_3(t) \end{bmatrix} . \tag{2.60}$$

The H_3 must be selected so that

$$H_2H_3^T = 0$$

and

$$H_3H_3^T = I .$$

The task of choosing H_3 is not as difficult as it may at first appear; at least it is trivially accomplished for an important class of systems that are discussed later in this section.

Let the inverse of T be defined as

$$T^{-1}(t) = [J_2(t)\, J_3(t)] , \tag{2.61}$$

where it can easily be shown that

$$J_2 = H_2^T(H_2H_2^T)^{-1}$$

$$J_3 = H_3^T$$

when the constraints on H_3 are invoked.

Define new state variables with subvectors $\underline{z}_2$ and $\underline{\xi}$ such that

$$\begin{bmatrix} \underline{z}_2 \\ \underline{\xi} \end{bmatrix} \triangleq \begin{bmatrix} H_2 \\ H_3 \end{bmatrix} \underline{x} = T\underline{x} . \tag{2.62}$$

Since T has an inverse, the $\underline{x}$ can be recovered,

$$\underline{x} = J_2\underline{z}_2 + J_3\underline{\xi} . \tag{2.63}$$

But $\underline{z}_2$ is known measurement data, so the problem of obtaining the minimum variance estimate of $\underline{x}$ reduces to that of estimating $\underline{\xi}$. As a result, the order of the filter that is required reduces, as expected, from $(n+p)$ to n . After the estimate of $\underline{\xi}$ is determined, the estimate of the state $\underline{x}$ is obtained from

$$\hat{\underline{x}}(t|t) = J_2(t)\underline{z}_2(t) + J_3(t)\hat{\underline{\xi}}(t|t) . \tag{2.64}$$

Consider the problem of estimating the vector $\underline{\xi}$. From Equation (2.63), it is seen that

$$\underline{\xi} = H_3\underline{x} ,$$

so a differential equation describing $\underline{\xi}$ is obtained by differentiating and substituting Equation (2.57):

$$\begin{aligned}\dot{\underline{\xi}} &= \dot{H}_3\underline{x} + H_3\dot{\underline{x}} \\ &= (\dot{H}_3 + H_3F)\underline{x} + H_3\underline{w} \\ &= [(\dot{H}_3 + H_3F)J_3]\underline{\xi} + (\dot{H}_3 + H_3F)J_2\underline{z}_2 + H_3\underline{w} . \end{aligned} \tag{2.65}$$

The term $(\dot{H}_3 + H_3F)J_2\underline{z}_2$ is a known forcing function so it can be treated in the fashion described in Section 2.1.3.3.

The measurement data $\underline{z}_1$ can also be expressed in terms of $\underline{\xi}$:

$$\underline{z}_1 = H_1J_2\underline{z}_2 + H_1J_3\underline{\xi} + \underline{v} .$$

Let
$$\underline{z} \triangleq \underline{z}_1 - H_1J_2\underline{z}_2 ; \tag{2.66}$$

so
$$\underline{z} = [H_1J_3]\underline{\xi} + \underline{v} . \tag{2.67}$$

Equations (2.65), (2.66), and (2.67) form a system to which the previously derived filter equations can be applied, since $\underline{v}$ is assumed to have a positive-definite covariance matrix. One point remains to be defined. Ordinarily, statistics are prescribed for $\underline{x}(t_0)$, but in this application it is necessary to specify statistics for $\underline{\xi}(t_0)$. To accomplish this, e that the measurement $\underline{z}(t_0)$ is known; so it can be used to estimate $\underline{x}(t_0)$ and, through Equation (2.6. to obtain initial conditions for $\underline{\xi}(t_0)$. Using the results for time-discrete filtering presented in Sec ion 2.1.4, one sees that the statistics of $\underline{\xi}(t_0)$, based on the *a priori* statistics for $\underline{x}(t_0)$ and the noise-free measurement data $\underline{z}_2(t_0)$, are

$$\begin{aligned}\hat{\underline{\xi}}(t_0) &= H_3(t_0)\hat{\underline{x}}(t_0|t_0) \\ &= H_3(t_0)P(t_0)H_2^T(t_0)[H_2(t_0)P(t_0)H_2^T(t_0)]^{-1}\underline{z}_2(t_0) , \end{aligned} \tag{2.68}$$

with error covariance matrix

$$\Pi(t_0) = H_3(t_0)[P(t_0) - K(t_0)H_2(t_0)P(t_0)]H_3^T(t_0) , \tag{2.69}$$

where the gain matrix is given by

$$K(t_0) = P(t_0)H_2^T(t_0)[H_2(t_0)P(t_0)H_2^T(t_0)]^{-1} . \tag{2.70}$$

Naturally, it is necessary to assume that the inverse of $H_2(t_0)P(t_0)H_2^T(t_0)$ exists for these relations to be valid. Since H_2 has already been assumed to have maximal rank, it is sufficient to require that $P(t_0)$ be positive-definite, although this is not a necessary condition.

Using the initial conditions (Eqns (2.68) and (2.69)), the minimum variance estimate of $\underline{\xi}(t)$ based on the data $(\underline{z}_1 - H_1J_2\underline{z}_2)$ is obtained by applying Equations (2.32), (2.11) and (2.12) (or Equations (2.32), (2.29) and (2.30) if the plant and measurement noises are correlated). There are no differentiations of measurement data required in this approach and the order of the filter has been reduced from $(n+p)$ to n. Of course, this solution requires that there exist some measurements containing white-noise. If the vector $\underline{z}_1$ in Equation (2.58) does not exist, additional manipulations must be introduced, as now discussed.

Suppose that the plant is described by Equation (2.57) but that the measurements are entirely noise-free (i.e. there is no white-noise in the data) and described by

$$\underline{z}_2 = H_2(t)\underline{x} , \tag{2.71}$$

where the subscript 2 has been retained to be consistent with Equation (2.58). To obtain a form that can be treated by the above procedure, differentiate $\underline{z}_2$ to obtain

$$\begin{aligned}\dot{\underline{z}}_2 &= H_2\dot{\underline{x}} + \dot{H}_2\underline{x} \\ &= (H_2F + \dot{H}_2)\underline{x} + H_2\underline{w} . \end{aligned} \tag{2.72}$$

To simplify the discussion, it shall be assumed that the covariance $H_2QH_2^T$ of the noise $H_2\underline{w}$ is positive-definite. The differentiated data $\dot{\underline{z}}_2$ can now be treated as additional measurement data, so that the complete system has the form

$$\dot{\underline{x}} = F(t)\underline{x} + \underline{w}$$

$$\underline{z}_1 \triangleq \dot{\underline{z}}_2 = [H_2(t)F(t) + \dot{H}_2(t)]\underline{x} + H_2(t)\underline{w}$$

$$\underline{z}_2 = H_2(t)\underline{x} ,$$

and now the preceding results can be applied without additional modification.

From Equation (2.66), the measurement data used in the filter will be

$$\underline{z} = \dot{\underline{z}}_2 - [H_2F + \dot{H}_2]J_2\underline{z}_2 , \tag{2.73}$$

where J_2 is defined in Equation (2.61). By Equation (2.67), the $\underline{z}$ is represented in terms of $\underline{\xi}$ as

$$\underline{z} = [H_2F + \dot{H}_2]J_2\underline{\xi} + H_2\underline{w} . \tag{2.74}$$

The filter state $\underline{\xi}$ is still described by Equation (2.65). The plant and measurement noises are correlated with covariance matrix

$$Q(t)\delta(t-\tau) \triangleq E[H_2(t)\underline{w}(t)\underline{w}^T(\tau)H_2^T(\tau)]$$

$$Q(t)\delta(t-\tau) = H_2(t)H_2^T(t)\delta(t-\tau) , \tag{2.75}$$

The initial conditions are still described by Equations (2.68) and (2.69) and Equations (2.32), (2.29) and (2.30) must be used for the filtering. Note also that the plant and measurement noise covariances for this system are $H_2QH_2^T$ and $H_2QH_2^T$, respectively.

From Equation (2.33), the estimate of $\underline{\xi}$ is given by

$$\dot{\hat{\underline{\xi}}} = GJ_2\hat{\underline{\xi}} + GJ_2\underline{z}_2 + K[\underline{z} - JJ_2\hat{\underline{\xi}}] , \tag{2.76}$$

where

$$G \triangleq \dot{H}_2 + H_2F$$

$$J \triangleq \dot{H}_2 + H_2F .$$

The $\underline{z}$ given by Equation (2.74) appears to indicate that a differentiation $\dot{\underline{z}}_2$ is required in order to obtain an estimate of $\underline{\xi}$. This requirement can be circumvented computationally by resorting to the following artifice. Define

$$\underline{\xi}^* \triangleq \hat{\underline{\xi}} - K(t)\underline{z}_2 ,$$

so that

$$\dot{\underline{\xi}}^* = \dot{\hat{\underline{\xi}}} - K\dot{\underline{z}}_2 - \dot{K}\underline{z}_2$$

$$= G_2J_2\hat{\underline{\xi}} + GJ_2\underline{z}_2 + K[\underline{z} - JJ_2\hat{\underline{\xi}}] - K\dot{\underline{z}}_2 - \dot{K}\underline{z}_2 .$$

Using Equation (2.73), this reduces to

$$\dot{\underline{\xi}}^* = G_2J_2\hat{\underline{\xi}} - KJ[J_2\underline{z}_2 + J_2\hat{\underline{\xi}}] + [GJ_2 - \dot{K}]\underline{z}_2 \tag{2.77}$$

and does not require the differentiation of data. The gain matrix K must be differentiated, however.

There is a special case[24,25] that is sufficiently important that it will be considered here and will complete the discussion of colored measurement noise. In the following the notation will be changed somewhat from that used elsewhere in this section, in order to deal not with the augmented state implied by Equation (2.57) but with a vector that is more closely akin to the state of the physical system. Let the system be described by

$$\dot{\underline{x}} = F(t)\underline{x} + \underline{w} \tag{2.78}$$

$$\underline{z} = H(t)\underline{x} + \underline{n} , \tag{2.79}$$

where $\underline{n}$ is a zero mean colored-noise process described by a shaping filter,

$$\dot{\underline{n}} = A(t)\underline{n} + \underline{v} , \tag{2.80}$$

and with initial covariance $E[\underline{n}(t_0)\underline{n}^T(t_0)] = N$ where $\underline{v}$ has a positive-definite covariance matrix. Assume, also, that $\underline{v}$ and $\underline{w}$ are uncorrelated. This is essentially the problem considered immediately above, except that $\underline{x}$ is an n-dimensional state unaugmented by measurement noise variables. To put this system in a form compatible with Equations (2.57) and (2.71), rewrite Equations (2.78) - (2.80) as

$$\begin{bmatrix} \dot{\underline{x}} \\ \dot{\underline{n}} \end{bmatrix} = \begin{bmatrix} F(t) & 0 \\ 0 & A(t) \end{bmatrix} \begin{bmatrix} \underline{x} \\ \underline{n} \end{bmatrix} + \begin{bmatrix} \underline{w} \\ \underline{v} \end{bmatrix} , \tag{2.81}$$

where

$$E\left\{ \begin{bmatrix} \underline{w} \\ \underline{v} \end{bmatrix} \begin{bmatrix} \underline{w}^T & \underline{v}^T \end{bmatrix} \right\} = \begin{bmatrix} Q & C \\ C^T & R \end{bmatrix} \delta(t-\tau)$$

and

$$\underline{z} = [H(t)\ I] \begin{bmatrix} \underline{x} \\ \underline{n} \end{bmatrix} . \tag{2.82}$$

For this problem the definition of the matrix H_1 required in the transformation T of Equation (2.60) can be defined explicitly as

$$H_1 = [I \quad 0] ,$$

so that transformation T is

$$T(t) = \begin{bmatrix} H & I \\ I & 0 \end{bmatrix} \triangleq \begin{bmatrix} H_2 \\ H_1 \end{bmatrix} \tag{2.83}$$

and

$$T^{-1}(t) = \begin{bmatrix} 0 & I \\ I & -H \end{bmatrix} \triangleq \begin{bmatrix} J_2 & J_1 \end{bmatrix} . \tag{2.84}$$

With this transformation, the filtering state $\underline{\xi}$ is

$$\underline{\xi} = \underline{x} ,$$

so that $\underline{x}$ is to be estimated directly. Obviously $\underline{x}$ is described by Equation (2.78), although it can be verified that Equation (2.65) reduces to this equation by substituting appropriately. The measurement given by Equation (2.66) is seen to be

$$\dot{\underline{z}} - A(t)\underline{z} = [HF - AH + \dot{H}]\underline{x} + H\underline{w} + \underline{v} . \tag{2.85}$$

The estimate for $\underline{x}$ can be written explicitly in the following manner.

$$\dot{\hat{\underline{x}}} = F(t)\hat{\underline{x}} + K(t)[\dot{\underline{z}} - A(t)\underline{z} - J(t)\hat{\underline{x}}] , \tag{2.86}$$

where

$$J \triangleq HF - AH + \dot{H} .$$

The initial condition for $\hat{\underline{x}}(t_0)$ is obtained from Equation (2.68) and is found to be

$$\hat{\underline{x}}(t_0) = M(t_0)H^T(t_0)[H(t_0)M(t_0)H^T(t_0) + N(t_0)]^{-1}\underline{z}(t_0) , \tag{2.87}$$

where $M(t_0)$ is the covariance of the initial state $\underline{x}(t_0)$.

The gain matrix $K(t)$ is given by

$$K(t) = [P(t)J^T(t) + Q(t)H^T(t)][H(t)Q(t)H^T(t) + R(t)]^{-1} \tag{2.88}$$

and the error covariance matrix is

$$\dot{P} = F(t)P + F^T(t)P + Q(t) - K(t)[H(t)Q(t)H^T(t) + R(t)]K^T(t) . \tag{2.89}$$

The initial conditions are determined from Equation (2.69) to be

$$P(t_0) = M(t_0) - M(t_0)H^T(t_0)[H(t_0)M(t_0)H^T(t_0) + N(t_0)]^{-1}H(t_0)M(t_0) . \tag{2.90}$$

These results reduce to those published by Bucy[24] and by Stear and Stubberud[25] after minor changes in the model and notation are introduced.

2.3 Linear Smoothing

In this section the problem of finding the linear, unbiased, minimum variance estimate of $\underline{x}(t)$, given the data $\underline{z}(\tau)$, $\tau > t$, is considered. The system defined by Equations (1.1) - (1.2) is treated and the innovations approach of Section 2.1.2 is used to derive the solution. The development follows Kailath[27].

Based on the results of section 2.1.2, suppose that the smoothed estimate is given by

$$\hat{\underline{x}}(t|\tau) = \int_{t_0}^{\tau} W_s(t,s)\underline{r}(s)\,ds , \tag{2.91}$$

where W_s is chosen so that the error in the estimate is orthogonal to the innovations process $\underline{r}(s)$

$$E\{[\underline{x}(t) - \hat{\underline{x}}(t|\tau)]^T\underline{r}(\sigma)\} = 0 , \qquad t_0 \leq \sigma < \tau .$$

From the orthogonality property, one obtains the condition that

$$\begin{aligned} E[\underline{x}(t)\underline{r}^T(\sigma)] &= \int_{t_0}^{\tau} W_s(t,s)E[\underline{r}(s)\underline{r}^T(\sigma)]\,ds \\ &= W_s(t,\sigma)R(\sigma) . \end{aligned} \tag{2.92}$$

But this implies that

$$\begin{aligned} \hat{\underline{x}}(t|\tau) &= \int_{t_0}^{\tau} E[\underline{x}(t)\underline{r}^T(s)]R^{-1}(s)\underline{r}(s)\,ds \\ &= \int_{t_0}^{t} E[\underline{x}(t)\underline{r}^T(s)]R^{-1}(s)\underline{r}(s)\,ds + \int_{t}^{\tau} E[\underline{x}(t)\underline{r}^T(s)]R^{-1}(s)\underline{r}(s)\,ds . \end{aligned} \tag{2.93}$$

The first term can be recognized as the filtered estimate; so Equation (2.93) reduces to

$$\hat{\underline{x}}(t|\tau) = \hat{\underline{x}}(t|t) + \int_{t}^{\tau} E[\underline{x}(t)\underline{r}^T(s)]R^{-1}(s)\underline{r}(s)\,ds . \tag{2.94}$$

Now consider $E[\underline{x}(t)\underline{r}^T(s)]$. From Equation (2.92), and assuming that $\sigma > t$,

$$\begin{aligned} W_s(t,\sigma)R(\sigma) &= E[\underline{x}(t)\underline{r}^T(\sigma)] \\ &= E[\underline{x}(t)\tilde{\underline{x}}^T(\sigma|\sigma)H^T(\sigma) + \underline{x}(t)\underline{v}^T(\sigma)] \\ &= E[\underline{x}(t)\tilde{\underline{x}}^T(\sigma|\sigma)]H^T(\sigma) \\ &= E\{[\hat{\underline{x}}(t|t) + \tilde{\underline{x}}(t|t)]\tilde{\underline{x}}^T(\sigma|\sigma)\}H^T(\sigma) \\ &= P(t,\sigma)H^T(\sigma) , \qquad \sigma > t , \end{aligned}$$

where

$$P(t,\sigma) \triangleq E[\tilde{\underline{x}}(t|t)\tilde{\underline{x}}^T(\sigma|\sigma)] . \tag{2.95}$$

Using this definition, the smoothed estimate becomes

$$\hat{\underline{x}}(t|\tau) = \hat{\underline{x}}(t|t) + \int_{t}^{\tau} P(t,s)H^T(s)R^{-1}(s)\underline{r}(s)\,ds . \tag{2.96}$$

This shows that $\hat{\underline{x}}(t|\tau)$ is the linear combination of the filtered estimate $\hat{\underline{x}}(t|t)$ and a correction term that contains the data not included in $\hat{\underline{x}}(t|t)$.

The error covariance matrix $P(t|\tau) \triangleq E[\tilde{\underline{x}}(t|\tau)\tilde{\underline{x}}^T(t|\tau)]$ can be defined in terms of the filtering error covariance matrix $P(t|t)$. Observe that the error in the smoothed estimate is

$$\tilde{\underline{x}}(t|\tau) = \tilde{\underline{x}}(t|t) - \int_{t}^{\tau} P(t,s)H^T(s)R^{-1}(s)\underline{r}(s)\,ds .$$

Since $\tilde{\underline{x}}(t|\tau)$ is orthogonal to the residual $\underline{r}(s)$, it is clear that

$$E[\tilde{\underline{x}}(t|\tau)\tilde{\underline{x}}^T(t|\tau)] = E[\tilde{\underline{x}}(t|t)\tilde{\underline{x}}^T(t|t)] - \int_{t}^{\tau} P(t,s)H^T(s)R^{-1}(s)E[\underline{r}(s)\tilde{\underline{x}}^T(t|t)]\,ds$$

$$P(t|\tau) = P(t|t) - \int_{t}^{\tau} P(t,s)H^T(s)R^{-1}(s)H(s)P(s,t)\,ds . \tag{2.97}$$

The integrand is non-negative-definite, so the effect of the additional data in the smoothed estimate is to reduce the error in the filtered estimate.

The correlation matrix $P(t,s)$ is easily determined. As defined, it is known that

$$P(t,s) \triangleq E[\tilde{\underline{x}}(t|t)\tilde{\underline{x}}^T(s|s)] , \qquad s > t .$$

But $\tilde{\underline{x}}(t|t)$ is the solution of

$$\dot{\tilde{\underline{x}}} = [F(t) - K(t)H(t)]\tilde{\underline{x}} + \underline{w}(t) - K(t)\underline{v}(t) ;$$

so

$$\tilde{\underline{x}}(s|s) = \Psi(s,t)\tilde{\underline{x}}(t|t) + \Psi(s,t)\int_t^s \Psi(t,\sigma)[\underline{w}(\sigma) - K(\sigma)\underline{v}(\sigma)]\, d\sigma ,$$

where $\Psi(s,t)$ is the fundamental solution obtained from the matrix differential equation

$$\frac{d}{ds}\Psi(s,t) = [F(s) - K(s)H(s)]\Psi(s,t) , \qquad \Psi(t,t) = I .$$

Using this solution, one obtains

$$P(t,s) = P(t|t)\Psi^T(s,t) .$$

The smoothed estimate becomes

$$\hat{\underline{x}}(t|\tau) = \hat{\underline{x}}(t|t) + P(t|t)\int_t^\tau \Psi^T(s,t)H^T(s)R^{-1}(s)\underline{r}(s)\, ds \tag{2.98}$$

and

$$\begin{aligned} P(t|\tau) &\triangleq E[\tilde{\underline{x}}(t|\tau)\tilde{\underline{x}}^T(t|\tau)] \\ &= P(t|t) - P(t|t)\int_t^\tau \Psi^T(s,t)H^T(s)R^{-1}(s)H(s)\Psi(s,t)\, ds\, P(t|t) \\ &\triangleq P(t|t) - P(t|t)M_s(t,\tau)P(t|t) , \end{aligned} \tag{2.99}$$

where

$$M_s(t,\tau) = \int_t^\tau \Psi^T(s,t)H^T(s)R^{-1}(s)H(s)\Psi(s,t)\, ds .$$

The matrix $M_s(t,\tau)$ is similar to the observability matrix introduced in Section 3.1. Equations (2.98) and (2.99) are the general smoothing equations. Three classes of smoothing problems have been discussed in the literature[28]:

(i) Fixed-interval smoothing[29]: The initial time t_0 and the final time τ are fixed.

(ii) Fixed-point smoothing[30]: The time t for which a smoothed estimate is determined is fixed, while the amount of data increases (i.e. τ increases).

(iii) Fixed-lag smoothing[28]: The time for which a smoothed estimate is determined is a fixed amount Δ behind the most recent data occurring at time τ.

These results can be derived[27] from the general results and will not be discussed here.

3. OBSERVABILITY AND THE BEHAVIOR OF THE ERROR COVARIANCE MATRIX

In this section two important concepts related to the linear estimation problem are discussed. The first of these is observability[13,31], that is, the property of a system which permits estimation of its state. The second concept is the stability of the estimate as defined by the behavior of the error covariance[13,32].

3.1 Observability of Deterministic Systems

Consider a system described by Equations (1.1) and (1.2). In contrast to Section 2, it is assumed at this point that $\underline{w}(t)$ is a deterministic signal and that $\underline{v}(t)$ is identically zero. A state $\underline{x}(t)$ of the resultant deterministic system is called *observable* if from the input $\underline{w}(\tau)$, $t_0 \leq \tau \leq t$, and the output $\underline{z}(t)$, $\underline{x}(t)$

can be completely determined. If all states $\underline{x}(t)$ corresponding to all admissible $\underline{z}(t)$ are observable, the system is called *completely observable*.

The solution of Equation (1.1) is given by

$$\underline{x}(t) = \Phi(t,t_0)\underline{x}(t_0) + \int_{t_0}^{t} \Phi(t,\tau)\underline{w}(\tau)\, d\tau \,, \tag{3.1}$$

where $\Phi(t,\tau)$ is the transition matrix for Equation (1.1),

$$\frac{d\Phi(t,\tau)}{dt} = F(t)\Phi(t,\tau) \,; \qquad \Phi(\tau,\tau) = I \,. \tag{3.2}$$

The $\underline{w}(\tau)$, $t_0 \leq \tau \leq t$, and $\Phi(t,\tau)$ completely define the second term appearing in Equation (3.1). Apparently if $\underline{x}(t_0)$ can be completely determined from knowledge of $Z(t)$, then $\underline{x}(t)$ can be completely determined, that is, the state $\underline{x}(t)$ will be observable. If this is true for an arbitrary state, then the system is completely observable.

3.1.1 Observability Criterion

Combining Equation (3.1) and Equation (1.2) with $\underline{v}(t)$ set to zero, the observation vector $\underline{z}(t)$ is given by

$$\underline{z}(t) = H(t)\underline{x}(t) = \int_{t_0}^{t} H(t)\Phi(t,\tau)\underline{w}(\tau)\, d\tau + H(t)\Phi(t,t_0)\underline{x}(t_0) \,. \tag{3.3}$$

Knowledge of $\underline{Z}(t)$ and $\underline{w}(\tau)$, $t_0 \leq \tau \leq t$, completely determines

$$\Delta\underline{z}(t) = \underline{z}(t) - \int_{t_0}^{t} H(t)\Phi(t,\tau)\underline{w}(\tau)\, d\tau$$

$$= H(t)\Phi(t,t_0)\underline{x}(t_0) \,; \qquad t_0 \leq \tau \leq t \,. \tag{3.4}$$

Thus it is sufficient to deal with $\Delta\underline{z}(\tau)$, $t_0 \leq \tau \leq t$, alone and seek to determine under what conditions $\underline{x}(t_0)$ can be completely determined from $\Delta\underline{z}(\tau)$, $t_0 \leq \tau \leq t$. These conditions are then conditions for the observability of the state $\underline{x}(t)$. If these conditions do not depend on $\underline{x}(t_0)$ or on $\Delta\underline{z}(\tau)$, $t_0 \leq \tau \leq t$, they are also conditions for complete observability of the system.

Now consider a linear function of the form

$$\int_{t_0}^{t} s^T(\tau)\Delta\underline{z}(\tau)\, d\tau \,, \tag{3.5}$$

where $s(\tau)$ is a matrix of piecewise continuous functions. The system is completely observable if, for some $t > t_0$, an $s(\tau)$ exists such that this function equals $\underline{x}(t_0)$ for arbitrary $\underline{x}(t_0)$. Since $\Delta\underline{z}(\tau)$ is linear in $\underline{x}(t_0)$, only linear functions need be considered.

Theorem 3.1

A necessary and sufficient condition that a system defined by Equations (1.1) and (1.2), with $\underline{w}(t)$ a deterministic function and $\underline{v}(t)$ identically zero, be completely observable is that the matrix

$$M(t_0,t) = \int_{t_0}^{t} \Phi^T(\tau,t_0)H^T(\tau)H(\tau)\Phi(\tau,t_0)\, d\tau \tag{3.6}$$

be positive-definite for some $t > t_0$.

Proof of Sufficiency

Assume that $M(t_0,t)$ is positive-definite and thus has an inverse. In Equation (3.5) let

$$s(\tau) = H(\tau)\Phi(\tau,t_0)M^{-1}(t_0,t) \,; \tag{3.7}$$

then

$$\int_{t_0}^{t} s^T(\tau)\Delta\underline{z}(\tau)\, d\tau = \int_{t_0}^{t} M^{-1}(t_0,t)\Phi^T(\tau,t_0)H^T(\tau)\Delta\underline{z}(\tau)\, d\tau$$

$$= M^{-1}(t_0,t)\int_{t_0}^{t} \Phi^T(\tau,t_0)H^T(\tau)H(\tau)\Phi(\tau,t_0)\, d\tau\, \underline{x}(t_0)$$

$$= \underline{x}(t_0) \,. \tag{3.8}$$

Since the condition is independent of $\underline{x}(t_0)$, the system is completely observable.

Proof of Necessity (by Contradiction)

Suppose the system is completely observable, but that $M(t_0,t)$ is singular and hence not positive-definite. This implies that a non-zero constant vector $\underline{p}$ exists such that

$$\underline{p}^T M(t_0,t)\underline{p} = \underline{p}^T \int_{t_0}^{t} \Phi^T(\tau,t_0) H^T(\tau) H(\tau) \Phi(\tau,t_0)\, d\tau\, \underline{p}$$

$$= \int_{t_0}^{t} [H(\tau)\Phi(\tau,t_0)\underline{p}]^T [H(\tau)\Phi(\tau,t_0)\underline{p}]\, d\tau$$

$$= 0 . \qquad (3.9)$$

Since $H(\tau)\Phi(\tau,t_0)\underline{p}$ is a continuous function for $t_0 \leq \tau \leq t$, it must be identically zero in order that Equation (3.9) be satisfied.

Now, if the system is completely observable, then for $\underline{x}(t_0) = \underline{p}$ there exists an $\alpha^*(\tau)$, $t_0 \leq \tau \leq t$, such that

$$\int_{t_0}^{t} \alpha^{*T}(\tau)\Delta\underline{z}(\tau)\, d\tau = \int_{t_0}^{t} \alpha^{*T}(\tau) H(\tau)\Phi(\tau,t_0)\underline{p}\, d\tau = \underline{p} . \qquad (3.10)$$

However, since $H(\tau)\Phi(\tau,t_0)\underline{p} = 0$, $t_0 \leq \tau \leq t$, the last equality cannot hold and the system is not completely observable.

When $H(t)$ and $F(t)$ are constant matrices, the complete observability criterion can be reduced to an algebraic criterion.

Theorem 3.2

A necessary and sufficient condition that a system defined by Equations (1.1) and (1.2), with $\underline{w}(t)$ a deterministic function, $\underline{v}(t)$ identically zero, and $H(t)$ and $F(t)$ constant matrices, be completely observable is that the matrix (dimension $n \times mn$)

$$U = [H^T, F^T H^T, \ldots, (F^T)^{n-1} H^T] \qquad (3.11)$$

have rank n (the dimension of the state $\underline{x}(t_0)$).

Proof of Sufficiency (by Contradiction)

Since the system is time-invariant, t_0 may be set equal to zero without loss of generality. Also

$$\Phi(t,0) = e^{Ft} . \qquad (3.12)$$

Let $h_1, h_2, \ldots, h_m$ be the columns of H^T. Assume that U has rank n but that the system is not completely observable, that is, $M(0,t)$ is singular. If $M(0,t)$ is singular then a non-zero $\underline{p}$ exists such that

$$\underline{p}^T e^{F^T t} h_i = 0 ; \quad 0 \leq \tau \leq t , \quad i = 1,2,\ldots,m . \qquad (3.13)$$

Now differentiate each of the m Equations (3.13) j times, $j = 0,1,2,\ldots,n-1$, thus generating

$$\underline{p}^T (F^T)^j e^{F^T t} h_i = 0 ; \quad i = 1,2,\ldots,m \text{ and } j = 0,1,\ldots,n-1 . \qquad (3.14)$$

Now Equations (3.14) must hold for all $t > 0$, since U is constant. In the limit as t approaches zero, Equations (3.14) become

$$\underline{p}^T (F^T)^j h_i = 0 ; \quad i = 1,2,\ldots,m \text{ and } j = 0,1,\ldots,n-1 . \qquad (3.15)$$

If the rank of U is n, then $\underline{p} = 0$, which contradicts the assumption that $\underline{p} \neq 0$.

Proof of Necessity

Assume the system is completely observable but that the rank of U is less than n. There then exists a non-zero $\underline{p}$ which satisfies Equation (3.15). Now form

$$\underline{p}^T e^{F^T t} H_i = \underline{p}^T \sum_{j=0}^{\infty} (F^T t)^j / j!\ \ h_i$$

$$= \underline{p}^T \sum_{j=0}^{n-1} \alpha_j (F^T t)^j h_i , \quad i = 1,2,\ldots,m . \qquad (3.16)$$

where the last equality follows from the Cayley-Hamilton theorem. It follows from Equation (3.16) that

$$\underline{p}^T M(0,t)\underline{p} = 0 \tag{3.17}$$

for some $p \neq 0$, thus contradicting the assumption that the system is completely observable.

3.1.2 *Observability and Least-Squares Estimates*

Consider again a system described by Equations (1.1) and (1.2). In this case it is assumed that $\underline{w}(t)$ is identically zero and $\underline{v}(t)$ is a vector white-noise with covariance matrix $R(t)\delta(t-\tau)$, where $R(t)$ is positive-definite. Now the observation vector can be written in the form

$$\underline{z}(t) = H(t)\Phi(t,t_0)\underline{x}(t_0) + \underline{v}(t) . \tag{3.18}$$

This system is called *completely observable* if, for every t_0 and every $\underline{x}(t_0)$, there exists a $t > 0$ such that an unbiased estimate $\hat{\underline{x}}(t_0)$, which is a linear function of $\underline{z}(t)$, can be constructed. The general criterion for complete observability is given in the following theorem.

Theorem 3.3

A necessary and sufficient condition that a system defined by Equations (1.1) and (1.2), with $\underline{w}(t)$ identically zero and $\underline{v}(t)$ a white-noise with autocorrelation matrix $R(t)\delta(t-\tau)$, $R(t)$ positive-definite, be completely observable is that the matrix

$$M(t_0,t) = \int_{t_0}^{t} \Phi^T(\tau,t) H^T(\tau) R^{-1}(\tau) H(\tau) \Phi(\tau,t)\, d\tau \tag{3.19}$$

be positive-definite for some $t > t_0$.

The proof of this theorem parallels that of Theorem 3.1.

If the system is time-invariant, that is, $H(t)$ and $F(t)$ are constant, then the complete observability criterion of Theorem 3.3 can be simplified to the algebraic criterion of the following theorem.

Theorem 3.4

A necessary and sufficient condition that the system defined in Theorem 3.3 be completely observable is that the matrix

$$U = [H^T, F^T H^T, \ldots, (F^T)^{n-1} H^T] \tag{3.20}$$

have rank n (the dimension of the state $\underline{x}(t)$).

The proof of this theorem parallels that of Theorem 3.2 and is not given.

3.2 The Matrix-Riccati Equation

3.2.1 *General Solution*

In Section 2 it was shown that a matrix differential equation of the Riccati type is of central importance in the linear estimation problem. In this section the analytic solution of this type of equation is discussed[16].

Consider the general matrix-Riccati equation of the form

$$\dot{W}(t) = W(t)A^T(t) + A(t)W(t) + W(t)B(t)W(t) + C(t) , \tag{3.21}$$

with the initial condition

$$W(t_0) = W_0 ,$$

where W_0 is a non-negative-definite matrix.

The matrices $A(t)$, $B(t)$, and $C(t)$ are $n \times n$ matrices of continuous functions. Further, $B(t)$ and $C(t)$ are assumed to be non-negative-definite for all $t \geq t_0$. Now consider the pair of linear matrix differential equations

$$\begin{aligned} \dot{Y}(t) &= A(t)Y(t) + C(t)Z(t) ; \qquad Y(t_0) = W_0 \\ \dot{Z}(t) &= -B(t)Y(t) - A^T(t)Z(t) ; \qquad Z(t_0) = I . \end{aligned} \tag{3.22}$$

It can be shown by direct substitution that

$$Y(t) = W(t)Z(t) \quad . \tag{3.23}$$

Differentiating this forms

$$\dot{Y}(t) = \dot{W}(t)Z(t) + W(t)\dot{Z}(t) \quad . \tag{3.24}$$

Using Equation (3.22), this becomes

$$A(t)Y(t) + C(t)Z(t) = \dot{W}(t)Z(t) - W(t)B(t)Y(t) - W(t)A^T(t)Z(t) \quad . \tag{3.25}$$

Substituting $Y(t) = W(t)Z(t)$ and collecting terms produces the equation

$$[\dot{W}(t) - W(t)A^T(t) - A(t)W(t) - W(t)B(t)W(t) - C(t)]Z(t) = 0 \quad . \tag{3.26}$$

Now if $Z(t)$ is nonsingular for all $t \geq t_0$, this equation is equivalent to Equation (3.21), and

$$W(t) = Y(t)Z^{-1}(t) \quad . \tag{3.27}$$

Note also that

$$W(t_0) = Y(t_0)Z^{-1}(t_0) = W_0 I = W_0 \quad ; \tag{3.28}$$

thus the initial condition is also satisfied.

To show that $Z(t)$ is nonsingular, it is noted that

$$\begin{aligned} \dot{Z}(t) &= -B(t)Y(t) - A^T(t)Z(t) \\ &= [-A^T(t) - B(t)W(t)]Z(t) \quad . \end{aligned} \tag{3.29}$$

Since $Z(t_0) = I$, then $Z(t)$ is a transition matrix and hence nonsingular.

It was seen in Section 2 that there are two forms of the error covariance equation. The first of these is

$$\begin{aligned} \dot{P}(t|t) &= P(t|t)[F(t) - K(t)H(t)]^T + [F(t) - K(t)H(t)]P(t|t) + K(t)R(t)K^T(t) + Q(t) \\ P(t_0) &= P_0 \quad . \end{aligned} \tag{3.30}$$

Comparing Equation (3.30) with Equation (3.21) the following equivalent set of equations is obtained by a comparison with Equations (3.22):

$$\left.\begin{aligned} \dot{Y}(t) &= [F(t) - K(t)H(t)]Y(t) + [K(t)R(t)K^T(t) + Q(t)]Z(t) \quad ; \qquad Y(t_0) = P_0 \\ \dot{Z}(t) &= -[F(t) - K(t)H(t)]^T Z(t) \quad ; \qquad Z(t_0) = I \quad . \end{aligned}\right\} \tag{3.31}$$

Now

$$Z(t) = \Psi^T(t_0, t) \quad , \tag{3.32}$$

where $\Psi(t, t_0)$ is a transition matrix describing the filter dynamics and satisfying

$$\begin{aligned} \frac{d\Psi(t, t_0)}{dt} &= [F(t) - K(t)H(t)]\Psi(t, t_0) \\ \Psi(t_0, t_0) &= I \quad . \end{aligned} \tag{3.33}$$

$Y(t)$ can now be written as

$$Y(t) = \Psi(t, t_0)P_0 + \int_{t_0}^{t} \Psi(t, \tau)\, K(\tau)R(\tau)K^T(\tau) + Q(\tau)\, \Psi^T(t_0, \tau)\, d\tau \quad . \tag{3.34}$$

Finally,

$$\begin{aligned} P(t|t) &= Y(t)Z^{-1}(t) \\ &= \Psi(t, t_0)\left[P_0 + \int_{t_0}^{t} \Psi(t_0, \tau)\right][K(\tau)R(\tau)K^T(\tau) + Q(\tau)]\Psi^T(t, \tau)\, d\tau]\Psi^T(t, t_0) \quad . \end{aligned} \tag{3.35}$$

The second form of the error covarinnce equation is obtained from Equation (3.30) by letting

$$K(t) = P(t|t)H^T(t)R^{-1}(t) . \tag{3.36}$$

The resulting equation is

$$\dot{P}(t|t) = P(t|t)F^T(t) + F(t)P(t|t) - P(t|t)H^T(t)R^{-1}(t)H(t)P(t|t) + Q(t) \tag{3.37}$$

$$P(t_0) = P_0 .$$

The set of equivalent linear equations is

$$\left.\begin{aligned} \dot{Y}(t) &= F(t)Y(t) + Q(t)Z(t); \qquad Y(t_0) = P_0 \\ \dot{Z}(t) &= H^T(t)R^{-1}(t)H(t)Y(t) - F^T(t)Z(t) ; \qquad Z(t_0) = I . \end{aligned}\right\} \tag{3.38}$$

Now let

$$\Theta(t,t_0) = \begin{bmatrix} \Theta_{11}(t,t_0) & \Theta_{12}(t,t_0) \\ \Theta_{21}(t,t_0) & \Theta_{22}(t,t_0) \end{bmatrix} \tag{3.39}$$

be the transition matrix corresponding to the matrix

$$\begin{bmatrix} F(t) & Q(t) \\ H^T(t)R^{-1}(t)H(t) & -F^T(t) \end{bmatrix} . \tag{3.40}$$

Therefore

$$Y(t) = \Theta_{11}(t,t_0)P_0 + \Theta_{12}(t,t_0) \tag{3.41}$$

$$Z(t) = \Theta_{21}(t,t_0)P_0 + \Theta_{22}(t,t_0) \tag{3.42}$$

and finally

$$\begin{aligned} P(t|t) &= Y(t)Z^{-1}(t) \\ &= [\Theta_{11}(t,t_0)P_0 + \Theta_{12}(t,t_0)][\Theta_{21}(t,t_0)P_0 + \Theta_{22}(t,t_0)]^{-1} . \end{aligned} \tag{3.43}$$

We now consider two special cases.

Case 1: Let $H = 0$, so that

$$\dot{P} = PF^T(t) + F(t)P + Q(t) . \tag{3.44}$$

In this case the Ricatti Equation (3.37) reduces to a linear matrix equation whose solution can be obtained using the general formalism. From Equation (3.38),

$$\dot{Z} = -F^T(t)Z , \qquad Z(t_0) = I .$$

But this equation is the adjoint of

$$\dot{\Phi} = F(t)\Phi , \qquad \Phi(t_0,t_0) = I$$

so

$$Z(t) = \Theta_{22}(t,t_0)Z(t_0) = \Phi^T(t_0,t) . \tag{3.45}$$

This implies that

$$\Theta_{21}(t,t_0) = 0$$

$$\Theta_{22}(t,t_0) = \Phi^T(t_0,t) .$$

Using Equation (3.45), one has

$$\dot{Y} = F(t)Y + Q(t)\Phi^T(t_0,t) ,$$

so it follows that

$$Y(t) = \Phi(t,t_0)Y(t_0) + \Phi(t,t_0)\int_{t_0}^{t}\Phi(t_0,\tau)Q(\tau)\Phi^T(t_0,\tau)\,d\tau \tag{3.46}$$

which implies that

$$\Theta_{11}(t,t_0) = \Phi(t,t_0)$$

$$\Theta_{12}(t,t_0) = \Phi(t,t_0)\int_{t_0}^{t}\Phi(t_0,\tau)Q(\tau)\Phi^T(t_0,\tau)\,d\tau\ .$$

Substituting these into Equation (3.43), the solution of the linear matrix differential equation is found to be

$$P(t|t) = \Phi(t,t_0)P(t_0)\Phi^T(t,t_0) + \Phi(t,t_0)\left\{\int_{t_0}^{t}\Phi(t_0,\tau)Q(\tau)\Phi^T(t_0,\tau)\,d\tau\right\}\Phi^T(t,t_0)\ . \tag{3.47}$$

Case 2: Let $Q = 0$, so that

$$\dot{P} = PF^T(t) + F(t)P - PH^T(t)R^{-1}(t)H(t)P\ . \tag{3.48}$$

In this case the variance equation has no forcing term and one can consider

$$\dot{Y} = F(t)Y\ .$$

This has the solution

$$Y(t) = \Phi(t,t_0)P(t_0) \tag{3.49}$$

and implies that

$$\Theta_{11}(t,t_0) = \Phi(t,t_0)$$

$$\Theta_{12}(t,t_0) = 0\ .$$

Using Equation (3.49), the equation for $\dot{Z}$ becomes

$$\dot{Z} = -F^T(t)Z + H^T(t)R^{-1}(t)H(t)\Phi(t,t_0)P_0\ .$$

This has the solution

$$Z(t) = \Phi^T(t_0,t) + \Phi^T(t_0,t)\int_{t_0}^{t}\Phi^T(\tau,t_0)H^T(\tau)R^{-1}(\tau)H(\tau)\Phi(\tau,t_0)\,d\tau\,P_0\ . \tag{3.50}$$

so that

$$\Theta_{21}(t,t_0) = \Phi^T(t_0,t)\int_{t_0}^{t}\Phi^T(\tau,t_0)H^T(\tau)R^{-1}H(\tau)\Phi(\tau,t_0)\,d\tau$$

$$\Theta_{22}(t,t_0) = \Phi^T(t_0,t)\ .$$

The solution of Equation (3.48), using Equations (3.49) and (3.50) in Equation (3.43), is seen to be

$$P(t|t) = \Phi(t,t_0)P_0\left[\Phi^T(t_0,t) + \Phi^T(t_0,t)\int_{t_0}^{t}\Phi^T(\tau,t_0)H^T(\tau)R^{-1}(\tau)H(\tau)\Phi(\tau,t_0)\,d\tau\,P_0\right]^{-1}$$

$$= \Phi(t,t_0)P_0\left[I + \int_{t_0}^{t}\Phi^T(\tau,t_0)H^T(\tau)R^{-1}(\tau)H(\tau)\Phi(\tau,t_0)\,d\tau\,P_0\right]^{-1}\Phi^T(t,t_0)\ . \tag{3.51}$$

But observe that the integral term is identically equal to the observability matrix $M(t_0,t)$ defined in Equation (3.19). This suggests the important role played by observability considerations in establishing the behavior of the error of the minimum variance estimator

3.2.2 *Solution for the Stationary Case*

Suppose in Equation (3.37) that the coefficient matrices F, H, R, and Q are constant.. In this case, one can obtain more insight into the character of the general solution and can define the steady-state solution. Form

$$N \triangleq \begin{bmatrix} F & Q \\ H^TR^{-1}H & -F^T \end{bmatrix}\ . \tag{3.52}$$

where N represents the matrix of coefficients implied in the system (3.38). It has been shown[33] that the eigenvalues of N are real and that they are symmetric relative to the origin. For this discussion suppose that the eigenvalues are distinct and let Λ be the $n \times n$ diagonal matrix containing the eigenvalues that are positive. Then, there exists a transformation T such that

$$D \triangleq \begin{bmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{bmatrix} = T^{-1}NT , \tag{3.53}$$

where

$$T \triangleq \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} .$$

It can be shown that

$$T^{-1} \triangleq \begin{bmatrix} T_{22}^T & -T_{12}^T \\ -T_{21}^T & T_{11}^T \end{bmatrix} . \tag{3.54}$$

Define matrices Ω and Σ such that

$$\begin{bmatrix} Y \\ Z \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} \Omega \\ \Sigma \end{bmatrix} \tag{3.55}$$

and

$$\begin{bmatrix} \dot{Y} \\ \dot{Z} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} \dot{\Omega} \\ \dot{\Sigma} \end{bmatrix} .$$

From Equation (3.53), it follows that

$$\begin{bmatrix} \dot{\Omega} \\ \dot{\Sigma} \end{bmatrix} = \begin{bmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{bmatrix} \begin{bmatrix} \Omega \\ \Sigma \end{bmatrix} . \tag{3.56}$$

Since this is a linear, constant coefficient system, the solution of Equation (3.56) is

$$\begin{bmatrix} \Omega(t) \\ \Sigma(t) \end{bmatrix} = \begin{bmatrix} e^{\Lambda t} & 0 \\ 0 & e^{-\Lambda t} \end{bmatrix} \begin{bmatrix} \Omega(0) \\ \Sigma(0) \end{bmatrix} , \tag{3.57}$$

so that Z and Y are

$$\begin{bmatrix} Y(t) \\ Z(t) \end{bmatrix} = \left\{ T \begin{bmatrix} e^{\Lambda t} & 0 \\ 0 & e^{-\Lambda t} \end{bmatrix} T^{-1} \right\} \begin{bmatrix} Y(0) \\ Z(0) \end{bmatrix} . \tag{3.58}$$

The transition matrix $\Theta(t)$ for this system is

$$\Theta(t) = T \begin{bmatrix} e^{\Lambda t} & 0 \\ 0 & e^{-\Lambda t} \end{bmatrix} T^{-1}$$

$$= \begin{bmatrix} T_{11}e^{\Lambda t}T_{22}^T - T_{12}e^{-\Lambda t}T_{21}^T & T_{12}e^{-\Lambda t}T_{11}^T - T_{11}e^{\Lambda t}T_{12}^T \\ T_{21}e^{\Lambda t}T_{22}^T - T_{22}e^{-\Lambda t}T_{21}^T & T_{22}e^{-\Lambda t}T_{11}^T - T_{21}e^{\Lambda t}T_{12}^T \end{bmatrix} . \tag{3.59}$$

Substitution of Equation (3.59) into Equation (3.43) yields the solution $P(t)$.

$$P(t|t) = [T_{12}e^{-\Lambda t}T_{11}^T - T_{11}e^{\Lambda t}T_{12}^T + (T_{11}e^{\Lambda t}T_{22}^T - T_{12}e^{-\Lambda t}T_{21}^T)P_0]^* \times$$
$$\times [T_{22}e^{-\Lambda t}T_{11}^T - T_{21}e^{\Lambda t}T_{12}^T + (T_{21}e^{\Lambda t}T_{22}^T - T_{22}e^{-\Lambda t}T_{21}^T)P_0]^{-1} . \tag{3.60}$$

Consider the steady-state value of P as $t \to \infty$. The terms involving $e^{-\Lambda t}$ will vanish so

$$\begin{aligned}\lim_{t\to\infty} P(t|t) &\triangleq P_\infty \\ &= [-T_{11}e^{\Lambda t}T_{12}^T + T_{11}e^{\Lambda t}T_{22}^T P_0][-T_{21}e^{\Lambda t}T_{12}^T + T_{21}e^{\Lambda t}T_{22}^T P_0]^{-1} \\ &= T_{11}e^{\Lambda t}[-T_{12}^T + T_{22}^T P_0][-T_{12}^T + T_{22}^T P_0]^{-1}e^{-\Lambda t}T_{21}^{-1} \\ &= T_{11}T_{21}^{-1} \quad . \end{aligned} \tag{3.61}$$

Equation (3.60) has a computational disadvantage, since only the term $e^{\Lambda t}$ appears as t increases without bound. It is difficult to retain accuracy as t becomes large because of the presence of $e^{\Lambda t}$ in both $Y(t)$ and $Z^{-1}(t)$. This problem can be circumvented by expressing the initial value problem (Eqn (3.38)) as a two-point boundary-value problem[34]. Equation (3.57) can be rewritten as

$$\begin{bmatrix} \Omega(0) \\ \Sigma(t) \end{bmatrix} = \begin{bmatrix} e^{-\Lambda t} & 0 \\ 0 & e^{-\Lambda t} \end{bmatrix} \begin{bmatrix} \Omega(t) \\ \Sigma(0) \end{bmatrix} .$$

It has been shown that

$$Y(t) = P(t|t)Z(t)$$

and

$$Y(0) = P(0)Z(0) .$$

Using Equation (3.55), it can be seen that

$$Y(0) = P(0)[T_{21}\Omega(0) + T_{22}\Sigma(0)]$$

and

$$Y(0) = T_{11}\Omega(0) + T_{12}\Sigma(0) .$$

These relations imply that

$$0 = [T_{11} - P(0)T_{21}]\Omega(0) + [T_{12} - P(0)T_{22}]\Sigma(0) ,$$

so that

$$\Sigma(0) = -[T_{12} - P(0)T_{22}]^{-1}[T_{11} - P_0 T_{21}]\Omega(0) \triangleq R\Omega(0) .$$

Using this, $\Omega(t)$ and $\Sigma(t)$ can be related by

$$\Sigma(t) = e^{-\Lambda t}\Sigma(0) = e^{-\Lambda t}R\Omega(0) = e^{-\Lambda t}Re^{-\Lambda t}\Omega(t) . \tag{3.62}$$

Thus, one finds that

$$Y(t) = T_{11}\Omega(t) + T_{12}\Sigma(t) = [T_{11} + T_{12}e^{-\Lambda t}Re^{-\Lambda t}]\Omega(t)$$

$$Z(t) = [T_{21} + T_{22}e^{-\Lambda t}Re^{-\Lambda t}]\Omega(t)$$

and

$$P(t|t) = [T_{11} + T_{12}e^{-\Lambda t}Re^{-\Lambda t}][T_{21} + T_{22}e^{-\Lambda t}Re^{-\Lambda t}]^{-1} . \tag{3.63}$$

In this form only negative exponentials appear and the steady-state value P_∞ expressed by Equation (3.61) follows immediately. Thus Equation (3.63) appears to have computational advantages.

3.3 Bounds for the Error Covariance Matrix

The error covariance matrix $P(t)$ of the Kalman-Bucy filter provides the primary statistical measure of the behavior of the estimate. Since the $P(t|t)$ does not depend upon measurement data, it can be examined to determine the theoretical effectiveness of a particular measurement process for a given dynamical system. In this section upper and lower bounds for the error covariance matrix are derived[30,32] that point out the intrinsically different roles played by the plant and measurement noise processes. The bounds indicate the importance of the deterministic observability considerations of Section 3.1 and are derived using some of the properties of the matrix-Riccati equation solution discussed in Section 3.2.

3.3.1 Some General Bounds

Consider the system defined by Equations (1.1) and (1.2). This system can be reinterpreted as the sum of two simpler systems, which allows the effects of the measurement and plant processes to be examined separately and thereby obtain insight into their influence.

Define System P as a system with plant noise but no measurement noise,

$$\dot{\underline{x}}^P = F(t)\underline{x}^P + \underline{w}\,, \qquad \underline{x}^P(t_0) = 0 \tag{3.64}$$

$$\underline{z}^P = H(t)\underline{x}^P\,, \tag{3.65}$$

and define System M to have measurement noise but no plant noise

$$\dot{\underline{x}}^M = F(t)\underline{x}^M\,, \qquad \underline{x}^M(t_0) = \underline{x}(t_0) \tag{3.66}$$

$$\underline{z}^M = H(t)\underline{x}^M + \underline{v}\,. \tag{3.67}$$

Note that $\underline{x}(t_0) = \underline{x}^M(t_0) + \underline{x}^P(t_0)$

and that
$$\dot{\underline{x}}^P + \dot{\underline{x}}^M = \frac{d}{dt}[\underline{x}^P + \underline{x}^M]$$
$$= F(t)[\underline{x}^P + \underline{x}^M] + \underline{w}\,,$$

so that
$$\underline{x}^P(t) + \underline{x}^M(t) = \underline{x}(t)\,. \tag{3.68}$$

Also, it follows that

$$\underline{z}(t) = \underline{z}^P + \underline{z}^M\,. \tag{3.69}$$

The minimum variance estimate of $\underline{x}(t)$, given $\underline{Z}(t)$, has been shown to be

$$\dot{\hat{\underline{x}}} = F(t)\hat{\underline{x}} + K(t)[\underline{z} - H(t)\hat{\underline{x}}]\,, \tag{3.70}$$

where the optimal gain $K(t)$ is defined by Equations (2.11) and (2.12). Considering Systems P and M, define linear estimates of $\underline{x}^P$ and $\underline{x}^M$, using the optimal gain $K(t)$ of Equation (3.70), as

$$\dot{\hat{\underline{x}}}^P = F(t)\hat{\underline{x}}^P + K(t)[\underline{z}^P - H(t)\hat{\underline{x}}^P] \tag{3.71}$$

$$\dot{\hat{\underline{x}}}^M = F(t)\hat{\underline{x}}^M + K(t)[\underline{z}^M - H(t)\hat{\underline{x}}^M]\,. \tag{3.72}$$

Note that $\hat{\underline{x}}^P$ and $\hat{\underline{x}}^M$ are *not* the minimum variance estimators for $\underline{x}^P$ and $\underline{x}^M$. However, it follows without difficulty that the minimum variance estimate of $\underline{x}$ is given by

$$\hat{\underline{x}} = \hat{\underline{x}}^P + \hat{\underline{x}}^M\,. \tag{3.73}$$

Write the error in the estimate as

$$\tilde{\underline{x}} = \underline{x} - \hat{\underline{x}}$$
$$= (\underline{x}^P + \underline{x}^M) - (\hat{\underline{x}}^P + \hat{\underline{x}}^M)$$
$$= (\underline{x}^P - \hat{\underline{x}}^P) + (\underline{x}^M - \hat{\underline{x}}^M)\,. \tag{3.74}$$

Using Equation (3.74), the error covariance matrix $P(t|t)$ can be expressed as

$$P(t|t) \triangleq E\{[\underline{x}(t) - \hat{\underline{x}}(t|t)][\underline{x}(t) - \hat{\underline{x}}(t|t)]^T\}$$
$$= E[(\underline{x}^P - \hat{\underline{x}}^P)(\underline{x}^P - \hat{\underline{x}}^P)^T] + E[(\underline{x}^P - \hat{\underline{x}}^P)(\underline{x}^M - \hat{\underline{x}}^M)^T]$$
$$+ E[(\underline{x}^M - \hat{\underline{x}}^M)(\underline{x}^P - \hat{\underline{x}}^P)^T] + E[(\underline{x}^M - \hat{\underline{x}}^M)(\underline{x}^M - \hat{\underline{x}}^M)^T]\,.$$

But
$$E[(\underline{x}^P - \hat{\underline{x}}^P)(\underline{x}^M - \hat{\underline{x}}^M)^T] = 0\,,$$

since $\underline{x}^P(t_0) = 0$ and $\underline{w}(t)$ and $\underline{v}(t)$ are uncorrelated white-noise processes; so

$$P(t|t) = E[(\underline{x}^P - \hat{\underline{x}}^P)(\underline{x}^P - \hat{\underline{x}}^P)^T] + E[(\underline{x}^M - \hat{\underline{x}}^M)(\underline{x}^M - \hat{\underline{x}}^M)^T]\,. \tag{3.75}$$

It is possible to state general bounds for $P(t|t)$. Suppose that $P^M(t|t)$ and $P^P(t|t)$ represent the minimum variance error covariance matrices for Systems M and P, respectively. Then, since $\hat{\underline{x}}^P$ and $\hat{\underline{x}}^M$ are *not* the minimum variance estimates for these systems, it must certainly follow from Equation (3.75) that

$$P(t|t) \geq P^{p}(t|t) + P^{m}(t|t) \ , \tag{3.76}$$

where the matrix notation $A \geq B$ implies that the matrix $A - B$ is non-negative-definite.

On the other hand, suppose that some gain other than $K(t)$ is used to obtain estimates of both $\underline{x}^{p}(t)$ and $\underline{x}^{m}(t)$. That is, a suboptimal gain $K^{s}(t)$ is used in *both* subsystems. Then, from the definition of $K(t)$ it must be true that

$$P(t|t) \leq P^{ps}(t|t) + P^{ms}(t|t) \ , \tag{3.77}$$

where P^{ps} and P^{ms} represent the error covariance matrices associated with the suboptimal gain $K^{s}(t)$.

3.3.2 Unforced Dynamical Plant with Noisy Measurement Data

Consider the system described by

$$\dot{\underline{x}} = F(t)\underline{x} \ , \qquad \underline{x}(t_0) = \underline{x}_0 \tag{3.78}$$

$$\underline{z} = H(t)\underline{x} + \underline{y} \ . \tag{3.79}$$

This system is identical with System M but the superscripts have been eliminated to simplify the notation. The minimum variance estimate of $\underline{x}(t)$, given $Z(t)$, has been shown to be given by

$$\dot{\hat{\underline{x}}} = F(t)\hat{\underline{x}} + K(t)[\underline{z} - H(t)\hat{\underline{x}}] \ , \tag{3.80}$$

where

$$K(t) = P(t|t)H^{T}(t)R^{-1}(t) \tag{3.81}$$

and

$$\dot{P} = F(t)P + PF^{T}(t) - PH^{T}(t)R^{-1}(t)H(t)P \ . \tag{3.82}$$

But in Section 3.2, it was shown (i.e. Equation (3.51)) that the solution of the matrix-Ricatti equation in this case is

$$P(t|t) = \Phi(t,t_0)[P^{-1} + M(t_0,t)]^{-1}\Phi^{T}(t,t_0) \ , \tag{3.83}$$

where $M(t_0,t)$ is the observability matrix (see Equation (3.19)) and is defined to be

$$M(t_0,t) = \int_{t_0}^{t} \Phi^{T}(\tau,t_0)H^{T}(\tau)R^{-1}(\tau)H(\tau)\Phi(\tau,t_0)\,d\tau \tag{3.84}$$

and $\Phi(t,t_0)$ is the transition matrix associated with the dynamical system (Eqn (3.78)).

From Equation (3.83), one sees that the behavior of the observability matrix essentially determines the behavior of $P(t|t)$. One can show under appropriate conditions that $M(t_0,t)$ is strictly increasing with t (i.e. $M(t_0,t_2) - M(t_0,t_1)$ is positive-definite for some $t_2 > t_1$).

It is clear from the definition that

$$\delta M(t_1,t_2) \triangleq M(t_0,t_2) - M(t_0,t_1) = \int_{t_1}^{t_2} \Phi^{T}(\tau,t_0)H^{T}(\tau)R^{-1}(\tau)H(\tau)\Phi(\tau,t_0)\,d\tau$$

so that

$$\delta M(t_1,t_2) \geq 0 \ .$$

To consider the asymptotic behavior of $M(t_0,t)$, and therefore $P(t|t)$, it is convenient to introduce the concept of *q-observability*.

Definition: A system is said to be *q-observable* for $q > 0$ on an interval (t_0,t_N), where t_N may be infinite, if and only if $\delta M(t_j,t_k)$ is positive-definite for every t_j, t_k such that $t_N \geq t_k > t_j \geq t_0$ and $t_k - t_j \geq q$.

q-observability is a slight generalization of complete observability and is introduced to insure that the observability matrix $M(t_0,t)$ is strictly increasing over intervals of duration q. For stationary systems there is no difference, so that this concept is useful only for time-varying systems. In this case, q-observability insures that the system is completely observable for every interval of duration q.

If a system is q-observable, then

$$M(t_0,t) > M(t_0,t-q) \ , \qquad t - q \geq t_0 \ .$$

Let the eigenvalues $M(t_0,t)$ and $M(t_0,t-q)$ be $(\lambda_1^t \geq \lambda_2^t \geq \lambda_n^t)$ and $(\lambda_1^{t-q} \geq \lambda_2^{t-q} \geq \ldots \geq \lambda_n^{t-q})$, respectively. It follows that

$$\lambda_1^t > \lambda_1^{t-q}\,; \qquad \lambda_2^t > \lambda_2^{t-q}\,, \ \ldots,\ \lambda_n^t > t_n^{t-q}\,.$$

Thus, $M(t_0,t)$ is strictly increasing in the sense that *all* eigenvalues are strictly increasing as $t \to \infty$, although some may remain constant for intervals of length less than q.

For a symmetric matrix, the largest eigenvalue serves as a matrix norm (i.e. the spectral norm), so one can write

$$\|M(t_0,t)\| = \lambda_1^t\,,$$

so that one sees that

$$\lim_{t\to\infty} \|M(t_0,t)\| = \infty \tag{3.86}$$

and

$$\lim_{t\to\infty} \|M^{-1}(t_0,t)\| = \lim_{t\to\infty} \frac{1}{\lambda_n^t} = 0\,. \tag{3.80}$$

Note that $M^{-1}(t_0,t)$ exists if the system is completely observable for t. But Equation (3.86) implies that $M^{-1}(t_0,t)$ converges to the zero matrix, since its norm vanishes.

Assuming q-observability, one sees from Equation (3.83) that the error covariance matrix $P(t|t)$ is given approximately by

$$P(t|t) \simeq \Phi(t,t_0)M^{-1}(t_0,t)\Phi^{-T}(t,t_0)$$

for t sufficiently large. Since P_0^{-1} is positive-definite, it follows that

$$\begin{aligned}\|P(t|t)\| &= \|\Phi(t,t_0)[P_0^{-1}+M(t_0,t)]^{-1}\Phi^T(t,t_0)\| \\ &\leq \|\Phi(t,t_0)\|^2\ \|[P_0^{-1}+M(t_0,t)]^{-1}\| \\ &\leq \|\Phi(t,t_0)\|^2\ \|M^{-1}(t_0,t)\|\,.\end{aligned} \tag{3.87}$$

As a result of Equation (3.87), one concludes the following:

The error covariance matrix $P(t)$ *for the q-observable system (Eqns (3.78) – (3.79)) vanishes as* $t \to \infty$ *if* $\|M^{-1}(t,t_0)\|$ *converges to zero faster than* $\|\Phi(t,t_0)\|^2$ *increases.* If the plant (Eqn (3.78)) is stable, then this is certainly true.

This conclusion indicates, for a large class of systems, that the effect of measurement noise is eliminated by filtering over a sufficiently large period of time. Reference to Equation (3.81) also indicates that the gain matrix $K(t)$ vanishes as the error covariance matrix tends to zero. As a result, the estimate $\underline{\hat{x}}$ tends to be characterized by a homogeneous, linear differential equation identical to the plant equation (Eqn (3.78)). This shows that the most recent measurement data have a decreasing influence on the estimate. While this behavior is understandable from an entirely theoretical viewpoint it should be observed that, unless Equations (3.78) and (3.79) actually constitute an exact model of the system, the convergence of the $P(t|t)$ can lead to unduly optimistic measures of the error in the estimate and can ultimately lead to filter "divergence" and nonsensical results[36,37].

3.3.3 Noisy Dynamical Plant and Noise-Free Measurement Data

Consider a system containing no measurement noise

$$\dot{\underline{x}} = F(t)\underline{x} + \underline{w}\,, \qquad \underline{x}(t_0) = 0 \tag{3.88}$$

$$\underline{z} = H(t)\underline{x}\,. \tag{3.89}$$

It should first be noted that Equations (3.88) and (3.89) are essentially identical with the system (Equations (2.40) and (2.53)) treated in Section 2.2.2. This case was discussed there, since the Kalman-Bucy equations can not be applied directly to this system. Instead, the filter is reduced to dimension $(n-m)$ and is based on the system described by Equations (2.55), (2.56), and (2.58). Certainly, it is still possible to discuss the error covariance matrix $P(t|t)$ associated with the error in the estimate of $\underline{x}(t)$ but, as a consequence of the considerations of Section 2.2.2, the $P(t|t)$ can have, at most, rank $n-m$.

The $n \times n$ error covariance matrix $P(t|t)$ can never vanish identically as long as the plant covariance matrix $Q(t)$ is non-zero. This is an immediate consequence of the discussion in Section 3.2.1 relating to

Equation (3.30). In Equation (3.30) the matrix $K(t)$ is arbitrary and the associated solution is given by Equation (3.35). The measurement noise covariance matrix $R(t)$ can be set equal to zero (i.e. no measurement noise) in these equations. It is obvious that the term

$$\int_{t_0}^{t} \Psi(t_0,\tau) Q(\tau) \Psi^T(t_0,\tau)\, d\tau$$

is non-negative definite and never vanishes. The $\Psi(t,t_0)$ is a transition matrix, so it is always nonsingular. Therefore the matrix $P(t|t)$ described by Equation (3.35) with $R(t)$ set equal to zero never vanishes for non-zero $Q(t)$. Thus, in contrast with the discussion of the preceding section, the error covariance matrix $P(t|t)$ associated with Equations (3.88) and (3.89) vanishes only under very special conditions, and is never positive-definite. Also it becomes apparent that the plant noise covariance matrix can serve to prevent the error covariance matrix from becoming too small in any principal direction.

3.3.4 Error Bounds

It was demonstrated in Section 3.3.1 that the error covariance matrix can be bounded by

$$P^p(t|t) + P^m(t|t) \leq P(t|t) \leq P^{ps}(t|t) + P^{ms}(t|t) . \tag{3.90}$$

But in Section 3.3.1 the $P^m(t|t)$ was expressed more precisely by Equation (3.83). If the gain $K^m(t)$ defined by Equations (3.81) and (3.82) were used in both the M and P systems, then Equation (3.83) can be also used in the upper bound and Equation (3.90) becomes

$$\Phi(t,t_0)[P_0^{-1} + M(t_0,t)]^{-1}\Phi^T(t,t_0) + P^p(t|t) \leq P^{ps}(t|t) + \Phi(t,t_0)[P_0^{-1} + M(t_0,t)]^{-1}\Phi^T(t,t_0) . \tag{3.91}$$

The lower bound can be relaxed by eliminating $P^p(t|t)$, so that

$$P(t|t) \geq \Phi(t,t_0)[P_0^{-1} + M(t_0,t)]^{-1}\Phi^T(t,t_0) . \tag{3.92}$$

It is shown in Chapter 6 that this bound is essentially the Cramer-Rao inequality.

Consider the error that one has if there were no filtering for system P. Then

$$\underline{x}^p(t) = \int_{t_0}^{t} \Phi(t,\tau)\underline{w}(\tau)\, d\tau .$$

Assuming that $\underline{w}$ has zero mean, and assuming that no measurements are processed, the best estimate is

$$\hat{\underline{x}}(t) = 0 .$$

Then, it follows that

$$P^{p0}(t|t) = E[(\underline{x} - \hat{\underline{x}})(\underline{x} - \hat{\underline{x}})^T] = \int_{t_0}^{t} \Phi(t,\tau) Q(\tau) \Phi^T(t,\tau)\, d\tau , \tag{3.93}$$

But the matrix integral has been referred to as a stochastic controllability matrix by Kalman[13]. Let

$$W(t_0,t) \triangleq \int_{t_0}^{t} \Phi(t,\tau) Q(\tau) \Phi^T(t,\tau)\, d\tau . \tag{3.94}$$

Note that Equation (3.93) is identical with Equation (3.47) of Section 3.2.1. Using Equation (3.93), the upper bound becomes

$$P(t|t) \leq \Phi(t,t_0)[P_0^{-1} + M(t_0,t)]^{-1}\Phi^T(t,t_0) + W(t_0,t) . \tag{3.95}$$

Equations (3.92) and (3.95) provide more specific upper and lower bounds for the minimum variance error covariance matrix. From the discussion of Section 3.3.2, it can be seen that the upper and lower bounds are defined by $W(t_0,t)$ and $P^p(t|t)$ as t increases, since the $P^m(t|t)$ tends to vanish for q-observable systems.

Kalman has derived bounds for the error covariance matrix which bear some resemblance to those presented above[13]. It was shown that

$$P(t|t) \leq \Phi(t,\tau^*)M^{-1}(\tau^*,t)\Phi^T(t,\tau^*) + W(\tau^*,t) , \tag{3.96}$$

where τ^* defines the time at which the system is completely observable and completely controllable (i.e. $M(\tau^*,t)$ and $W(\tau^*,t)$ are positive-definite). Equation (3.96) is very similar to Equation (3.95). Kalman's lower bound is

$$[W^{-1}(\tau^*,t) + \Phi^T(\tau^*,t)M(\tau^*,t)\Phi(\tau^*,t)]^{-1} \leq P(t|t) \quad . \tag{3.97}$$

This matrix will vanish for a q-observable and q-controllable system and so does not appear to contain as much information as Equation (3.92).

A linear dynamical system with transition matrix $\Psi(t,t_0)$ is said to be uniformly asymptotically stable if

$$\|\Psi(t,t_0)\| \leq \alpha e^{-\beta(t-t_0)} \qquad \text{for all} \quad t \geq t_0 \quad ,$$

where α , $\beta > 0$. It is possible to prove that, if the system (Equations (1.1) and (1.2)) is completely controllable and completely observable, then the filter is uniformly asymptotically stable. That is, the norm of the transition matrix obtained as the solution of the matrix differential equation

$$\dot{\Psi}(t,t_0) = [F(t) - K(t)H(t)]\Psi(t,t_0) \quad , \qquad \Psi(t_0,t_0) = I$$

satisfies the inequality stated above. The proof of this result is omitted and the reader is directed to Kalman[13].

3.4 Behavior of Error Covariance for Stationary Systems

To obtain more insight into the behavior of the error covariance matrix, it is informative to consider stationary systems. In this case, the asymptotic behavior of the error covariance can be established more specifically and the upper bound can be investigated more thoroughly.

Consider the stationary linear system

$$\dot{\underline{x}} = F\underline{x} + \underline{w} \tag{3.98}$$

$$\underline{z} = H\underline{x} + \underline{v} \quad , \tag{3.99}$$

where F and H are constant matrices. Assume also that the plant and measurement noise covariance matrices Q and R are constant.

As has been shown above, the minimum variance error covariance matrix

$$\dot{P} = FP + PF^T - PH^TR^{-1}HP + Q \quad , \qquad P(t_0) = P_0 \quad . \tag{3.100}$$

Let P_∞ be the steady-state value of Equation (3.100), as discussed in Section 3.2.2. This matrix can be determined from Equation (3.61) and is the solution of the algebraic equation

$$0 = FP_\infty + P_\infty F^T - P_\infty H^TR^{-1}HP_\infty + Q \quad . \tag{3.101}$$

Use P_∞ to form a suboptimal gain for filter application

$$K_w = P_\infty H^TR^{-1} \quad . \tag{3.102}$$

This gain can be shown to be the gain obtained for the "classical" Wiener filter. The error covariance matrix for this suboptimal gain is obtained, using Equation (2.9) or Equation (3.30),

$$\dot{P}_w = (F - K_wH)P_w + P_w(F - K_wH)^T + (K_wRK_w^T + Q) \quad , \qquad P_w(t_0) = P(t_0) \quad . \tag{3.103}$$

This is a linear matrix equation of the form treated in Section 3.2.1 with the steady-state value P_∞ . The fact that $P_w(\infty) = P_\infty$ will be confirmed below.

Let
$$\delta P(t|t) \triangleq P(t|t) - P_\infty \quad , \tag{3.104}$$

where $P(t|t)$ is the minimum variance error covariance matrix. Also, define

$$\delta P_w(t|t) \triangleq P_w(t|t) - P_\infty \quad . \tag{3.105}$$

and let δP and δP_w be assigned the initial conditions

$$\delta P(t_0) = \delta P_w(t_0) = P_0 - P_\infty \triangleq \delta_0 \quad . \tag{3.106}$$

Since the system is stationary, the initial time t_0 can be assumed to be zero without loss of generality. Further, assume that

$$P_0 \geq P_\infty \quad .$$

This is a natural assumption and states that the initial uncertainty is greater than the steady-state error.

It follows from Equations (3.104), (3.101), and (3.100) that

$$\dot{\delta P} = (F - K_w H)\delta P + \delta P(F - K_w H)^T - \delta P H^T R^{-1} H \delta P \ . \tag{3.107}$$

But this is a matrix-Riccati equation with no forcing term. This was shown in Section 3.2.1 to have the solution

$$\delta P(t|t) = \Psi(t,t_0)\delta_0[I + M(t_0,t)\delta_0]^{-1}\Psi^T(t,t_0) \ , \tag{3.108}$$

where Ψ is the transition matrix associated with the linear system with coefficient matrix $(F - K_w H)$. But $\delta P(t|t)$ vanishes as t tends to infinity if the system is completely observable, thereby confirming that P_∞ is the steady-state solution of Equation (3.100).

One can also obtain, from Equations (3.105), (3.101), and (3.103),

$$\dot{\delta P}_w = (F - K_w H)\delta P_w + \delta P_w(F - K_w H)^T \ . \tag{3.109}$$

This is a homogeneous linear matrix equation and, from Equation (3.47), it follows that δP_w tends to vanish if $(F - K_w H)$ represents a stable system. This is indeed the case for an observable and controllable system, so that the steady-state value of P_w is established to be P_∞. Thus, the asymptotic behavior for the suboptimal gain K_w must be very similar to that for the minimum variance gain.

Let $$A \triangleq F - K_w H \ .$$

It is true, from Equation (3.47), that

$$\delta P_w(t|t) = e^{At}\delta_0 e^{A^T t} \ , \tag{3.110}$$

since the transition matrix $\Psi(t,0)$ associated with

$$\dot{\Psi}(t,0) = A\Psi(t,0)$$

is $$\Psi(t,0) = e^{At} \ .$$

One also finds, using Equation (3.51), that

$$\delta P(t|t) = e^{At}[\delta_0^{-1} + M(0,t)]^{-1} e^{A^T t} \ . \tag{3.111}$$

The difference between $P_w(t|t)$ and $P(t|t)$ can be established using Equations (3.110) and (3.111). First, note that

$$\begin{aligned} \delta P(t|t) &= e^{At}[I + \delta_0 M(0,t)]^{-1}\delta_0 e^{A^T t} \\ &= e^{At}[I + \delta_0 M(0,t)]^{-1} e^{-At}\delta P_w(t|t) \ . \end{aligned} \tag{3.112}$$

Rearranging terms, one has

$$P_w(t|t) - P(t|t) = \delta P_w(t|t) - \delta P(t|t) = e^{At}\delta_0 M(0,t) e^{-At}\delta P(t|t)$$

and

$$\delta P(t|t) = e^{At}M^{-1}(0,t)\delta_0^{-1} e^{-At}[P_w(t|t) - P(t|t)] \ . \tag{3.113}$$

Equating Equations (3.111) and (3.113), it follows that

$$P_w(t|t) - P(t|t) = e^{At}\delta_0 M(0,t)[\delta_0^{-1} + M(0,t)]^{-1} e^{A^T t} \ . \tag{3.114}$$

Equation (3.114) provides a relation for the difference between the error covariance of the suboptimal filter and the minimum variance filter. An inequality is easily derived from Equation (3.114) that can be used to describe the behavior of the error covariance. Let λ_1 and λ_n denote the largest and smallest eigenvalues of A. (Note that λ_1 and λ_n are negative since A represents a stable system.) Then the spectral norm of the transition matrix is

$$\|e^{At}\| = e^{\lambda_1 t} \tag{3.115}$$

and

$$\|e^{-At}\| = e^{-\lambda_n t} \ . \tag{3.116}$$

and it follows that

$$\begin{aligned}\|M(0,t)\| &= \left\| \int_0^t e^{A^T\tau} H^T R^{-1} H e^{A\tau}\, d\tau \right\| \\ &\leq \int_0^t \|e^{A^T\tau}\| \, \|H^T R^{-1} H\| \, \|e^{A\tau}\| \, d\tau \\ &= \|H^T R^{-1} H\| \int_0^t e^{2\lambda_1 \tau}\, d\tau \\ &= \frac{\|H^T R^{-1} H\|}{2\lambda_1} \left(e^{2\lambda_1 t} - 1 \right) \\ &= \frac{\|H^T R^{-1} H\|}{2|\lambda_1|} \left(1 - e^{2\lambda_1 t} \right) . \end{aligned} \tag{3.117}$$

One also sees that

$$\| [\delta_0^{-1} + M(0,t)]^{-1} \| \leq \|\delta_0\| . \tag{3.118}$$

Using Equations (3.115) through (3.118) in (3.114), it is found that

$$\|P_w(t|t) - P(t|t)\| \leq e^{2\lambda_1 t} \|\delta_0\|^2 \, \frac{\|H^T R^{-1} H\|}{2|\lambda_1|} \left(1 - e^{2\lambda_1 t} \right) . \tag{3.119}$$

The time-dependent factor reaches a maximum which is determined by considering

$$\frac{d}{dt} \left[e^{2\lambda_1 t} \left(1 - e^{2\lambda_1 t} \right) \right] = 0 .$$

This implies that

$$2e^{2\lambda_1 t} = 1 ,$$

so that the maximum occurs at

$$t = \frac{1}{2\lambda_1} \log_e \tfrac{1}{2} .$$

The maximum is easily found to be $\frac{1}{4}$, so that the maximum difference is

$$\|P_w(t|t) - P(t|t)\| \leq \frac{\|\delta_0\|^2 \, \|H^T R^{-1} H\|}{8|\lambda_1|} \qquad \text{for all } t . \tag{3.120}$$

The difference between the suboptimal and optimal error covariances is seen to depend upon the initial uncertainty δ_0, the signal-to-noise ratio as described by $\|H^T R^{-1} H\|$ and the smallest time-constant $|\lambda_1|$ of the filter dynamics. As one should expect, the larger the initial uncertainty or the better the signal-to-noise ratio, the greater the error introduced by the suboptimal gain. On the other hand, if the filter system has a large time constant, the suboptimal gain does not necessarily cause a significant deterioration of filter performance.

To summarize, the determination of the steady-state value of the optimal filter permits the definition of a suboptimal gain (i.e. the gain of the Wiener filter) which in many cases may provide satisfactory filter performance. A bound on the norm of the difference between the optimal and suboptimal error covariance matrices is determined which involves the intuitively obvious elements of the system which determine filter performance. More extensive discussion of this special problem is provided by Singer[31].

4. CONCLUSION

In this concluding section the intent is to state the principal results and the highlights of Sections 2 and 3 in order to allow the casual reader to sidestep the morass of details found there and to provide a support for the more interested reader with which he may anchor himself as he slides into morass.

Section 1.2 provides a statement of the mathematical model in terms of Equations (1.1) and (1.2) and then [illegible] the Unbiased Minimum Variance, Linear Estimation Problem. These aspects shall not be repeated here.

Section 2.1: Linear Prediction and Filtering

Section 2.1 contains the solution of the problem stated in Section 1.2 in terms of the Kalman-Bucy filter equations. This is first accomplished in Section 2.1, using arguments which, hopefully, appeal to intuition and provide some understanding of the basic character of the solution. This development is put on a more rigorous basis in Section 2.1.2, wherein the properties of the measurement residual process

$$\underline{r}(t) \triangleq \underline{z}(t) - H(t)\hat{\underline{x}}(t|t) \tag{4.1}$$

are examined and then used to derive the filter equations. Thus, in the first two paragraphs, it is shown in two ways that the solution of the linear filtering and prediction problem stated in Section 1.2 is provided by the system

$$\dot{\hat{\underline{x}}}(t|t) = F(t)\hat{\underline{x}}(t|t) + K(t)\underline{r}(t) , \tag{4.2}$$

where $\hat{\underline{x}}(t_0|t_0) = E[\underline{x}(t_0)]$.

The optimal gain $K(t)$ is given by

$$K(t) = P(t|t)H^T(t)R^{-1}(t) , \tag{4.3}$$

where the error covariance matrix

$$P(t|t) \triangleq E\{[\underline{x}(t) - \hat{\underline{x}}(t|t)][\underline{x}(t) - \hat{\underline{x}}(t|t)]^T\} \tag{4.4}$$

is described by the matrix-Ricatti equation

$$\frac{d}{dt}P = F(t)P + PF^T(t) - PH^T(t)R^{-1}(t)H(t)P + Q(t) , \tag{4.5}$$

where $P(t_0|t_0) = M$.

Some additional aspects of the filtering problem are considered in Section 2.1.3. First, the concept of orthogonal projections is introduced and it is proven that the estimate that is formed as a linear function of data minimizes the mean-square error if, and only if, the error in the estimate is orthogonal to the data. This result is used to derive the Wiener-Hopf equation for the system (Equations (1.1) and (1.2)).

Next, the original problem is generalized to allow cross-correlation $C(t)$ between the plant and measurement noise processes to be non-zero. It is shown that the correlation causes the gain matrix $K(t)$ to become

$$K(t) = [P(t|t)H^T(t) + C(t)]R^{-1}(t) \tag{4.6}$$

and the error covariance matrix is modified to become

$$\dot{P} = F(t)P - PF^T(t) - K(t)[C^T(t) + H(t)P] + Q(t) . \tag{4.7}$$

The effect of a deterministic forcing function $\underline{d}(t)$ in the plant is considered and it is shown that this term requires a trivial change to the estimate. Equation (4.2) is modified to

$$\dot{\hat{\underline{x}}} = F(t)\hat{\underline{x}} + \underline{d}(t) + K(t)\underline{r}(t) . \tag{4.8}$$

Finally Section 2.1.3 is concluded by proving that the conditional mean $E[\underline{x}(t)|\underline{Z}(\tau)]$ provides the unbiased, minimum variance estimate for a random variable $\underline{x}(t)$. This result is useful in giving a probabilistic interpretation of minimum variance estimates. For the linear problem described in Section 1.2, it follows from this property that the unbiased minimum variance estimate of $\underline{x}(t)$ for the system (Equations (1.1) and (1.2)) is actually a linear estimate, thereby confirming that the linear Kalman-Bucy filter provides optimal estimate even when nonlinear estimates are considered.

The time-discrete system equivalent to the continuous-time model described by Equations (1.1) and (1.2) is introduced in Section 2.1.4 and the filter equations for this system are stated without proof.

Section 2.2: The Colored-Noise Problem

In Section 2.2, the white-noise assumptions of Section 1.2 are relaxed to allow the plant and measurement noise processes to exhibit correllation between different times (i.e. the processes are colored). The restriction that the process, say $\underline{n}(t)$, can be described by a linear shaping filter

$$\dot{\underline{n}} = A(t)\underline{n} + \underline{v}(t) \tag{4.9}$$

is introduced, however. The $\underline{y}(t)$ appearing in Equation (4.9) is a white-noise process. The class of problems with colored-noise $\underline{n}$ described by a shaping filter (Eqn (4.9)) is treated by augmenting the noise variables to the state. It is seen that this approach is satisfactory for colored plant noise but is inadequate when there are measurement data which do not contain white-noise.

In Section 2.2.2 the solution for the colored measurement noise problem is determined by noting that the treatment of data without white-noise is accomplished considering a transformation of the state that allows the dimension of the filter equations to be reduced. Several special cases are treated that have appeared in the literature.

Section 2.3: Linear Smoothing

To complete the discussion of linear estimation theory, the linear smoothing problem is solved using the innovations (or residual) process $\underline{\gamma}(t)$ introduced in Section 2.1.2. It is shown that the estimate of $\underline{x}(t)$ given the data $Z(\tau)$, $\tau > t$, has the general form

$$\underline{\hat{x}}(t|\tau) = \underline{\hat{x}}(t|t) + P(t|t)\int_t^\tau \Phi^T(s,t)H^T(s)R^{-1}(s)\underline{\gamma}(s)\,ds\,, \tag{4.10}$$

with associated error covariance matrix described by

$$P(t|\tau) \triangleq E\{[\underline{x}(t) - \underline{\hat{x}}(t|\tau)][\underline{x}(t) - \underline{\hat{x}}(t|\tau)]^T\} = P(t|t) - P(t|t)M_s(t,\tau)P(t|t)\,, \tag{4.11}$$

where

$$M_s(t,\tau) = \int_t^\tau \Psi^T(s,t)H^T(s)R^{-1}(s)H(s)\Psi(s,t)\,ds\,.$$

Equations (4.10) and (4.11) represent the general solution of the smoothing problem. Three classes of smoothing problems, characterized by the time t at which the estimate is to be determined and the interval τ for which data exists, are stated and the solution for each is given.

Section 3.1: Observability of Deterministic System

The concept of observability is introduced in terms of a deterministic linear system obtained from Equations (1.1) and (1.2) by considering the plant noise $\underline{w}(t)$ as known and the measurement noise $\underline{v}(t)$ to be identically zero. If all states $\underline{x}(t)$ of the deterministic system can be determined from knowledge of the input $\underline{w}(\tau)$, $t_0 \leq \tau \leq t$ and the output $\underline{Z}(t)$, then the system is completely observable. It is shown that the system is completely observable if, and only if, the observability matrix

$$M(t_0,t) = \int_{t_0}^t \Phi^T(\tau,t_0)H^T(\tau)R^{-1}(\tau)H(\tau)\Phi(\tau,t_0)\,d\tau \tag{4.12}$$

is positive-definite for some $t > t_0$.

When the linear system is stationary it is shown that the observability criterion can be expressed in terms of the $n \times mn$ matrix

$$U = [H^T F^T H^T \ldots (F^T)^{n-1}H^T]\,. \tag{4.13}$$

The system is completely observable if this matrix has rank n.

In Section 3.2 it is shown that the least-squares estimate of an arbitrary initial state exists if the system is completely observable. That is, a least-squares estimator exists if, and only if, the observability matrix is positive-definite.

Section 3.2: The Matrix-Riccati Equation

The error covariance matrix was found in Section 2 to be described by a matrix-Riccati equation. This nonlinear matrix differential equation is investigated in this section and it is shown that it can be rewritten by a system of linear matrix differential equations. Thus, the matrix-Riccati equation

$$\dot{P} = F(t)P + PF^T(t) - PH^T(t)R^{-1}(t)H(t)P + Q(t)\,; \qquad P(t_0) = P_0 \tag{4.14}$$

can be written as

$$\dot{Y} = F(t)Y + Q(t)Z\,; \qquad Y(t_0) = P_0 \tag{4.15}$$

$$\dot{Z} = H^T(t)R^{-1}(t)H(t)Y - F^T(t)Z\,; \qquad Z(t_0) = I\,. \tag{4.16}$$

The solution of Equation (4.14) is obtained from the solutions of Equations (4.15) and (4.16) as

$$P(t) = Y(t)Z^{-1}(t) . \tag{4.17}$$

More explicit solutions can be obtained when either $H(t)$ or $Q(t)$ vanish. These solutions are derived. Also, additional insights are possible into the character of Equation (4.14) when the system is stationary. This special case is discussed in some detail.

Section 3.3: Bounds for the Error Covariance Matrix

The error covariance matrix provides the primary statistical measure of the performance of the estimator. This statistic is actually independent of the measurement data so it can be investigated before the Kalman-Bucy filter equations are used to process data for a given system. Upper and lower bounds for this matrix are established in this system in such a manner that the fundamentally different roles of the plant noise and the measurement noise processes are clearly displayed. It is shown that $P(t|t)$ is bounded below by

$$P(t|t) \geq \Phi(t,t_0)[P_0^{-1} + M(t_0,t)]^{-1}\Phi^T(t,t_0) + P^p(t|t) , \tag{4.18}$$

where $P^p(t|t)$ is a non-negative-definite (never positive-definite) matrix and the first term vanishes as t becomes large for q-observable systems. (This term is defined in Section 3.3.) The upper bound is given by

$$P(t|t) \leq \Phi(t,t_0)[P_0^{-1} + M(t_0,t)]^{-1}\Phi^T(t,t_0) + W(t_0,t) , \tag{4.19}$$

where the stochastic controllability matrix $W(t_0,t)$ is defined to be

$$W(t_0,t) \triangleq \int_{t_0}^{t} \Phi(t,\tau)Q(\tau)\Phi^T(t,\tau)\, d\tau .$$

Section 3.4: Behavior of the Error Covariance for Stationary Systems

The error covariance matrix for stationary systems can be studied in more detail using the steady-value $\lim P(t|t) = P_\infty$. A suboptimal gain K_W equivalent to the Wiener filter gain can be defined that is useful in examining the influences that affect the relative performance of the optimal and Wiener filters. Letting $P_W(t|t)$ denote the error covariance matrix of a filter using the constant Wiener gain, it is shown that the difference is bounded by

$$\|P_W(t|t) - P(t|t)\| \leq \frac{\|\delta_0\|^2 \|H^T R^{-1} H\|}{8|\lambda_1|} \qquad \text{for all } t , \tag{4.20}$$

where $\delta_0 = P(t_0) - P_\infty$, and where $\|H^T R^{-1} H\|$ represents the signal-to-noise ratio of the measurement system and $|\lambda_1|$ is the smallest time constant of the filter dynamics.

This concludes a summary of the presentation found in Sections 2 and 3. Many topics have been omitted from this discussion and other topics have been dealt with somewhat superficially. Many of these omissions and superficialities are given their deserved airing in subsequent chapters and their variety and number provide an indication of the breadth, depth, and importance of unbiased, minimum variance linear estimation theory in engineering theory and application.

REFERENCES

1. Gauss, K.F. — *Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections.* Dover, New York, 1963.

2. Wiener, N. — *The Extrapolation, Interpolation and Smoothing of Stationary Time Series.* Wiley, New York, 1949.

3. Kolmogorov, A.N. — *Interpolation und Extrapolation von Stationären Zufälligen Folgen.* Bulletin of the Academy of Sciences, USSR, Mathematical Series, Vol.5, 1941, pp.3-14.

4. Bode, H.W.; Shannon, C.E. — *A Simplified Derivation of Linear Least-Squares Smoothing and Prediction Theory.* Proceedings, Institute of Radio Engineers, Vol.38, 1950, pp.417-425.

5. Zadeh, L.A.; Ragazzini, J.R. — *An Extension of Wiener's Theory of Prediction.* Journal of Applied Physics, Vol.21, 1950, pp.645-655.

6. Shinbrot, M. — *Optimization of Time-Varying Linear Systems with Nonstationary Inputs.* Transactions, American Society of Mechanical Engineers, Vol.80, 1958, pp.457-462.

7. Hsieh, H.; Leondes, C.T. — *On the Optimum Synthesis of Multipole Control Systems in the Wiener Sense.* Institute of Radio Engineers, Transactions on Automatic Control, Vol.AC-4, No.2, Nov.1959, pp.18-29.

8. Carlton, A.G. — *Linear Estimation in Stochastic Processes.* The Johns Hopkins University, Applied Physics Laboratory, Bumblebee Series Report 311, Baltimore, Maryland, 1962.

9. Carlton, A.G.; Follin, J.W., Jr. — *Recent Developments in Fixed and Adaptive Filtering.* AGARDograph 21, 1956.

10. Swerling, P. — *First-Order Error Propagation in a Stage-Wise Smoothing Procedure for Satellite Observations.* Journal of the Astronautical Sciences, Vol.6, 1959, pp.46-52.

11. Kalman, R.E. — *A New Approach to Linear Filtering and Prediction Problems.* Journal of Basic Engineering, Vol.82D, 1960, pp.95-108.

12. Kalman, R.E.; Bucy, R.S. — *New Results in Linear Filtering and Prediction Theory.* Journal of Basic Engineering, Vol.83D, 1961, pp.95-108.

13. Kalman, R.E. — *New Methods in Wiener Filtering Theory.* Proceedings of the First Symposium of Engineering Application of Random Function Theory and Probability. Wiley, New York, 1963, pp.270-388.

14. Coddington, E.A.; Levinson, N. — *Theory of Ordinary Differential Equations.* McGraw-Hill, 1956.

15. Bucy, R.S.; Joseph, P.D. — *Filtering for Stochastic Processes with Applications to Guidance.* Interscience Publishers, New York, 1968.

16. Kailath, T. — *An Innovations Approach to Least-Squares Estimation. Part I: Linear Filtering in Additive White-Noise.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-13, No.6, Dec.1968, pp.646-655.

17. Athans, M.; Tse, E. — *A Direct Derivation of the Optimal Linear Filter Using the Maximum Principle.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-12, No.6, Dec.1967, pp.690-698.

18. Reid, W.T. — *A Matrix Differential Equation of Riccatti Type.* American Journal of Mathematics, Vol.68, 1946, pp.237-246.

19. Sorenson, H.W. — *Kalman Filtering Techniques.* Chapter 5 of "Advances in Control Systems", Vol.3, C.T.Leondes, Editor, Academic Press, New York, 1966.

20. Sorenson, H.W. — *Controllability and Observability of Linear, Stochastic, Time-Discrete Control Systems.* Chapter 2 of "Advances in Control Systems", Vol.6, C.T.Leondes, Editor, Academic Press, New York, 1968.

21. Stear, E.B. — *Synthesis of Shaping Filters for Nonstationary Stochastic Processes and Their Use.* University of California, Report 61-50, Los Angeles, California, 1961.

22. Cox, H. — *Estimation of State Variables via Dynamic Programming.* 1964 JACC Proceedings, June 1964, pp.376-381.

23. Bryson, A.E. Johansen, D.E. — *Linear Filtering for Time-Varying Systems Using Measurements Containing Colored-Noise.* Institute of Electrical and Electronic Engineers. Transactions on Automatic Control, Vol.AC-10, Jan.1965, pp.4-10.

24. Buoy, R.S. — *Optimal Filtering for Colored-Noise.* Journal of Mathematical Analysis and its Applications, Vol.20, Oct.1967, pp.1-8.

25. Stear, E.B. Stubberud, A.R. — *Optimal Filtering for Gauss-Markov Noise.* International Journal of Control, Vol.8, No.2, 1968, pp.123-130.

26. Sarachik, P.E. — *State Estimation from Measurements with Correlated Noise without Using Differentiators.* To appear in Institute of Electrical and Electronic Engineers, Transactions on Automatic Control.

27. Kailath, T. Frost, P. — *An Innovation Approach to Least-Squares Estimation. Part II: Linear Smoothing in Additive White-Noise.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-13, No.6, Dec.1968, pp.655-660.

28. Meditch, J.S. — *On Optimal Linear Smoothing Theory.* Journal of Information and Control, Vol.10, 1967, pp.598-615.

29. Bryson, A.S. Frazier, M. — *Smoothing for Linear and Nonlinear Dynamic Systems.* Aeronautics Systems Division, Wright-Patterson AFB, Ohio, Technical Report ASD-TDR-63-119, February 1969.

30. Meditch, J.S. — *On Optimal Fixed-Point Linear Smoothing.* International Journal of Control, Vol.6, 1967, pp.189-199.

31. Kalman, R.E. — *Mathematical Description of Linear Dynamical Systems.* Journal SIAM Control, Series A, Vol.1, No.2, 1963, pp.152-192.

32. Sorenson, H.W. — *On the Error Behavior in Linear Minimum Variance Estimation Problems.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-12, No.5, 1967, pp.557-562.

33. O'Donnell, J.J. — *Asymptotic Solution of the Matrix Ricatti Equation of Optimal Control.* Proceedings, 4th Annual Allerton Conference on Circuit and System Theory, pp.577-586.

34. Vaughn, D.R. — *A Negative Exponential Solution for the Matrix Ricatti Equation.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-14, No.1, Feb.1969, pp.72-75.

35. Singer, R.A. — *The Design and Synthesis of Linear Multivariable Systems with Application to State Estimation.* Stanford Electronics Laboratory Report 6302-8, Stanford, 1968.

36. Denham, W.F. Pines, S. — *Sequential Estimation when Measurement Function Nonlinearity is Compurable to Measurement Error.* AIAA Journal, Vol.4, No.6, June 1966, pp.1071-1076.

37. Schlee, F.H. et al. — *Divergence in the Kalman Filter.* AIAA Journal, Vol.5, No.6, June 1967, pp.1114-1120.

CHAPTER 2 - FURTHER COMMENTS ON THE DERIVATION OF KALMAN FILTERS

SECTION I - DERIVATION OF THE KALMAN FILTERING EQUATIONS FROM ELEMENTARY STATISTICAL PRINCIPLES

by

P.M.Barham and D.E.Humphriva

Avionics Department, Royal Aircraft Establishment,
Farnborough, Hampshire, England

CHAPTER 2 – FURTHER COMMENTS ON THE DERIVATION OF KALMAN FILTERS

SECTION 1 – DERIVATION OF THE KALMAN FILTERING EQUATIONS FROM ELEMENTARY STATISTICAL PRINCIPLES

P.M. Barham and D.E. Humphries

1. INTRODUCTION

In less than a decade Kalman filtering has become an established technique within the resources of the Aerospace systems designer. Unfortunately the fact that it is an advanced statistical concept, mainly described by mathematicians on the assumption that the reader is familiar with advanced statistical theory, has too often had the effect that the practical engineer has accepted a set of equations while feeling insufficiently qualified to attempt a full appreciation of their significance. The aim in this chapter is to provide a simple, though rigorous, derivation of the Kalman filtering equations in a form which should appeal to the less mathematical reader. For this reason more steps have been included in the derivation of the equations than would be necessary for those fully familiar with statistical methods.

The necessary statistical theory is developed in Section 2 as an extension of familiar and elementary sampling theory. Whereas a knowledge of simple matrix notation and manipulation is unavoidable for the development and application of the practical multi-dimensional filter, the theory is first developed for a simple single-dimension case which leads directly from the concepts of mean value and standard deviation.

2. OPTIMUM COMBINATION OF INDEPENDENT ESTIMATES

The Kalman filtering process consists of combining two independent estimates of a variable to form a weighted mean. The weighting factor is chosen to yield a mean with minimum variance and hence maximum probability. One of the estimates is derived by updating a previous best estimate in accordance with the known equations of motion and the other estimate is obtained from a measurement. The form of the required weighting factor is derived below, first for a single-dimension case and secondly for the general multi-dimensional form required for the Kalman filter.

2.1 The Single-Dimension Case

Let x_1 and x_2 be two independent estimates of a quantity x with variances σ_1^2 and σ_2^2 respectively. It is required to combine these estimates to form a weighted mean corresponding to the overall best estimate ($\bar{x}$) where "best estimate" means the minimum variance estimate.

Now the general form of the weighted mean of x_1 and x_2 is

$$\bar{x} = (1-w)x_1 + wx_2 . \tag{2.1}$$

Thus the expected or mean value of $\bar{x}$ [written $E(\bar{x})$] is given by

$$E(\bar{x}) = (1-w)\,E(x_1) + wE(x_2) . \tag{2.2}$$

By definition the variance of a quantity x is

$$E[(x-E(x))^2] . \tag{2.3}$$

Hence the variance (σ^2) of $\bar{x}$ is given by

$$\begin{aligned}\sigma^2 &= E[(\bar{x}-E(\bar{x}))^2] = E[((1-w)x_1 + wx_2 - (1-w)E(x_1) - wE(x_2))^2] \\ &= E[(1-w)^2(x_1-E(x_1))^2 + w^2(x_2-E(x_2))^2 - 2w(1-w)(x_1-E(x_1))(x_2-E(x_2))] \\ &= (1-w)^2E[(x_1-E(x_1))^2] + w^2E[(x_2-E(x_2))^2]\end{aligned}$$

or

$$\sigma^2 = (1-w)^2\sigma_1^2 + w^2\sigma_2^2 \tag{2.4}$$

since

$$E[(x_1-E(x_1))(x_2-E(x_2))] = 0 ,$$

as x_1 and x_2 are independent estimates so that $(x_1-E(x_1))$ and $(x_2-E(x_2))$ are uncorrelated. To determine the value of w for which σ^2 is minimum, Equation (2.4) may be partially differentiated with respect to w. Thus

$$\frac{\partial\sigma^2}{\partial w} = -2(1-w)\sigma_1^2 + 2w\sigma_2^2 = 0 .$$

giving the optimum value θ of weighting factor as

$$\theta = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} . \tag{2.5}$$

Substitution in Equation (2.1) gives

$$\hat{x} = \frac{\sigma_2^2 x_1 + \sigma_1^2 x_2}{\sigma_1^2 + \sigma_2^2} , \tag{2.6}$$

and in Equation (2.4) gives the variance of $\hat{x}$ as

$$\hat{\sigma}^2 = \frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2} . \tag{2.7}$$

If x_2 is considered to be a measurement used to improve an updated esimtate x_1, Equations (2.5) and (2.7) can be put into the more convenient forms

$$\hat{x} = x_1 - \theta(x_1 - x_2) \tag{2.8}$$

and

$$\hat{\sigma}^2 = \sigma_1^2(1-\theta) , \tag{2.9}$$

which more clearly show how the estimate and its variance are "improved" by the measurement in readiness for further updating. Equation (2.8) also shows the close analogy to negative feedback.

2.3 The Multi-Dimensional Case

Let X_1 and X_2 be two $n \times 1$ matrices representing two independent estimates (i.e. with uncorrelated errors) of an n-dimensional vector quantity X. P_1 and P_2 are two $n \times n$ matrices representing the variances of X_1 and X_2 respectively.

In general the weighted mean of X_1 and X_2 is given by

$$\bar{X} = (I - W)X_1 + WX_2$$

or

$$\bar{X} = X_1 - W(X_1 - X_2) , \tag{2.10}$$

where W is an arbitrary $n \times n$ weighting matrix and I is the unit matrix. The best estimate of X (denoted by $\hat{X}$) will be given by (2.10) when W is chosen so that the variance of $\bar{X}$ is a minimum in the sense considered below.

In many practical cases, however, the two estimates are not of equal dimensions, one of them being some function (M, say) of the individual elements of X (for example, a measurement relating to some only of the elements of X). In general M may be an $m \times n$ rectangular matrix operating on X to give an m-dimensional estimate Y.

The general problem therefore is that of forming an optimum estimate ($\hat{X}$) of X from one estimate (X_1 with variance P') of X and an estimate (Y_2 with variance V) of $Y(=MX)$. If we let the weighting matrix in (2.10) be $W = KM$, where K is another arbitrary weighting matrix, then

$$\begin{aligned} \bar{X} &= X_1 - KM(X_1 - X_2) \\ &= X_1 - K(MX_1 - Y_2) \\ &= (I - KM)X_1 + KY_2 . \end{aligned} \tag{2.11}$$

By definition the variance $\bar{P}$ of $\bar{X}$ is given by

$$\bar{P} = E\{[\bar{X} - E(\bar{X})][\bar{X} - E(\bar{X})]^T\}$$

and similar expressions may be written for P' and V, the variances of X_1 and Y_2 respectively, where, as in Section 2.1, E denotes the expected or mean value. Substituting for $\bar{X}$ from Equation (2.11) we have

$$\begin{aligned} \bar{P} &= E\{[(I - KM)X_1 + KY_2 - (I - KM)E(X_1) - KE(Y_2)][(I - KM)X_1 + KY_2 - (I - KM)E(X_1) - KE(Y_2)]^T\} \\ &= (I - KM)E\{[X_1 - E(X_1)][X_1 - E(X_1)]^T\}(I - KM)^T + KE\{[Y_2 - E(Y_2)][Y_2 - E(Y_2)]^T\}K^T , \end{aligned}$$

since X_1 and Y_2 are uncorrelated. Thus

$$\bar{P} = (I - KM)P'(I - KM)^T + KVK^T . \tag{2.12}$$

It is now required to find the value of K which minimises $\bar{P}$ in the sense that each diagonal element of $\bar{P}$ shall be a minimum variance. Expanding Equation (2.12) gives

$$\begin{aligned} \bar{P} &= (I - KM)P'(I - KM)^T + KVK^T \\ &= P' + (KM)P'(KM)^T - (KM)P' - P'(KM)^T + KVK^T \\ &= P' + K(MP'M^T + V)K^T - K(MP') - (MP')^TK^T , \end{aligned}$$

since $(AB)^T = B^TA^T$ and P' is a symmetrical matrix so that $P'^T = P'$. Now consider

$$\begin{aligned}(KS-A)(KS-A)^T &= K(SS^T)K^T - KSA^T - AS^TK^T + AA^T \\ &= K(SS^T)K^T - K(SA^T) - (SA^T)^TK^T + AA^T \\ &= K(MP'M^T+V)K^T - K(MP') - (MP')^TK^T + AA^T ,\end{aligned}$$

if $$SS^T = MP'M^T + V$$

and $$SA^T = MP' ,$$

whence $$\begin{aligned}\bar{P} &= P' + (KS-A)(KS-A)^T - AA^T \\ &= P' + (KS-A)(KS-A)^T - P'M^T(MP'M^T+V)^{-1}MP' . \end{aligned} \tag{2.13}$$

The only term in this expression for $\bar{P}$ which is dependent on K is $(KS-A)(KS-A)^T$. As this is of the form BB^T, the diagonal elements consist of the sums of squares of the individual elements of $(KS-A)$, and, as such, cannot have negative values. Thus each diagonal element of $\bar{P}$ shall be a minimum variance if K takes the value which makes $(KS-A) = 0$; that is, if

$$K = P'M^T[MP'M^T+V]^{-1} . \tag{2.14}$$

Substituting in Equation (2.13) with $(KS-A) = 0$ gives

$$P = \bar{P} = P' - KMP' \tag{2.15}$$

as the variance for the best estimate of X which, from Equation (2.11), becomes

$$\bar{\bar{X}} = X_1 - K(MX_1 - Y_2) , \tag{2.16}$$

where K takes the value given in Equation (2.14).

3. REPRESENTATION OF PHYSICAL PROCESSES

Classical physics tells us that, in the absence of external disturbance, the future behaviour of a system may be derived from its present state by application of the known equations of motion. It is now realised that, due to the quantum nature of fundamental physical processes, this is only effectively true in the macro case when the quantum noise can be ignored and only the statistical mean need be considered. Where this is not the case the expected or mean state of the system may be predicted by the equations of motion, but imposed upon this will be a random contribution due to quantum events, whose probability distribution only will be known.

Similar considerations apply to any external disturbance. Some of these are predictable and can, if desired, be included as parameters within our chosen system; this will also apply to any bias which may be present in otherwise unpredictable disturbances. It is thus seen that the behaviour of any physical system may be considered to consist of one part which is precisely predictable from the known equations of motion and a second part which is random about zero mean, but whose probability distribution may be determined. Such a process, in which the statistics of the future behaviour of a system are purely a function of its state at a particular time and the statistical structure of the perturbing process, is termed a Markov Process. It is clear that, by a suitable choice of the variables included, any physical process can be represented in this way.

Assuming a Markov process of this kind, its true state X_{k+1} at time t_{k+1} may be represented as

$$X_{k+1} + \Phi_kX_k + G_kU_k , \tag{3.1}$$

where X_k was its true state at time t_k. Here X_k and X_{k+1} are $n \times 1$ matrices representing the n parameters necessary to define the state of the system (termed the state vector).

Φ_k is an $n \times n$ matrix representing the known dynamics or equations of motion of the system (termed the transition matrix).

U_k represents the random inputs (with zero means) to the system. In general, if there are l sources of random input, U_k will be an $l \times 1$ matrix and is known as plant noise.

G_k is an $n \times l$ matrix representing the effects of the l elements of noise on the n parameters of the system.

The choice of elements to include in the state vector is, in practice, somewhat arbitrary. Although theoretically X_k should include all the systematic elements and U_k all the random elements affecting the system, this itself is a choice dependent upon the state of the art. Many effects which were considered random or unpredictable a century ago are now well understood processes. On the other hand some influences, although fully understood and predictable, may be so remote from the system under consideration that their effects may conveniently be included as part of the random input.

The representation of a Markov process by Equation (3.1) does not necessitate a stationary random process. The form of Φ_k and G_k and the statistical structure of U_k (i.e. its variance Q_k) may vary from step to step. Indeed the requirement is not so much that they should be constant even within a single step, provided that the

integrated effect shall be known for each individual step. In most cases, of course, the step size will be chosen to ensure that the transition matrix, the variance of the plant noise and its effect on the system may all be considered constant. Finally is should be noted that, as the step size approaches the zero limit, Equation (3.1) collapses to the more familiar differential form.

4. DERIVATION OF THE KALMAN FILTER

It is necessary to consider three quite distinct sets of equations in order to devise the Kalman filter. First we must represent the true state of the system in order to be able to study its statistical structure. Secondly, we must consider how we may forecast the state of the system from a previous best estimate and what variance this forecast will have. Finally we must determine a best estimate of the updated state, together with its variance, by suitably combining its forecast value with a measurement.

In this and subsequent sections true values will be represented by non-accented characters such as X_k , a forecast value which has not been subjected to a measurement will be represented by a prime (e.g. X'_{k+1}) and a best estimate by a hat (e.g. $\hat{X}_{k+1}$). Where a measurement is available this last will be obtained by combining the forecast with the measurement. In the absence of a measurement the forecast value will be the best estimate.

4.1 The True State Equations

We have already shown in Section 3 that the true state X_{k+1} at time t_{k+1} is derived from its state X_k at time t_k by the equation

$$X_{k+1} = \Phi_k X_k + G_k U_k \,. \tag{4.1}$$

Here Φ_k , the transition matrix, is known, as is G_k representing the effect of a noise input on the elements of the state vector. We know that statistically U_k has a zero mean and variance Q_k , but we have no knowledge of the actual noise contribution U_k to this particular step.

Similarly, when we make a measurement, the quantity measured will be some function M_{k+1} of the individual elements of the state vector, but the actual reading will also have a noise contribution N_{k+1} . Thus the actual quantity determined by the measurement will be in the form

$$Y_{k+1} = M_{k+1} X_{k+1} + N_{k+1} \,. \tag{4.2}$$

M_{k+1} is, in general, an $m \times n$ matrix which operates on the n elements of X_{k+1} to give the $m \times 1$ matrix representing the m elements of the measurement Y_{k+1} . In most cases M_{k+1} will be a simple row matrix yielding a single parameter for measurement and frequently all elements of the row except one will be zeros, so that the effect of M_{k+1} will be to isolate a single element of X_{k+1} . N_{k+1} is the "measurement noise" and has zero mean with variance R_{k+1} ,but the actual contribution to a particular measurement is unknown.

We are then able to describe the true behaviour of the system by Equation (4.1) and of the measurement process by (4.2).

4.2 The Forecasting Process

If the error in $\hat{X}_k$ (the best estimate of the true state X_k) is ϵ_k , then

$$\hat{X}_k = X_k + \epsilon_k \,.$$

Since the plant noise U_k has zero mean, it is easily seen that the best forecast is simply

$$\begin{aligned} X'_{k+1} &= \Phi_k \hat{X}_k \\ &= \Phi_k X_k + \Phi_k \epsilon_k \\ &= X_{k+1} - G_k U_k + \Phi_k \epsilon_k \,. \end{aligned} \tag{4.3}$$

Now ϵ_k , the error of $\hat{X}_k$, will have the same variance as $\hat{X}_k$ itself (i.e. P_k) and hence the variance of $\Phi_k \epsilon_k$ is $\Phi_k P_k \Phi_k^T$. Similarly the variance of $G_k U_k$ is $G_k Q_k G_k^T$, where Q_k is the variance of U_k . Thus, since X_{k+1} is a true value, the error of X'_{k+1} is represented by the last two terms of Equation (4.3). Since these two error terms (original error and error increment) are independent, the variance of X'_{k+1} is given by

$$P'_{k+1} = \Phi_k P_k \Phi_k^T + G_k Q_k G_k^T \,. \tag{4.4}$$

4.3 The Measurement Process

The measurement process consists of a determination of Y_{k+1} , which is defined in Equation (4.2), and comparing this with a forecast (X'_{k+1}) in order to obtain a best estimate ($\hat{X}_{k+1}$).. Now using Equations (2.14), (2.15) and (2.16) the best estimate of X_{k+1} will be

$$\hat{X}_{k+1} = X'_{k+1} - K_{k+1} (M_{k+1} X'_{k+1} - Y_{k+1}) \,. \tag{4.5}$$

where

$$K_{k+1} = P'_{k+1}M^T_{k+1}[M_{k+1}P'_{k+1}M^T_{k+1} + R_{k+1}]^{-1} \tag{4.6}$$

and R_{k+1} is the variance of Y_{k+1}.

The variance of $\hat{X}_{k+1}$ is given by

$$P_{k+1} = P'_{k+1} - K_{k+1}M_{k+1}P'_{k+1} . \tag{4.7}$$

5. CONCLUSION

If the true state of a Markov process and a measurement are defined by the equations

$$X_{k+1} = \Phi_k X_k + G_k U_k . \tag{5.1}$$

where X_k is the state vector at time t_k

Φ_k is a transition matrix representing the equation of motion

U_k is the plant noise with variance Q_k

G_k represents the effect of the noise input on the state vector and

$$Y_{k+1} = M_{k+1}X_{k+1} + N_{k+1} . \tag{5.2}$$

where Y_{k+1} is the quantity measured

M_{k+1} represents its relationship with X_{k+1}

N_{k+1} is the measurement noise with variance R_{k+1}.

then the best estimate $\hat{X}_k$ of the system and its variance P_k can be updated by the equations

$$X'_{k+1} = \Phi_k \hat{X}_k \tag{5.3}$$

$$P'_{k+1} = \Phi_k P_k \Phi^T_k + G_k Q_k G^T_k \tag{5.4}$$

$$K_{k+1} = P'_{k+1}M^T_{k+1}[M_{k+1}P'_{k+1}M^T_{k+1} + R_{k+1}]^{-1} \tag{5.5}$$

$$\hat{X}_{k+1} = X'_{k+1} - K_{k+1}(M_{k+1}X'_{k+1} - Y_{k+1}) \tag{5.6}$$

and

$$P_{k+1} = P'_{k+1} - K_{k+1}M_{k+1}P'_{k+1} . \tag{5.7}$$

ACKNOWLEDGEMENT

The authors wish to record their debt to Mr D.E.Williams of Mathematics Department, Royal Aircraft Establishment, for his helpful criticism of this work.

CHAPTER 2 - FURTHER COMMENTS ON THE DERIVATION OF KALMAN FILTERS

SECTION II - GAUSSIAN ESTIMATES AND KALMAN FILTERING

by

Y. Genin*

Applied Mathematics Group
MBLE Research Laboratories

* 2 Avenue van Becelaere, Brussels 17, Belgium

NOTATION

e_i	generalized noise vector random variable corrupting the generalized measurement z^i of the state x_i
e_i'	generalized noise vector random variable corrupting the generalized measurement z^{i-1} of the state x_i
f_i	deterministic vector forcing function in discrete, linear dynamical system at t_i
q_j	eigenvector corresponding to eigenvalue λ_j of matrix M
t_i	i^{th} sampling time
v	random noise vector variable corrupting the measurement z of the constant state x
v_k	k^{th} random vector variable formed by partioning v in subvectors
v_i	white noise sequence corrupting measurement z_i at t_i
v^i	generalized random vector variable formed by grouping the v_i for i running from 1 up to i in one vector
w_{i-1}	white noise sequence affecting discrete, linear dynamical system at t_{i-1}
w^{i-1}	generalized random vector variable formed by grouping the w_{i-1} for i running from 1 up to i in one vector
x	constant state vector
$\hat{x}_k$	estimate of x based upon the k first measurement subvectors z_k
x_i	state of discrete, linear dynamical system at t_i
x^i	generalized state vector formed by grouping the x_i for i running from 1 up to i in one vector
$\hat{x}_i$	estimate of x_i based upon z^i
$\hat{x}_i'$	estimate of x_i based upon z^{i-1} (extrapolation of $\hat{x}_{i-1}$)
y_i	auxiliary vector defined by $\hat{x}_i = P_i y_i$
y_i'	extrapolation of y_{i-1} from t_{i-1} to t_i
z	measurement vector of the state x
z_k	k^{th} measurement subvector formed by partitioning z in subvectors
z_i	measurement vector of the state x_i at time t_i
z^i	generalized measurement vector formed by grouping the z_i for i running from 1 up to i in one vector
C_i	covariance matrix of e_i , $E(e_i e_i^T)$
C_i'	covariance matrix of e_i' , $E[(e_i')(e_i')^T]$
D_i	gain matrix for accumulative filter
P^i	auxiliary matrix defining e^i as function of w^{i-1} and v^i , $e_i = -H^i P^i w^{i-1} + v^i$
H	observation matrix of state x
H_k	submatrix of H defined by $z_k = H_k x_k + v_k$
H_i	observation matrix of state x_i at t_i
H^i	generalized observation matrix
K_k	gain matrix for Kalman filter
M	non-negative definite matrix
P	covariance matrix of the error $(\hat{x} - x)$

P_k	covariance matrix of the error $(\hat{x}_k - x)$
P_i	covariance matrix of the error $(\hat{x}_i - x_i)$
P_i'	covariance matrix of the error $(\hat{x}_i' - x_i)$
Q_{i-1}	covariance matrix of w_{i-1}
Q^{i-1}	covariance matrix of w^{i-1}
R	covariance matrix of v
R_k	covariance matrix of v_k
R_i	covariance matrix or v_i
R^i	covariance matrix of v^i
W	positive definite weighting matrix
λ_j	j^{th} eigenvalue of M
$\Phi_{i,i-1}$	state transition matrix from t_{i-1} to t_i
Φ^i	generalized state transition matrix
I	identity matrix
0	zero vector
0	zero matrix
$()^T$	transpose of ()
$()^{-1}$	inverse of ()
$E()$	mean value of ()
δ_{ki}	Kronecker delta
$\sum_{j=1}^{n}$	summation symbol for j running from 1 up to n

CHAPTER 2 - FURTHER COMMENTS ON THE DERIVATION OF KALMAN FILTERS

SECTION II - GAUSSIAN ESTIMATES AND KALMAN FILTERING

Y.Genin

1. INTRODUCTION

The aim of this contribution is twofold. First, it will give the reader, who is familiar with mean squares techniques, sufficient background to gain easy access to Kalman filter theory. To this end, the intimate connection between the Kalman and the Gaussian estimation theory will be pointed out: more precisely, it will be shown that the application of results obtained by Gauss, at least in essence, to discrete stochastic linear processes leads to the Kalman filter.

The first part recalls the main points of the Gauss theory[6], as slightly extended by Plackett[12]. The core of this section is a restatement of two classical theorems, which are of primary importance in modern linear filtering techniques: the first constitutes the cornerstone of recursive filtering, while the second offers its dual form, i.e. accumulative filtering.

The second part is devoted to the application of these theorems to discrete stochastic processes and shows how the Kalman filter and its dual form fit naturally in the framework provided by Gauss.

2. THE GAUSSIAN LEAST SQUARES THEORY

2.1 Statement of the Problem

Consider the following equation of measurements:

$$z = Hx + v \,, \tag{1}$$

where z is a N-dimensional vector of measurements, x a n-dimensional constant state vector to be determined ($n < N$), H a $N \times n$ matrix of maximal rank and v a N-dimensional vector of errors with zero mean and positive definite covariance matrix R.

The problem is to find an estimate $\hat{x}$ of the unknown state x as a linear combination of the measurements z,

$$\hat{x} = Az \,, \tag{2}$$

so that the estimate $\hat{x}$ be unbiased and have the smallest variance for each of its elements.

2.2 A Deterministic Problem

Let us consider a particular problem, in which Equation (1) holds[6]. The z vector is known and is assumed to take the value Z. Let us temporarily disregard probabilistic and statistical considerations and seek a "good" approximation X of the state x, in the form of a linear combination of the known measurements Z. If we arbitrarily decide to measure the quality of the approximation by adopting the usual mean squares criterion

$$(Z - HX)^T W (Z - HX) = \text{minimum} \,, \tag{3}$$

with W a positive definite weighting matrix, the best approximation is readily found to be

$$X = (H^T W H)^{-1} H^T W Z \,, \tag{4}$$

which solves this simple deterministic problem.

2.3 Theorem 1

Coming back to our main problem (Section 2.1), we shall now prove the following fundamental theorem. *The unbiased estimate $\hat{x}$, which is a linear combination of the measurements z and has the smallest variance for each of its components, is given by*

$$\hat{x} = (H^T R^{-1} H)^{-1} H^T R^{-1} z \,. \tag{5}$$

It has the same form as Equation (4), the solution of the deterministic problem in which the weighting matrix W *is set equal to* R^{-1} .

The estimate is readily proved to be unbiased, for its mean value

$$E(\hat{x}) = (H^T R^{-1} H)^{-1} H^T R^{-1} E(z)$$

may be written in view of Equation (1),

$$E(\hat{x}) = (H^T R^{-1} H)^{-1} H^T R^{-1} [Hx + E(v)]$$

$$= x . \tag{6}$$

The covariance matrix P of the estimate $\hat{x}$ is defined by

$$P = E[(\hat{x} - x)(\hat{x} - x)^T]$$

which, in this case, has the form

$$P = (H^T R^{-1} H)^{-1} H^T R^{-1} E(vv^T) R^{-1} H (H^T R^{-1} H)^{-1}$$

$$= (H^T R^{-1} H)^{-1} . \tag{7}$$

Before proving that each diagonal element of the P matrix is minimized by Equation (5), let us first establish the following important lemma. *Lemma. Let* M *be any* $n \times n$ *non-negative definite matrix. Then the unbiased estimate* $\hat{x}$ *, which is a linear combination of the measurements* z *and minimizes the quadratic form*

$$E[(\hat{x} - x)^T M (\hat{x} - x)] \tag{8}$$

is again given by Equation (5).

Suppose this proposition were false and let $\hat{x} = Bz$, the solution with a matrix B different from $(H^T R^{-1} H)^{-1} H^T R^{-1}$. Without restriction, M may be assumed to be symmetric, for (8) is a homogeneous quadratic form. If it has multiple eigenvalues, it may be submitted to the Gauss-Schmidt orthogonalization procedure[3] so that, in any case, (8) may be written

$$\sum_{j=1}^{n} \lambda_j q_j^T E[(\hat{x} - x)(\hat{x} - x)^T] q_j , \tag{9}$$

with q_j the eigenvector corresponsing to the eigenvalue λ_j of the matrix M . Note that the eigenvalues λ_j are all real and non-negative, for M is a non-negative definite real symmetric matrix. On the other hand, since the estimate $\hat{x}$ is unbiased, we have

$$BH = I , \tag{10}$$

with I the identity matrix. The quadratic form (9) is therefore equivalent to

$$\sum_{j=1}^{n} \lambda_j q_j^T R R B^T q_j ,$$

which may be transformed into

$$\sum_{j=1}^{n} \lambda_j q_j^T B R B^T q_j = \sum_{j=1}^{n} \lambda_j q_j^T (H^T R^{-1} H)^{-1} q_j + \sum_{j=1}^{n} \lambda_j q_j^T [B - (H^T R^{-1} H)^{-1} H^T R^{-1}] R [B - (H^T R^{-1} H)^{-1} H^T R^{-1}]^T q_j \tag{11}$$

in view of Equation (10) and the identity

$$B = (H^T R^{-1} H)^{-1} H^T R^{-1} + [B - (H^T R^{-1} H)^{-1} H^T R^{-1}] . \tag{12}$$

Since in the summation (11) each term is non-negative, we are left with

$$B = (H^T R^{-1} H)^{-1} H^T R^{-1} , \tag{13}$$

which proves the lemma.

A direct consequence of the lemma is that the estimate Equation (5) has the smallest variance for each of its elements, as can be seen by choosing all elements of M equal to zero, except for a unit element anywhere on the diagonal.

2.4 Theorem 2 (Recursive Estimation)

Suppose now a partition of the measurements vector z in two subvectors z_{k-1}, z_k. Partitioning v and H accordingly, Equation (1) may be written

$$\left.\begin{aligned} z_{k-1} &= H_{k-1}x + v_{k-1} \\ z_k &= H_k x + v_k \end{aligned}\right\} \tag{14}$$

and we assume the set v_{k-1} to be uncorrelated with the set v_k, so that R has the form

$$R = \begin{bmatrix} R_{k-1} & 0 \\ 0 & R_k \end{bmatrix} . \tag{15}$$

Let $\hat{x}_{k-1}$ and P_{k-1} be the minimum variance unbiased estimate and its associated covariance matrix, defined on the subset z_{k-1} only. Then *the minimum variance unbiased estimate* $\hat{x}_k$ *defined on the set of measurements* $[z_{k-1}, z_k]$ *may be obtained without reprocessing the subset* z_{k-1} *and is given by*

$$\hat{x}_k = \hat{x}_{k-1} + K_k(z_k - H_k\hat{x}_{k-1}) , \tag{16}$$

with K_k a gain matrix defined by

$$K_k = P_{k-1}H_k^T[H_kP_{k-1}H_k^T + R_k]^{-1} , \tag{17}$$

while the covariance matrix P_k is obtained by .

$$P_k = (I - K_kH_k)P_{k-1} . \tag{18}$$

To demonstrate the above relations, note that, in view of Equation (7), the inverse covariance matrix is

$$\begin{aligned} P_k^{-1} &= H^TR^{-1}H \\ &= H_{k-1}^TR_{k-1}^{-1}H_{k-1} + H_k^TR_k^{-1}H_k \\ &= P_{k-1}^{-1} + H_k^TR_k^{-1}H_k . \end{aligned} \tag{19}$$

This expression has a classical form in matrix algebra, known as the Frobenius form, so that the inverse may be readily obtained:

$$\begin{aligned} P_k &= P_{k-1} - P_{k-1}H_k^T[H_kP_{k-1}H_k^T + R_k]^{-1}H_kP_{k-1} \\ &= (I - K_kH_k)P_{k-1} . \end{aligned} \tag{20}$$

On the other hand, in view of Equation (5), the new estimate $\hat{x}_k$ has the form

$$\begin{aligned} \hat{x}_k &= P_kH^TR^{-1}z \\ &= P_k[H_{k-1}^TR_{k-1}^{-1}z_{k-1} + H_k^TR_k^{-1}z_k] . \end{aligned} \tag{21}$$

Combination of Equations (20) and (21) yields the result

$$\begin{aligned} \hat{x}_k &= (I - K_kH_k)P_{k-1}[H_{k-1}^TR_{k-1}^{-1}z_{k-1} + H_k^TR_k^{-1}z_k] \\ &= (I - K_kH_k)\hat{x}_{k-1} + (I - K_kH_k)P_{k-1}H_k^TR_k^{-1}z_k \\ &= (I - K_kH_k)\hat{x}_{k-1} + P_{k-1}H_k^TR_k^{-1}z_k - \\ &\quad - P_{k-1}H_k^T[H_kP_{k-1}H_k^T + R_k]^{-1}[H_kP_{k-1}H_k^T + R_k]R_k^{-1}z_k + \\ &\quad + P_{k-1}H_k^T[H_kP_{k-1}H_k^T + R_k]^{-1}z_k \\ &= \hat{x}_{k-1} + K_k(z_k - H_k\hat{x}_{k-1}) , \end{aligned} \tag{22}$$

which completes the proof of the theorem.

The generalization to an arbitrary number of subsets of measurements z_k uncorrelated with each other is obvious.

2.5 Theorem 3 (Accumulative Estimation)

With the same definitions as above for z_{k-1}, z_k, R_{k-1}, R_k, $\hat{x}_{k-1}$ and P_{k-1}, let us now denote by $\hat{x}_k$ the minimum variance unbiased estimate defined on the subset z_k only, and let P_k be the corresponding covariance matrix; then, *the minimum variance unbiased estimate* $\hat{x}$ *defined on the two subsets* $[z_{k-1}, z_k]$ *is directly obtained by*

$$\hat{x} = P[P_{k-1}^{-1}\hat{x}_{k-1} + P_k^{-1}\hat{x}_k] \quad , \tag{23}$$

with a covariance matrix P computed from the relation

$$P^{-1} = P_{k-1}^{-1} + P_k^{-1} \quad . \tag{24}$$

The proof of this theorem is a direct consequence of Equations (5) and (7); in fact, we have

$$\begin{aligned} P^{-1} &= H^T R^{-1} H \\ &= H_{k-1}^T R_{k-1}^{-1} H_{k-1} + H_k^T R_k^{-1} H_k \\ &= P_{k-1}^{-1} + P_k^{-1} \end{aligned} \tag{25}$$

and

$$\begin{aligned} \hat{x} &= PH^T R^{-1} z \\ &= P[H_{k-1}^T R_{k-1}^{-1} z_{k-1} + H_k^T R_k^{-1} z_k] \\ &= P[P_{k-1}^{-1}\hat{x}_{k-1} + P_k^{-1}\hat{x}_k] \quad . \end{aligned} \tag{26}$$

The generalization to an arbitrary number of subsets of measurements uncorrelated with each other is again obvious.

2.6 Observability

From Equation (7) the existence condition for a minimum variance unbiased estimate is readily deduced: the unknown state vector is observable if, and only if, the matrix H has maximal rank.

2.7 Comments

Theorem 2 and its dual form (Theorem 3) have received numerous applications in the field of satellite orbit determination, especially Reference 4; the filtering procedures of Swerling[14], Claus[5] and Battin[1,2] are direct consequences of them. Theorem 2 is of primary importance in modern filtering theory, for it constitutes the cornerstone of the Kalman recursive filter.

It can be proved[11] that, for normal processes, linear least squares estimation yields the same solution as non-linear least squares estimation. Moreover, in that case, the least squares estimate turns out to be identical to the maximum likelihood estimate, which requires to maximize the conditional probability density of measuring z, assuming x.

The minimum variance unbiased estimate is often called Gaussian estimate, Markov estimate, for obvious reasons in view of Theorem 2 or weighted least squares estimate, underlining the occurrence of the covariance matrix R^{-1} in the quadratic form (9).

Let us finally emphasize that the minimum variance unbiased estimate requires no *a priori* information concerning the error distribution function.

3. THE KALMAN-BUCY FILTERING THEORY

3.1 The Kalman Recursive Filter

It will be shown in this section that the Kalman filter, almost simultaneously derived by Kalman and Bucy[8], Battin[1] and Swerling[14] is a direct consequence of Theorems 1 and 2, when applied to first-order discrete linear systems.

Consider an n-dimensional non-constant state vector, taking the value x_i at time t_i and obeying the following first-order discrete linear equation ($i = 1, \ldots, N$):

$$x_i = \Phi_{i,i-1}x_{i-1} + w_{i-1} \quad , \tag{27}$$

where w_{i-1} is a white noise random vector sequence with zero mean and positive definite covariance matrix Q_{i-1}

$$E(w_{i-1}) = 0$$

$$E(w_{i-1}w_{i-1}^T) = Q_{i-1}$$

$$E(w_i w_k^T) = \delta_{ki} Q_i \ ,$$

while $\Phi_{i,i-1}$ is the system state transition matrix from time t_{i-1} to time t_i. At each time t_i, the state vector x_i is observed through the measurement equation

$$z_i = H_i x_i + v_i \ , \tag{28}$$

with the same definitions for z_i, H_i and v_i as in Theorem 2. It is further assumed that the two white noise random sequences v_i and w_{i-1} are uncorrelated.

Then, the problem may be formulated as follows. Find from the measurements $(z_i, z_{i-1}, \ldots, z_1)$ the minimum variance unbiased estimate $\hat{x}_i$ (covariance matrix P_i) of the state x_i, depending linearly on the measurements, assuming the minimum variance unbiased estimate $\hat{x}_{i-1}$ (covariance matrix P_{i-1}) to be known from the measurements $(z_{i-1}, z_{i-2}, \ldots, z_1)$.

In order to apply the theorems of the preceeding sections to the problem at hand, let us introduce the vectors x^i, z^i, w^{i-1}, v^i defined by the following recurrence relations:

$$x^i = \begin{bmatrix} x_i \\ \hline x^{i-1} \end{bmatrix} , \qquad z^i = \begin{bmatrix} z_i \\ \hline z^{i-1} \end{bmatrix}$$

$$w^{i-1} = \begin{bmatrix} w_{i-1} \\ \hline w^{i-2} \end{bmatrix} , \qquad v^i = \begin{bmatrix} v_i \\ \hline v^{i-1} \end{bmatrix} .$$

If we similarly* define the matrices Q^{i-1}, R^i, H^i, Φ^i, F^i,

$$Q^{i-1} = E[(w^{i-1})(w^{i-1})^T]$$

$$R^i = E[(v^i)(v^i)^T]$$

$$H^i = \left[\begin{array}{c|c} H_i & 0 \\ \hline 0 & H^{i-1} \end{array}\right]$$

$$\Phi^i = \left[\begin{array}{c} I \\ \hline \Phi^{i-1}\Phi_{i,i-1} \end{array}\right]$$

$$F^i = \left[\begin{array}{c|c} 0 & 0 \\ \hline \Phi^{i-1}\Phi_{i,i-1} & F^{i-1} \end{array}\right] ,$$

the whole set of Equations (27) and (28) for i running from 1 up to i, may then be globally written

$$z^i = H^i\Phi^i x_1 - H^i F^i w^{i-1} + v^i \ , \tag{29}$$

or equivalently

$$z^i = H^i\Phi^i x_1 + e_i \ , \tag{30}$$

with e_i the random vector variable

$$e_i = - H^i F^i w^{i-1} + v^i \ , \tag{31}$$

with zero mean and covariance matrix C_i :

$$C_i = H^i F^i Q^{i-1}(H^i F^i)^T + R^i \ . \tag{32}$$

---

* The matrix $\Phi_{i,i-1}$ is assumed to have an inverse $\Phi_{i,i-1}^{-1}$.

Note that Equation (30) has the form of Equation (1) so that, by Theorem 1, the estimate $\hat{x}_{i-1}$ may be written

$$\hat{x}_{i-1} = P_{i-1}(H^{i-1}\Phi^{i-1})^T C_{i-1}^{-1} z^{i-1} \tag{33}$$

$$P_{i-1} = [(H^{i-1}\Phi^{i-1})^T C_{i-1}^{-1}(H^{i-1}\Phi^{i-1})]^{-1} , \tag{34}$$

in view of Equations (5) and (7).

Before calculating $\hat{x}_i$, let us first compute the minimum variance unbiased estimate $\hat{x}_i'$ (covariance matrix P_i') of the state x_i , based upon the measurements $(z_{i-1}, z_{i-2}, \ldots, z_1)$ only.

Introducing (27) in the generalized measurement equation

$$z^{i-1} = H^{i-1}\Phi^{i-1}x_{i-1} + e_{i-1} , \tag{35}$$

one obtains

$$z^{i-1} = H^{i-1}\Phi^{i-1}\Phi_{1,i-1}^{-1}x_1 + e_i' , \tag{36}$$

where e_i' is a new random vector variable:

$$e_i' = e_{i-1} - H^{i-1}\Phi^{i-1}\Phi_{1,i-1}^{-1}w_{i-1} , \tag{37}$$

with zero mean and covariance matrix C_i' :

$$C_i' = C_{i-1} + H^{i-1}\Phi^{i-1}\Phi_{1,i-1}^{-1}Q_{i-1}(H^{i-1}\Phi^{i-1}\Phi_{1,i-1}^{-1})^T .$$

As a Frobenius form is recognized in the above equation, $(C_i')^{-1}$ is known, i.e.

$$(C_i')^{-1} = C_{i-1}^{-1} - C_{i-1}^{-1}H^{i-1}\Phi^{i-1}\Phi_{1,i-1}^{-1}[(\Phi_{1,i-1}^{-1})^T P_{i-1}^{-1}\Phi_{1,i-1}^{-1} + Q_{i-1}^{-1}]^{-1}(H^{i-1}\Phi^{i-1}\Phi_{1,i-1}^{-1})^T C_{i-1}^{-1} \tag{38}$$

so that the covariance matrix P_i' , which may be written, in view of Theorem 1,

$$(P_i')^{-1} = (H^{i-1}\Phi^{i-1}\Phi_{1,i-1}^{-1})^T (C_i')^{-1}(H^{i-1}\Phi^{i-1}\Phi_{1,i-1}^{-1}) , \tag{39}$$

reduces to

$$\begin{aligned}(P_i')^{-1} &= (\Phi_{1,i-1}^{-1})^T P_{i-1}^{-1}\Phi_{1,i-1}^{-1} - (\Phi_{1,i-1}^{-1})^T P_{i-1}^{-1}\Phi_{1,i-1}^{-1}[(\Phi_{1,i-1}^{-1})^T P_{i-1}^{-1}\Phi_{1,i-1}^{-1} + Q_{i-1}^{-1}]^{-1}(\Phi_{1,i-1}^{-1})^T P_{i-1}^{-1}\Phi_{1,i-1}^{-1} \\ &= [\Phi_{1,i-1}P_{i-1}\Phi_{1,i-1}^T + Q_{i-1}]^{-1} ,\end{aligned} \tag{40}$$

as can be easily verified, using Equations (34) and (38). Thus, P_i' is equal to

$$P_i' = \Phi_{1,i-1}P_{i-1}\Phi_{1,i-1} + Q_{i-1} . \tag{41}$$

A similar manipulation of $\hat{x}_i'$, which by Theorem 1 is defined to be

$$\hat{x}_i' = P_i'(H^{i-1}\Phi^{i-1}\Phi_{1,i-1}^{-1})^T (C_i')^{-1} z^{i-1} , \tag{42}$$

leads via Equations (33), (34), (38) and (41) to the result

$$\hat{x}_i' = \Phi_{1,i-1}\hat{x}_{i-1} . \tag{43}$$

We are now in position to simply derive the Kalman filter equations, for all the conditions are satisfied to apply Theorem 2 with

$$\begin{aligned} x_k &= x_i \\ P_{k-1} &= P_i' \\ \hat{x}_{k-1} &= \hat{x}_i' \\ H_k &= H_i \\ R_k &= R_i . \end{aligned}$$

so that the estimate $\hat{x}_i$ is immediately given by Equations (16), (17) and (18):

$$\hat{x}_i = \hat{x}_i' + K_i(z_i - H_i\hat{x}_i') \tag{44}$$

$$K_i = P_i'H_i^T(H_iP_i'H_i^T + R_i)^{-1} \tag{45}$$

$$P_i = (I - K_iH_i)P_i' \ . \tag{46}$$

Equations (41), (43), (44), (45) and (46) constitute the Kalman filter. The physical meaning of these equations is evident: Equations (41) and (43) extrapolate the estimate $\hat{x}_{i-1}$ from time t_{i-1} to time t_i through the state transition equation (27) while Equations (44), (45) and (46) update this extrapolated estimate on account of the new available measurements z_i .

The dynamical system model (27) may be readily extended to include deterministic forcing terms f_{i-1} :

$$x_i = \Phi_{i,i-1}x_{i-1} + f_{i-1} + w_{i-1} \ . \tag{47}$$

This clearly affects the extrapolated estimate $\hat{x}_i'$ only and Equation (43) must be replaced by

$$\hat{x}_i' = \Phi_{i,i-1}\hat{x}_{i-1} + f_{i-1} \ , \tag{48}$$

the other filtering equations remaining unchanged.

The Kalman filter is capable of other important extensions; although out of the scope of the present contribution let us quote the simultaneous estimation of imprecisely known parameters, entering the dynamical system linearly[13], the replacement of white noise sequences for w_{i-1} , v_i by wide-sense Markov sequences[13] or by sequences correlated with respect to each other[10], stochastic optimization, and so on ...

3.2 The Accumulative Filter

It has been shown in the preceding section, that the Kalman filter is a direct application of Theorem 2 to first-order discrete linear systems; it is therefore natural to ask whether Theorem 3 may be applied to the same problem[7,8].

Consider Equations (27) and (28) of Section 3.1 with the same definitions and assumptions and let us introduce a n-dimensional vector y_i such that at any time t_i the following relation holds:

$$\hat{x}_i = P_iy_i \ . \tag{49}$$

It is easily verified from Theorems 1 and 2 that such a vector may always be defined. Furthermore, suppose that the vector y_{i-1} and the inverse covariance matrix P_{i-1}^{-1} are available at time t_{i-1} and let S_i be a matrix defined by

$$S_i = (\Phi_{i,i-1}^{-1})^T P_{i-1}^{-1}\Phi_{i,i-1}^{-1} \ . \tag{50}$$

Equations (40) and (41) may then be written

$$P_i' = S_i^{-1} + Q_{i-1} \tag{51}$$

$$(P_i')^{-1} = D_iS_i \ , \tag{52}$$

with D_i a matrix given by the relation

$$D_i = I - S_i[S_i + Q_{i-1}^{-1}]^{-1} \ . \tag{53}$$

Similarly, the extrapolated y_{i-1} vector, i.e. y_i' , can be written

$$y_i' = D_i(\Phi_{i,i-1}^{-1})^T y_{i-1} \ . \tag{54}$$

in view of Equation (43).

Use of Theorem 3 yields the new vector y_i at time t_i :

$$y_i = y_i' + H_i^TR_i^{-1}z_i \ , \tag{55}$$

and the inverse estimate covariance matrix

$$P_i^{-1} = (P_i')^{-1} + H_i^TR_i^{-1}H_i \ . \tag{56}$$

The estimate $\hat{x}_i$ is then available via

$$\hat{x}_i = P_i y_i \,. \tag{57}$$

Equations (50), (52), (53), (54), (55) and (56) constitute the accumulative filter equations for the problem at hand. It in important, to note that the calculation of the estimate $\hat{x}_i$ is made independently of the accumulative filter: it means that any error in the inversion of P_i^{-1} does not affect the succeeding estimates. For the general case, however, the accumulative filter is not a practical tool to solve the problem, in view of the $n \times n$ matrix inversion required in Equation (53). The situation is completely different in the absence of plant noise; in such cases ($w_{i-1} = 0$ for any i), the M_i matrix reduces to the identity matrix so that the accumulative filter equations simply become

$$B_i = (\Phi_{i,i-1}^{-1})^T P_{i-1}^{-1} \Phi_{i,i-1} \tag{50}'$$

$$(P_i')^{-1} = B_i \tag{52}'$$

$$M_i = I \tag{53}'$$

$$y_i' = (\Phi_{i,i-1}^{-1})^T y_{i-1} \tag{54}'$$

$$y_i = y_i' + H_i^T R_i^{-1} z_i \tag{55}'$$

$$(P_i)^{-1} = (P_i')^{-1} + H_i^T R_i^{-1} H_i \,. \tag{56}'$$

Any numerical matrix inversion has completely disappeared in the scheme, which is not the case in the Kalman filter: the computation procedure is made particularly stable and this may be of primary importance when weakly observable systems are dealt with.

3.3 Observability Condition

In view of Theorem 1 and Equation (29), an observability condition is easily obtained: the state x_i is observable if, and only if, the matrix $(H^i\Phi^i)^T C_i^{-1}(H^i\Phi^i)$ has an inverse or, more compactly, if and only if the matrix $H^i\Phi^i$ has a generalized inverse (C_i is positive definite).

4. CONCLUSION

The modern recursive filtering theory has been shown to proceed from an idea originally due to Gauss, at least in essence. More precisely, simple theorems arising from the minimum variance estimation of a constant state, linearly depending on the measurements, have been applied to discrete first-order stochastic linear systems, by the way of a restatement of the problem. As a result, the Kalman filter and its dual form have been obtained by simple algebraic manipulations.

REFERENCES

1. Battin, R.H. — *A Statistical Optimizing Navigation Procedure for Space Flight.* American Rocket Society Journal, Vol.32, 1962, pp.1681-1682.

2. Battin, R.H. — *Astronautical Guidance.* McGraw-Hill, New York, 1937.

3. Bellman, R. — *Introduction to Matrix Analysis.* McGraw-Hill, New York, 1960.

4. Blackmann, R.E. — *Methods of Orbit Refinement.* Bell System Technical Journal, Vol.43, 1965, pp.885-909.

5. Claus, A.J. — *Orbit Determination for Communication Satellites from Angular Data Only.* American Rocket Society 17th Annual Meeting, Los Angeles, 1962.

6. Gauss, K.F. — *Theory of the Motion of the Heavenly Bodies Moving About the Sun in Conic Sections.* Dover, New York, 1963.

7. Genin, Y. — *A Note on Linear Minimum Variance Estimation Problems.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-13, 1968, p.103.

8. Genin, Y. — *Gaussian Estimates and Modern Linear Discrete Filtering Methods.* Revue A, Vol.13, 1968, pp.70-78.

9. Kalman, R.E. and Bucy, R.S. — *New Results in Linear Filtering and Prediction Theory.* Journal of Basic Engineering, Vol.83D, 1961, pp.95-108.

10. Kalman, R.E. — *Fundamental Study of Adaptive Control Systems.* Technical Report ASD-TR 61, Vol.1, (NASA N 62-15355).

11. Papoulis, A. — *Probability, Random Variables and Stochastic Processes.* McGraw-Hill, New York, 1965.

12. Rosenbrock, H.H. — *On the Connection Between Discrete Linear Filters and Some Formulae of Gauss.* Reprints of AFRA Congress, Paper 1, Paris, May 4-7, 1965.

13. Sorenson, H.W. — *Kalman Filtering Techniques.* Advances in Control Systems, Vol.3, Chapter 5, Academic Press, 1966.

14. Swerling, P. — *A Proposed Stagewise Differential Correction Procedure for Satellite Tracking and Prediction.* Rand Corporation, Paper P-1292, 1958.

CHAPTER 3 - COMPUTATIONAL TECHNIQUES IN KALMAN FILTERING

by

Stanley F. Schmidt

Analytical Mechanics Associates, Inc.
264 Sheridan Avenue
Palo Alto, California, USA

CHAPTER 3 -- COMPUTATIONAL TECHNIQUES IN KALMAN FILTERING

Stanley F. Schmidt

1. INTRODUCTION

In the application of Kalman filtering theory, the mathematical formulation of the problem and computational techniques utilized may depend heavily on the computational capabilities available. This interface is particularly strong in real time aerospace applications wherein considerations of weight, power, and reliability generally require the selection of techniques which minimize onboard computational requirements.

The discussion of this chapter will emphasize computing techniques for Kalman filtering in these onboard type problems. The techniques are applicable but not necessarily essential when large general purpose computing facilities are available.

A note in passing is that the Kalman filter is a special data processing technique. It was originally introduced[1,2] for the onboard application because of its reduction in computational requirements over other methods available at the time.

Numerous investigators[3-8] have reported various problems with and "fixes" for Kalman filters. The problems in general fall into the following categories:

(a) Loss of positive definiteness in the covariance matrix resulting from numerical errors.

(b) Improper mathematical model, leading to a divergence of the estimates from measurements.

(c) Nonlinear phenomena generally aggravated by a poor selection of the starting estimate.

Of these problems, (a) has probably been experienced to some degree by almost everyone working in the field. Reasonably efficient (from computational considerations) solutions for this problem based on square root algorithms will be given in this chapter.

For Item (b), techniques exist for compensating for errors in the mathematical model. The techniques, although useful, still leave a good deal of "cut and try" in finding suitable solutions. Some relevant material on this subject will be covered in this chapter.

Item (c) will not be covered here. The mechanization equations described use linearization about the current best estimate of state. This factor and the occasional requirement of specially designed starting calculations (based on the raw data) have removed nonlinear problems in the practical applications with which the author is familiar.

2. MATHEMATICAL FORMULATION AND DEFINITIONS

One of the first problems which must be addressed in applying Kalman filtering theory is the mathematical formulation of the problem. This section will define the overall filtering problem from a practical viewpoint.

2.1 Problem Statement and Discussion

(1) A true set of dynamical equations describing the dynamic behavior of the system is believed to be represented by a set of vector differential equations

$$\dot{\bar{X}} = \bar{F}(\bar{X},\bar{C},\bar{U},t) \ . \tag{2.1}$$

One therefore develops a set of equations for the system of the form

$$\dot{X} = F(X,C,U,t) \ , \tag{2.2}$$

where

X = a vector describing the fundamental state variables*

C = a vector describing constants used in the system equations

* The term "fundamental state variables" is used for those time-varying quantities whose initial conditions must be specified in solving (2.2) (e.g., the position and velocity vectors of an orbiting spacecraft).

U = a vector of forcing functions acting on the dynamic system

t = the independent variable, time.

(ii) A vector, Y, of measurements is available which is believed related to the fundamental state variables in the form

$$\bar{Y} = G(\dot{\bar{X}}, \bar{V}, T) + \bar{q}(t) . \tag{2.3}$$

A model of these measurements is therefore developed of the form

$$Y = G(X, V, t) + q(t) , \tag{2.4}$$

where

V = a vector describing assumed constants and/or time-varying states in the measurement model

q = the assumed random error in the measurement.

Before proceeding further in the definitions some discussion of the notation and meaning of the phrases "believed to be" and "assumed to be" is in order.

"Believed to be" is defined as meaning: in conformance with such scientific knowledge as is possessed at the present time.

"Assumed to be" is used in the context of known approximation, such as truncations of infinite series representations, and so on.

(iii) An initial estimate of the state vector at a time t_0 may be given, or is assumed, as

$$\hat{X}(t_0) = \hat{X}_0 . \tag{2.5}$$

(iv) Other initial conditions and time-varying forcing functions which are assumed are*

$$\hat{C}(t_0) = \hat{C}_0 \tag{2.6}$$

$$\hat{V}(t_0) = \hat{V}_0 \tag{2.7}$$

$$\hat{U}_0(t) = \hat{U}_0 . \tag{2.8}$$

(v) The errors in the previously defined quantities are defined by

$$E(X_0 - \hat{X}_0) = E(\tilde{X}_0) = 0 \tag{2.9}$$

$$E(\tilde{X}_0 \tilde{X}_0^T) = \text{covariance matrix} = P_0 \tag{2.10}$$

$$E[q(t)] = 0 \tag{2.11}$$

$$E(qq^T) = Q(t) \tag{2.12}$$

$$E(C - \hat{C}_0) = E(\tilde{C}_0) = 0 \tag{2.13}$$

$$E(\tilde{C}_0 \tilde{C}_0^T) = P_{C_0} \tag{2.14}$$

$$E[U(t) - \hat{U}_0(t)] = E[\tilde{U}(t)] = 0 \tag{2.15}$$

$$E(\tilde{U}\tilde{U})^T = P_{U_0}(t) \tag{2.16}$$

$$E(V - \hat{V})_0 = E(\tilde{V}) = 0 \tag{2.17}$$

$$E(\tilde{V}\tilde{V}^T) = P_{V_0} \tag{2.18}$$

The use of the expected value operator E in Equations (2.9) to (2.18) requires some discussion. Nominal usage of the operator implies an ensemble average. For example, if the errors in $\hat{X}$ are considered as injection errors of a space vehicle, the ensemble represents an infinite number of launch vehicles of the same type launched from the same location. If the basic causes of injection errors are random in this ensemble and it is meaningful for the problem, then (2.9) is a true statement.

In general one cannot define a meaningful ensemble for quantities such as errors in the estimate of the vectors like C and V. Hence one really should consider the initial values as estimates and the covariance matrix as a confidence level. This is, in fact, how one uses the quantities and the problem of defining the ensemble is avoided. Equations (2.3), (2.7) and probably (2.15) are not true statements in the strict statistical sense.

---

* As a result of limited knowledge and known approximation in computations, one may desire to consider a portion of $U(t)$ to be a random forcing function. This will be discussed later in this chapter as a possible means of dynamic model compensation.

2.3 Problem Objectives

The objective of filtering (or data processing) considered herein is to find:

> Algorithms for use in a given computer which process the measurement data and provide *an estimate of the state variables which is sufficiently accurate for real time command and control purposes.*

Additional factors which must be considered in this problem include:

(a) Computer speed.
(b) Computer memory availability.
(c) Computational errors.
(d) Measurement data availability.
(e) Development time schedule.

This problem statement contains some of the practical considerations which arise in most applications of Kalman filtering. Theory is used to provide approaches for finding a practical solution. By no means will this solution be optimal in the theoretical sense. This is because exact mathematical models of all quantities, as well as practical performance indices which consider all the factors, cannot be defined.

Subsequent discussion will assume that preliminary considerations of computer speed and memory have led to the requirement of data processing algorithms of a "one pass" nature. That is, once measurements at any time point have been processed they must be discarded. This assumption is made to restrict the scope of subsequent discussion to modified forms of the Kalman filter.

A practical solution of the filtering problem in any specific example is found by the use of existing theory and practical experience to define starting algorithms. Simulation is then used for further validation and/or modifications. Final validation and/or modifications then takes place in tests of the real system.

Before introducing the basic algorithms, some additional definitions are required. If the gradient of (2.2) is taken with respect to X, C, and U, a set of time-varying linear differential equations is obtained*:

$$\dot{x} = F(t)x(t) + B(t)c + D(t)\,u(t)\,, \tag{2.19}$$

where

$$F(t) = \nabla_X F(X,C,U,t)\big|_{X(t)\,=\,\hat{X}(t)}$$

$$B(t) = \nabla_C F(X,C,U,t)\big|_{C(t)\,=\,\hat{C}(t)}$$

$$D(t) = \nabla_U F(X,C,U,t)\big|_{U(t)\,=\,\hat{U}(t)}$$

Since (2.19) is linear, the general solution may be written in the form

$$x(t) = \Phi(t;t_0)x(t_0) + \Phi_c(t;t_0)c + \int_{t_0}^{t} \Phi(t;\tau)D(\tau)u(\tau)\,d\tau\,. \tag{2.20}$$

The transition matrix $\Phi(t;t_0)$ and the sensitivity to constant forces or control can be found by solving

$$\dot{\Phi} = F(t)\Phi\,, \qquad \Phi(t_0;t_0) = I \tag{2.21}$$

$$\dot{\Phi}_c = F(t)\Phi_c + B(t)(I)\,, \qquad \Phi_c(t_0;t_0) = 0\,. \tag{2.22}$$

The integral term of (2.20) may be solved by approximating $u(t)$ as a constant for small time increments, Δ:

$$\dot{\Phi}_u = F(t)\Phi_u + D(t)I\,, \qquad \Phi_u(t_0+n\Delta;\, t_0+n\Delta) = 0\,. \tag{2.23}$$

It is usually convenient to consider an augmented state vector of the form

$$\hat{Z} = \begin{pmatrix} \hat{X} \\ \hat{C} \\ \hat{U}_c \end{pmatrix} \begin{matrix} \text{- fundamental state variables} \\ \text{- constant terms in equations of motion} \\ \text{- constant forces.} \end{matrix}$$

In this instance the augmented state deviation obeys

$$z(t) = \Phi_z(t;t_0)z(t_0) + \int_{t_0}^{t} \Phi_z(t;\tau)D(\tau)u_0(\tau)\,d\tau \tag{2.24}$$

if linear assumptions are valid.

---

* Lower case letters are used to denote small deviations from the corresponding upper case value. That is, a vector Z is considered as $Z = \hat{Z} + z$.

The gradient of the measurement equation (2.4), is taken with respect to X and V to find linear equations for the i^{th} measurement*

$$y_i = h_i x + g_i v + q_i \,. \tag{2.25}$$

where

$$h_i = \nabla_X G(X,V,t_i)\big|_{X(t)=\hat{X}(t)} \tag{2.26}$$

$$g_i = \nabla_V G(X,V,t_i)\big|_{V=0} \,. \tag{2.27}$$

One may desire to relate the vector of measurements to the vector, z_0, at a fixed epoch. This may be done by using (2.24)† and (2.25) to obtain

$$y = Hz_0 + Gv + q \,. \tag{2.28}$$

In (2.28) the vector y may consist of measurements at different time points. The vector z_0 can be at a fixed epoch. This fact can be used to obtain an algorithm for processing small batches of data or for a simplified form of data compression[10]. It is generally desirable to consider a further augmentation of the state to include the measurement parameters, V. Then (2.28) would be written in the equivalent form

$$y = (H\ G)\begin{pmatrix} z_0 \\ v \end{pmatrix} + q \triangleq \mathcal{H} z_0 + q \,. \tag{2.29}$$

In the case where it is meaningful or necessary, one may write the equation for updating the covariance matrix $P_z(t)$ in time

$$P_z(t) = \Phi_z(t;t_0) P_z(t_0) \Phi_z^T(t;t_0) + R \,. \tag{2.30}$$

The matrix R in (2.30) gives the added uncertainty caused by random forcing functions in the time interval $t - t_0$.

2.3 Kalman Filter Equations

A modified form of the Kalman filter for nonlinear systems is given here for processing the measurements in a sequential manner. These algorithms are readily derivable from Kalman's original equations for the discrete filter[9]. The algorithm is conveniently stated in two parts:

(i) Between the i and $i+1$ measurement times,

$$\hat{Z}(t_{i+1}) = \hat{Z}(t_i) + \int_{t_i}^{t_{i+1}} \dot{\hat{Z}}(\tau)\, d\tau \tag{2.31}$$

$$\dot{\Phi}_z = \left(\nabla \dot{\hat{Z}}\Big|_{Z=\hat{Z}}\right) \Phi_z \,, \quad \Phi_z(t_i;t_i) = I \,. \tag{2.32}$$

(ii) At measurement times t_{i+1},

$$P_z(t_{i+1}) = \Phi_z P_z(t_i) \Phi_z^T + R \triangleq P_b \tag{2.33}$$

$$\hat{Z}_a = \hat{Z}_b + K(Y - \hat{Y}) \tag{2.34}$$

$$P_a = P_b - KHP_b \tag{2.35}$$

$$K = P_b H^T (HP_b H^T + Q)^{-1} \tag{2.36}$$

$$\hat{Y} = G(\hat{X},\hat{V},t_{i+1}) = \text{computed measurement.}$$

Equation (2.31) is a simple statement of integrating the system dynamic equations (2.2). The $\hat{Z}$ notation is used to imply a time update of all those components of the state which are assumed non-constant. The transition matrix can be calculated by integrating the variational equations implied in (2.32) with the appropriate initial conditions. At the measurement time t_{i+1}, the covariance matrix $P_z(t_{i+1})$ is calculated by (2.33). Equations (2.34), (2.35), and (2.36) are then used to calculate the new estimate of state, $\hat{Z}_a$, and the new covariance matrix P_a. These new values result from including the measurements Y at time t_{i+1} in both the estimate and the covariance matrix.

Note that if the influence of random forcing functions is approximated by the use of (2.23), then (2.33) should be updated at Δ time intervals between the measurements.

---

* The random error in measurement is assumed to be small and therefore considered to give a small random deviation from a noiseless measurement.

† The effects of the random variation of $u(t)$ are assumed negligible in writing this expression. Otherwise, a time correlated error must be added to equation (2.28).

3. SQUARE ROOT FORMULATIONS

As already mentioned, numerical errors frequently cause a problem when the covariance matrix equations (2.33) and (2.35) are used. The usual symptoms of the problem are negative diagonal terms occuring in the covariance matrix after processing data. The problem can occur after the first measurement in some instances[10]. It is more of a "nuisance" than of a serious type, as one usually can find a remedy by adding or increasing R of (2.33). This is one of the techniques discussed in Section 4 for model error compensation. The other technique given in Section 4 also can be used as a remedy for this numerical problem.

This numerical difficulty can also be removed by square root methods. The square root methods also provide a significant improvement in numerical accuracy of the covariance matrix. Since the covariance matrix is the weighting factor containing the effects of all past measurements, improving the accuracy of this matrix can be important in some problems.

The square root of P is defined as W and satisfies the relationship

$$WW^T = P = \sum_{i=1}^{n} w_i w_i^T . \tag{3.1}$$

where

w_i = i^{th} column vector of the matrix W

n = the number of columns of W, which will be defined as the dimension of P.

With this definition, if P is positive definite, the column vectors w_i are linearly independent. It is generally desirable to force the column vectors of W to be linearly independent. Then when P is positive semi-definite some column vector or vectors will be null. The rank of P is equal to the number of non-zero column vectors of W in this formulation.

In the case of the square root matrix one needs computation algorithms for

(a) starting the problem (e.g., given P_0 find W_0),

(b) propagating W in time (square root of (2.33)),

(c) modifying W for including measurements (square root of (2.35)).

Subsequent material will summarize algorithms for these purposes.

3.1 Initialization of the Square Root Matrix

One obvious way of specifying the initial square root matrix W_0 is to define the column vectors of W_0. One should note that when the columns of W_0 are made linearly independent, each column specifies an independent error source. All error sources lie in different directions in the state space.

An alternate problem is that P_0 is given and a suitable square root matrix W_0 is to be defined. For this usage and from the preceding discussion it is seen that the desired algorithm will construct a set of column vectors of the matrix W_0 which are linearly independent. Consider the application of the following equation for $B_1 = P_0$:

$$B_{k+1} = B_k - \frac{B_k v_k v_k^T B_k}{v_k^T B_k v_k} . \tag{3.2}$$

In (3.2) v_k is an arbitrary (non-zero) column vector. Note that the quantity $v_k^T B_{k+1} v_k$ is zero. In other words, (3.2) removes all errors of P_0 in the direction of the vector v_k. If v_k is chosen such that its k^{th} element is unity and the remaining elements are zero, the k^{th} row and column from B_k are removed. By letting $k = 1,\ldots,n$, then $B_{n+1} = 0$. The quantity n is the dimension of P_0 (the number of state variables).

Therefore, by defining the column vectors of W_0 by

$$w_k = \left\{ \begin{array}{ll} \dfrac{B_k v_k}{\sqrt{(v_k^T B_k v_k)}} & \text{if } v_k^T B_k v_k \neq 0 \\ 0 & \text{if } v_k^T B_k v_k = 0 \end{array} \right\} \quad \text{for } k = 1,\ldots,n, \tag{3.3}$$

the resulting W_0 matrix will be of lower triangular form. From filtering considerations one can view the algorithm as the equivalent of making n perfect observations of the state vector which reduce the error in all components to zero ($B_{n+1} = 0$). The error vector removed after each step of the algorithm is stored so that the matrix P_0 can be recovered from W_0. This algorithm is given in a slightly different form in Reference 11.

The algorithm provides a means of calculating a lower triangular matrix T such that

$$TBT^T = \text{a diagonal matrix}, \tag{3.4}$$

where, in (3.4),

$$B = \text{a positive semi-definite matrix.}$$

Another use of the algorithm is in finding a sequence of observations whose random errors are uncorrelated from a sequence whose errors are correlated. For example, if

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} x + \begin{pmatrix} q_1 \\ q_2 \end{pmatrix} \tag{3.5}$$

and

$$E\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}(q_1 q_2) = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix} = Q, \tag{3.6}$$

then the observations given by

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{bmatrix} y_1 \\ a_1 y_1 + y_2 \end{bmatrix} = \begin{bmatrix} h_1 \\ a_1 h_1 + h_2 \end{bmatrix} x + \begin{bmatrix} q_1 \\ a_1 q_1 + q_2 \end{bmatrix} \tag{3.7}$$

have uncorrelated random errors if

$$a_1 = -\rho\frac{\sigma_2}{\sigma_1}. \tag{3.8}$$

In this instance

$$T = \begin{pmatrix} 1 & 0 \\ -\dfrac{\rho\sigma_2}{\sigma_2} & 1 \end{pmatrix}. \tag{3.9}$$

Since

$$TQT^T = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & (1-\rho^2)\sigma_2^2 \end{bmatrix}, \tag{3.10}$$

the measurement z_1 is the same as y_1 and its random error is σ_1^2. The measurement z_2 is a linear combination of y_1 and y_2 so chosen that its random error is uncorrelated with the random error of z_1. The variance of the random error in z_2 approaches zero as the correlation factor ρ approaches unity.

3.2 Measurement Update of the Square Root Matrix

For a single (scalar) measurement, (2.35) may also be written in square root form as*

$$W_a = W_b\left[I - \frac{W_b^T H^T H W_b}{a}\right], \tag{3.11}$$

where $W_a W_a^T = P_a$

and $$a = (HW_b W_b^T H^T + Q)\left(1 + \sqrt{\left[\frac{Q}{HW_b W_b^T H^T + Q}\right]}\right).$$

Multiple measurements (at a given time point) may be treated as sequential scalar measurements if their random errors are uncorrelated. The previous material illustrated how an uncorrelated sequence may always be defined; hence the square root formulation of (3.11) is general. Other formulations for multiple measurements are also available[12].

---

* This algorithm is given in Reference 12 and will be referred to as Potter's method.

3.3 Time Update of the Square Root Matrix

Two methods for updating the square root matrix in time for random forcing functions are given. Both assume that the matrix R of Equation (2.33) is available in factored form such that

$$R = \sum_{i=1}^{j} s_i s_i^T = SS^T , \tag{3.12}$$

where

s_i are linearly independent vectors

j = number of random forcing functions.

3.3.1 Method No. 1

A square root matrix time update algorithm can be developed by noting that

$$P^{-1}(t) = \left[\Phi P \Phi^T + \sum_{i=1}^{j} s_i s_i^T\right]^{-1} \tag{3.13}$$

As a result of the well-known matrix indentity

$$[B + tt^T]^{-1} = B^{-1} - B^{-1}t(t^T B^{-1} t + 1)^{-1} t^T B^{-1} , \tag{3.14}$$

and Potter's square root method for equations like the right-hand side of (3.14), one method for obtaining the time update of the square root matrix is readily developed. The steps are as follows:

(i) Define $CC = [\Phi W(t_0)]^{-1}$.

(ii) Operate j times on the matrix C, using Equation (3.11) with $H = s_i$, $i = 1, j$ and $Q = 1$. Let the result be C_j.

(iii) Then

$$W(t) = C_j^{-1} .$$

The author is not aware of any extensive use of this technique, as a result of the obvious drawback of requiring two matrix inversions for every time update. Reference 14 refers to this technique as Potter's method.

3.3.2 Method No. 2

An alternate technique which does not require inversion can be developed in a relatively simple manner from the material presented thus far.

Examination of Equation (3.2) illustrates that it may be factored as follows

$$A_{k+1} A_{k+1}^T = M_k A_k A_k^T M_k^T = B_{k+1} , \tag{3.15}$$

where

$$M_k = \left[I - \frac{A_k (A_k^T v_k v_k^T)}{v_k^T A_k A_k^T v_k}\right] . \tag{3.16}$$

Hence, if the matrix A is reduced using

$$A_{k+1} = M_k A_k , \quad k = 1, \ldots, n, \tag{3.17}$$

and the W matrix of Equation (3.3) is stored, a lower triangular matrix in W will result. Furthermore, this lower triangular matrix has the property that

$$WW^T = A_1 A_1^T . \tag{3.18}$$

It should therefore be obvious that, if

$$A_1 = [\Phi W \vdots S] , \tag{3.19}$$

the resulting square root matrix W is the desired square root matrix which includes the effects of the random forcing functions. This method was developed by the author[18] but appears to be equivalent in practice to the technique outlined in Reference 14.

The special choice of vectors (the k^{th} element unity and all other elements zero) makes the matrix M_k very simple to calculate. Furthermore, in (3.17) each step of the algorithm makes the k^{th} row of the matrix A_{k+1} zero. As a result the number of operations is not as large at it may appear at first glance.

To illustrate the simplicity of the algorithm, a simple example will be carried through each step. Let

$$\Phi W = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad S = I . \tag{3.20}$$

Then

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \tag{3.21}$$

$$A_1^T v_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

Define

$$w_1^{-n} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = A_1 A_1^T v_1$$

$$M_1 = \left[I - \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} \frac{(1\ 0\ 0)}{2} \right] = \begin{bmatrix} 0 & 0 & 0 \\ -1/2 & 1 & 0 \\ -1/2 & 0 & 1 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} 0 & 0 & 0 \\ -1/2 & 1 & 0 \\ -1/2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & -1/2 & 1 & 0 \\ 1/2 & 1 & 1 & -1/2 & 0 & 1 \end{bmatrix}$$

$$w_1 = \begin{pmatrix} 2/\sqrt{2} \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} = \text{first column of } W(t)$$

$$w_2^{-n} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & -1/2 & 1 & 0 \\ 1/2 & 1 & 1 & -1/2 & 0 & 1 \end{pmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1 \\ 0 \\ -1/2 \\ 1 \\ 0 \end{bmatrix} = \begin{pmatrix} 0 \\ 2.5 \\ 1.5 \end{pmatrix}$$

$$M_2 = \begin{bmatrix} 0 \\ I - \begin{matrix} 2.5 \\ 1.5 \end{matrix} \dfrac{(0\ 1\ 0)}{2.5} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & \dfrac{-1.5}{2.5} & 1 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -0.6 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5 & 1 & 0 & -0.5 & 1 & 0 \\ 0.5 & 1 & 1 & -0.5 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0.2 & 0.4 & 1 & -0.2 & -0.6 & 1 \end{bmatrix}$$

$$w_2 = \begin{pmatrix} 0 \\ \sqrt{2.5} \\ 1.5/\sqrt{2.5} \end{pmatrix} = \text{second column of } W(t)$$

$$w_3^{-n} = \begin{pmatrix} 0 \\ 0 \\ 2.6 \end{pmatrix}$$

$$M_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A_4 = 0$$

$$w_3 = \begin{pmatrix} 0 \\ 0 \\ \sqrt{2.6} \end{pmatrix} = \text{third column of } W(t) .$$

The final answer is

$$W(t) = \begin{pmatrix} 2/\sqrt{2} & 0 & 0 \\ 1/\sqrt{2} & \sqrt{2.5} & 0 \\ 1/\sqrt{2} & 1.5/\sqrt{2.5} & \sqrt{2.6} \end{pmatrix} . \tag{3.22}$$

As a check one may compute $A_1A_1^T$ from Equation (3.21) and compare the result to $W(t)W^T(t)$ of Equation (3.22).

$$A_1A_1^T = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

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

$$W(t)W^T(t) = \begin{bmatrix} 2/\sqrt{2} & 0 & 0 \\ 1/\sqrt{2} & \sqrt{2.5} & 0 \\ 1/\sqrt{2} & 1.5/\sqrt{2.5} & \sqrt{2.6} \end{bmatrix} \begin{bmatrix} 2/\sqrt{2} & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{2.5} & 1.5/\sqrt{2.5} \\ 0 & 0 & \sqrt{2.6} \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 4 \end{bmatrix} , \text{ which checks.}$$

As noted from the example, the following two facts about the method are significant:

(i) The dimension of M_k and A_k reduces (effectively) by one at each step.

(ii) The final solution for $W(t)$ is a lower triangular matrix, that is, the elements above the diagonal of $W(t)$ are zero. This fact can be used to reduce the number of calculations for the measurement update for the first measurement processed after the time update algorithm.

Both these facts result from the specific choice of the v_k vectors.

3.4 Recommended Algorithms for Square Root Implementations

Time Update Equations to a Measurement

Let $A_1 = (\Phi W \vdots S)$. (3.23)

Calculate the column vectors of $W(t)$ in accordance with the n-step algorithm $k = 1,\ldots,n$ where $v_k(l) = 0$, $l \neq K$, and $v_k(k) = 1$.

$$W_k = A_k A_k^T v_k \tag{3.24}$$

$$M_k = I - \frac{W_k v_k^T}{W_k(k)} \tag{3.25}$$

$$A_{k+1} = M_k A_k \tag{3.26}$$

$$w_k = W_k / \sqrt{(W_k(k))} = k^{th} \text{ column of } W(t) . \tag{3.27}$$

Measurement Update (for a Single Measurement)

Compute the vectors

$$v = W^T H^T \tag{3.28}$$

$$u = Wv \tag{3.29}$$

$$a = (v^T v + Q)\left(1 + \sqrt{\left[\frac{Q}{v^T v + Q}\right]}\right). \tag{3.30}$$

Modify W in accordance with

$$W = W - uv^T/a . \tag{3.31}$$

3.5 Comparative Computation Requirements

Although the proposed method should enhance the accuracy it may be costly in machine time. The proposed method is compared here with the more conventional method using the covariance matrix (Equations (2.33), (2.35) and (2.36).

The comparison data given assume that none of the matrices or vectors have zeros in prescribed locations. Also P is assumed an $n \times n$ matrix and S an $n \times m$ matrix.

Number of Operations for a Time Update

(i) Conventional Method[*†]

	M & D	*A & S*
Equation (2.33)	$2n^3 + n^2m$	$2n^3 + n^2m + n^2$

(ii) Square Root Method

	M & D	*A & S*	*Square Root*
Equation (3.23)	n^3	n^3	
Equation (3.24)	$\sum_{i=1}^{n} (n+m)(n+1-i)$	$\sum_{i=1}^{n} (n+m)(n+1-i)$	0
Equation (3.25)	$\sum_{i=1}^{n} (n-i)$	0	0
Equation (3.26)	$\sum_{i=1}^{n} (n+m)(n-i)^2$	$\sum_{i=1}^{n} (n+m)(n-i)^2$	0
Equation (3.27)	$\sum_{i=1}^{n} (n-i)$	0	0

* M & D means multiplications and/or divisions.

† A & S means additions and/or subtractions.

As an example assume that $n = 10$ and $m = 1$. The use of the tabulated equations for number of operations gives the following:

	Equivalent M & D	*M & D*	*A & S*	*Square Roots*
Conventional Method	2550	2100	2200	0
Square Root	5837	4630	4785	10

The column labeled "Equivalent M & D" assumes

(i) Square roots take 5 times as long as multiplications.

(ii) Additions take 1/6 of the time of multiplications.

For this example, the square root method takes twice as much computer time as the conventional method. If the advantage of symmetry of the covariance matrix for the conventional method was forced, then it would be about 3 times as fast as the modified square root method.

Number of Operations for a Measurement Update (Scalar Observation)

(i) *Conventional Method*

	M & D	*A & S*
Equations (2.35) and (2.36)	$3n^2 + n$	$2n^2 + n + 1$

(ii) *Square Root Method*

	If Triangular On Start		*Not Triangular*	
	M & D	*A & S*	*M & D*	*A & S*
Equation (3.28)	$\sum_{i=1}^{n} n + 1 - i$	$\sum_{i=1}^{n} n + 1 - i$	n^2	n^2
Equation (3.29)	$\sum_{i=1}^{n} n + 1 - i$	$\sum_{i=1}^{n} n + 1 - i$	n^2	n^2
Equation (3.30)	$n + 2$	2	$n + 2$	2
Equation (3.31)	$n^2 + n$	n^2	$n^2 + n$	n^2

One square root is required in Equation (3.30).

The number of operations in the various columns indicate whether or not W is triangular after a time update. This advantage can only be used for the first measurement processed at a given time point.

For comparative data assume that $n = 10$ and a single measurement is processed.

	Equivalent M & D	*M & D*	*A & S*	*Square Root*
Conventional	352	310	211	0
Square root	387	322	302	1
Square root (triangular)	287	232	212	1

The "Equivalent M & D" column is computed using the same relationships as stated for the previous case. The square root method has a slight advantage for the first measurement processed if W is triangular. However, this advantage is lost for subsequent measurements.

If one takes advantage of symmetry for the conventional method and other obvious improvements in techniques (for example forming the vector $PH^T/(HPH^T+Q)$ instead of dividing n^2 times), then the conventional method will be about twice as fast for the example cited.

3.5.1 *Influence of Coding Techniques*

Coding can significantly alter the previous computational time estimates. For example, if the square root method were coded as

$$W_+ = W\left[I - \frac{W^T H^T H W}{a}\right] \tag{3.32}$$

and the conventional method as

$$P_+ = [I-KH]P[I-KH]^T + KQK^T , \tag{3.33}$$

where

$$K = PH^T(HPH^T+Q)^{-1} ,$$

then the comparison is as follows for a single measurement. No advantage of W being triangular is used.

	M & D	*A & S*	*Square Root*
Conventional	$2n^3 + 3n^2 + 3n$	$2n^3 + 2n^2 + 2n + 1$	0
(n = 10)	2330	2221	0
Square root	$n^3 + 2n^2 + n + 1$	$n^3 + n^2 + 2n + 1$	1
(n = 10)	1211	1122	1

Several conclusions are evident:

(i) One should not code the square root algorithm using Equation (3.32) if speed of computations is of any importance. There are no obvious advantages of (3.32) for accuracy either, so it is definitely not recommended. The previously outlined procedure, where no matrix multiplications are used, is decidedly superior.

(ii) Equation (3.33) is extremely poor from speed considerations for the conventional method. It requires at least 10 times as long as (2.35) for the example of $n = 10$. This speed difference will get even worse with large values of n. Equation (3.33) has the advantage that the covariance matrix P should remain positive definite. However, there are numerous other ways of doing this, so that the use of (3.33) is not recommended.

(iii) Proper coding for both techniques should always make the conventional method superior with respect to computation time.

3.6 Formulation Using the Information Matrix

One may readily obtain a set of filtering equations where the inverse of the covariance matrix (the information matrix) is used. The equations are as follows.

3.6.1 *Time Update*

Let $$P^{-1} \triangleq \Lambda ; \tag{3.34}$$

then $$P^{-1}(t) = [\Phi P \Phi^T + R]^{-1} = \Lambda(t) , \tag{3.35}$$

defining $$\Phi^{-1T} P^{-1} \Phi^{-1} = R_1 = \Phi^{-1T} \Lambda \Phi^{-1} \tag{3.36}$$

and $$R = \sum_{i=1}^{j} s_i s_i^T . \tag{3.37}$$

Then the j-step equation, $$B_{k+1} = B_k - B_k s_i (s_i^T B_k s_i + 1)^{-1} s_i^T B_k , \quad k = 1, j \tag{3.38}$$

yields the result that $$P^{-1}(t) = B_{j+1} = \Lambda(t) . \tag{3.39}$$

Note that the adjoint equations

$$[\dot{\Phi}^{-1}]^T = -P^T(t)[\Phi^{-1}]^T \Phi^{-1T}(t_0, t_0) = I , \tag{3.40}$$

may be used in place of the transition matrix equations (2.32). This avoids any requirement for matrix inversion. It should therefore be obvious that, if the square root matrix C is defined such that

$$CC^T = \Lambda , \tag{3.41}$$

then the square root matrix C may be updated in time using Potter's factorization of (3.38). This method was given before as part of Method No. 1 for the time update of the square root of the covariance matrix. The quantity ΦW was inverted before, using Potter's algorithm. Here, no inversions are required, however, since the information matrix formulation is used.

3.6.2 Measurement Update

In this instance the square root of an equation of the form

$$\Lambda_a = \Lambda + H^T Q^{-1} H \tag{3.42}$$

is to be calculated. The measurements are defined by

$$y = Hx + q$$

and

$$E(qq^T) = Q .$$

Quite obviously this is the identical problem to the time update of the square root of the covariance matrix; hence, if one defines

$$A_1 = [C \vdots H^T \sqrt{(Q^{-1})}] , \tag{3.43}$$

and uses equations (3.24) to (3.27), the columns of the new lower triangular matrix C are given by the w_k of Equation (3.27).

3.6.3 State Change at Measurements

It is well known that the gain K of Equation (2.34) may also be written in the form

$$K = P_a H^T Q^{-1} . \tag{3.44}$$

The C obtained after including measurements is such that

$$[CC^T] = P_a^{-1} , \tag{3.45}$$

if C has no zero column vectors. The matrix C will be of full rank if the problem is initialized with a positive definite covariance matrix. This initialization is *not* necessary in the inverse formulation, so the gain K of Equation (3.44) should be written as

$$K = [CC^T]^{\#} H^T Q^{-1} . \tag{3.46}$$

In (3.46) the (#) symbol reads the pseudo-inverse. The fact that an *a priori* information matrix is not required with this formulation can be an advantage in some problems.

As a result of the manner of calculating C, the rank of the information matrix is available:

rank = n − number of diagonal terms of C which are identically zero.

Hence, if K is calculated by (3.46), the knowledge of when the pseudo-inverse is required is available. Also, since C is lower triangular, inversion of C directly (when it is full rank) is recommended. The recommended manner of computing the state change is

$$K(Y - \hat{Y}) = (D^{-1})^T (D^{-1} (H^T Q^{-1} (Y - \hat{Y}))) , \tag{3.47}$$

Calculations are to be carried out in (3.47) starting with innermost bracketed terms. In (3.47)

$$D^{-1} = C^{-1} , \quad \text{if } C \text{ is invertible;}$$

otherwise,

$$D^{-1} = [F^T F]^{-1} C , \tag{3.48}$$

The matrix F in (3.48) is defined by shifting all non-zero columns of C to the left to form a lower dimension matrix. In (3.48) D^{-1} is the square root of the pseudo-inverse of the information matrix.

3.6.4 Discussion

The inverse formulation given here is approximately the same as that of Reference 14. The method given uses Potter's equation for the time update, which is believed superior to the method of Reference 14 from computational considerations. Reference 15 gives this measurement update square root technique for purposes of improving numerical accuracy of the weighted least square type filter. In this latter instance the estimated states are held at a fixed epoch and all the measurements during a given interval are processed in a "batch" scheme. That is, the state is changed only as a result of all the measurements in the batch.

Reference 14 contains several examples which illustrate the superior accuracy of the square root formulation. An example is as follows.

Let
$$C = \begin{bmatrix} 1 & 1 \\ 1 & 1+e \end{bmatrix}, \tag{3.49}$$

where e is some small number. If CC^T is formed, then

$$CC^T = \begin{bmatrix} 2 & 2+e \\ 2+e & 2+2e+\underline{e^2} \end{bmatrix}. \tag{3.50}$$

The vector outer product given by

$$\begin{pmatrix} \sqrt{2} \\ (2+e)/\sqrt{2} \end{pmatrix} (\sqrt{2} \;\; (2+e)/\sqrt{2}) = \begin{bmatrix} 2 & 2+e \\ 2+e & 2+2e+\underline{e^2/2} \end{bmatrix} \tag{3.51}$$

is also formed. Comparing (3.50) and (3.51) shows that the only way of distinguishing rank (2) in (3.50) is the underlined e^2 as opposed to the $e^2/2$ of (3.51). If an 8 decimal digit computer is considered, then quite obviously when C is squared, (3.50), rank is lost if e is smaller than about 4×10^{-4}. If C is reduced to lower triangular form, using previous algorithms, then

$$C' = \begin{bmatrix} \sqrt{2} & 0 \\ (2+e)/\sqrt{2} & e/\sqrt{2} \end{bmatrix}.$$

An examination of the algorithm shows that the relative size of e to unity is important, rather than e^2. Numerical rank would not be lost here unless e is less than approximately 1×10^{-7} for an 8 decimal digit computer.

Another point of importance is that, once C has full rank, it appears impossible to lose rank numerically by adding very accurate observations. This is not true of the information matrix formulation.

4. TECHNIQUES FOR COMPENSATION OF MODELING ERRORS

The statement of the problem of Section 2 was made in a manner to emphasize the fact that many types of mathematical modeling errors can exist. As was mentioned in the real time onboard problem, one desires to minimize computational requirements. Hence, the modeling problem is further aggravated by the desire to keep calculations and word lengths to a minimum. As a result;

(a) The number of state variables is minimized by omitting as many error sources as possible.

(b) Closed form approximations are used as much as possible to minimize computer speed requirements.

(c) The minimum word length may be sought for reducing computer costs, size, and weight.

(d) Fixed point arithmetic is desirable from computer complexity considerations.

The overall accuracy attainable depends on

(A) The frequency and accuracy of, and the state space spanned by, the observations.

(B) The accuracy of the models used in fitting the data.

(C) The magnitude and spectrum of random forcing functions acting on the system.

The interrelationships between computational requirement considerations and attainable accuracy are very complex in practical problems.

In this section techniques for dynamic model compensation in Kalman filters are addressed. The techniques presented will tend to "lock" the estimate of the state to the most recent measurements and thus prevent a divergence of the estimate from the measurements. This divergence occurs when one attempts to fit data over relatively long time arcs with a poor dynamic model.

4.1 Orbit Determination Example

As an example of model error, an orbit determination problem where position and velocity of the spacecraft are the only assumed state variables is considered. If such a constraint is imposed on the number of state variables, then it is easy to show that model errors generally make it impossible to find a solution for the position and velocity which fits the actual data.

In this instance

$$\dot{\hat{X}} = F(\hat{X}, t) \tag{4.1}$$

$$\hat{Y} = G(\hat{X}, t). \tag{4.2}$$

The problem of finding the estimate of the initial state, $\hat{X}_0$, which minimizes the loss function

$$L = (Y - \hat{Y})^T Q^{-1} (Y - \hat{Y}) \tag{4.3}$$

is considered.

Equation (4.1) imposes a constraint on the time history of the solution for $\hat{X}$ if $\hat{X}_0$ is given. Equation (4.2) gives the functional relationship for computing the observation from the estimate of $\hat{X}$.

The real measurements obey reality rather than (4.1) and (4.2). Hence, if one uses Equations (4.1) and (4.2), a value of $\hat{X}_0$ for which the residuals $Y - \hat{Y}$ are small and random about zero mean, over all time, implies there are no errors in the mathematical models.

Since a weighted least squares fit is used in this example, one can hypothesize the typical residual behavior of the solution for errors in (4.1) and (4.2). A pictorial representation is shown in Figure 1. The term "short arc" means, for example, the data during a single pass of a ground station for a low altitude satellite.

As shown in Figure 1, for a short arc, dynamic model errors (errors in Equation (4.1)) will not prevent a solution for $\hat{X}_0$ where residuals are essentially random with zero mean. This can be seen by considering a Taylor series solution of (4.1). Let $\hat{X}_0$ represent the three position components at the beginning of a data arc. Then

$$\hat{X}(t - t_0) = \hat{X}_0 + \hat{\dot{X}}_0 (t - t_0) + \hat{\ddot{X}}_0 \left(\frac{t - t_0}{2!}\right)^2 + \underline{\quad\quad} \tag{4.4}$$

In the equation, $\hat{X}_0$ and $\hat{\dot{X}}_0$ are free to be estimated, and a nearly perfect fit can be assured as long as $(t - t_0)$ is small, regardless of the constraining effect of the dynamic model.

The underlined terms of the infinite series (Equation (4.4)) are dependent on the dynamic model; hence, errors in the dynamic model will become effective over long time arcs. One therefore will not be able to find a value of $\hat{X}_0$ and $\hat{\dot{X}}_0$ which produces small residual errors, for example, over a multiple pass. Dynamic model errors for data over long time arcs should cause effects of the type shown in Figure 1(b).

Measurement errors such as biases will also have only a small effect, for short time arcs, as shown in Figure 1(c). The error in the estimate of state will include the bias caused by the measurement model error in this instance. Only when measurements of many different types are processed should biased offset residuals occur for short data arcs. This latter results from the fact that no single values of $\hat{X}_0$ and $\hat{\dot{X}}_0$ can be selected which fit multiple measurements (more than 6) in the presence of biases or other measurement model errors. A measurement model error, however, is, in general, bounded. Hence, if a long data arc over many stations is visualized, measurement model errors would be expected to cause residuals as shown in Figure 1(d). Since the solution is constrained to obey the dynamic model, measurement model errors will have less and less effect on orbit determination accuracy as more data is processed. In these examples the errors are assumed to be small (not mistakes). Typical measurement model errors are station location errors, timing errors and biases.

The real problem, of course, includes both measurement and dynamic model errors. Dynamic errors may have both periodic and secular effects on the residuals. Hence Figure 1 can only indicate trends to be expected. Also, in the general situation, what is meant by short and multiple pass is not readily defined.

Figure 2 depicts general trends of the effects of measurement and dynamic model errors on state estimation accuracy. The figure is intended to indicate that dynamic errors cause a ever-increasing state estimation error as more data points, over a large time arc, are included. Measurement errors have the opposite effect.

One desires to have a compensation technique which gives the best balance between these two effects. This will be called dynamic model compensation and will be discussed in the next section. One should recognize, however, that both measurement and dynamic model errors exist and the resulting effect of the dynamic compensation can cause the measurement errors to be dominant.

4.2 Dynamic Model Compensation

As indicated in the previous section, some means of preventing the state estimation errors from growing indefinitely is necessary. It should be reasonably obvious that one way of preventing this growth is to gradually reduce the influence of past measurements in determining the estimate from current measurements. Assume that totally false measurements (outliers) caused by malfunctions can be removed. The remaining measurements can then be treated to give the theoretical value based on the correct model plus an error term. It is reasonably likely that upper bounds may be placed on the measurement error. That is, the error which is caused, for example, by calibration, atmospheric effects, receiver noise, and so forth, cannot be larger than some specified number. If this is the case, then the error in the computed measurement should be no worse than this upper bound.

As an example, suppose the measurement is the propagation delay of an r.f. signal between transmittal and reception from a spacecraft. It is reasonable to assume that the two-way range may be computed from this measurement and the upper bound on the error in this computed quantity prescribed. If the residual is greater than this bound, one must conclude that too much emphasis has been placed on past data in establishing the computed orbit. The author does not mean to imply that such "worst case" considerations are to be used. These remarks

are only made to give some reasons why a reasonable degree of trust in the basic measurements must be made to develop a compensation scheme.

Two different approaches for weighting out the influence of past data in the determination of the current estimate are as follows:

(i) Increase the *a priori* covariance matrix (or decrease the information matrix). Arguments for using this approach are as follows: Equations of motion which are invalid have been used to update the estimate of state. One should therefore increase the *a priori* covariance matrix in accordance with the errors involved in the time-updating of the estimate. The difficulties in using this approach lie in defining the real error sources. Their formulation can also become extremely complex. Hence, for practical usages, it is perhaps better to say that pseudo-errors are introduced to cause an increase in the *a priori* covariance matrix. These pseudo-errors can be of two types:

(a) random forcing functions,

(b) errors attributed to inaccuracies of constants in the equations of motion.

(ii) Overweight the most recent data. In this approach it is also recognized that the *a priori* covariance matrix may be overly optimistic. The matrix, however, is not modified on the basis of adding the effects produced by pseudo-errors in the dynamic model. Instead, a non-optimal filter algorithm is adopted which attaches a greater significance to the recent observations than the optimal filter does. The *a posteriori* covariance matrix is modified to conform with the non-optimal algorithm.

Either of these approaches has many variations to suit any specific problem. Although the philosophy is different, both provide a means of developing a dynamic model compensation technique.

4.2.1 *Pseudo-Random Forcing Function Approach*

In the equation for time update of the covariance matrix, (2.33), the quantity R represents an added growth in error caused by random forcing functions. Since dynamic model or computation errors would cause an additional error to exist, the natural fix is to increase R. As already mentioned, if

$$R = \sum_{i=1}^{j} s_i \sigma_i^2 s_i^T , \tag{4.5}$$

then the direction vectors s_i and the variances σ_i^2 are to be selected for compensation of the dynamic model errors.

In the real time onboard problem it is likely that a far more accurate model can be analytically defined than is possible to calculate. An approach for defining the errors is to use a general purpose computer to calculate the difference between the assumed "exact" solution and the approximate solution as computed by a simulated onboard computer. This difference (or error) data is calculated for a number of initial conditions and time durations. The quantities s_i and σ_i^2 are then adjusted to give a reasonable approximation of the error growth. Such a simulation is likely essential in validating the approximations and onboard computational techniques. As a result this approach should not require a large additional effort. The alternate approach of defining s_i and σ_i^2 (as well as j) by "cut and try" experiments with the real system is definitely *not* recommended for *complex* systems.

Model errors and computational errors generally cause bias type errors in the estimate. The error sources may be modeled and their effects included in the filter in a more direct manner. The theory for such a technique is given in Reference 16. For measurement type error sources this technique is reasonably efficient. For dynamic errors, however, it requires extensive calculations and does not necessarily give large reductions in state estimation error over the pseudo-random approach[3].

4.2.2 *Direct Overweighting of Most Recent Data*

The pseudo-random forcing approach gives a higher weighting to the more recent measurements by causing an additive covariance matrix growth between measurements. There is no reason why one should not consider modifications of the filter algorithm for achieving the overweighting in a direct manner.

One such modification is shown in (4.6). This modification is for sequential processing of the observations (one at a time).

$$\hat{X}_a = \hat{X}_b + (PH^T + H^T Q\epsilon / HH^T)(Y - \hat{Y}_b)/(HPH^T + Q) . \tag{4.6}$$

The scalar ϵ in Equation (4.6) is a control gain.

If linearity prevails, then $H\hat{X}_a = \hat{Y}_a$ = the computed measurement after the observation. Multiplication of (4.6) by H gives

$$H\hat{X}_a = H\hat{X}_b + \left(\frac{HPH^T + Q\epsilon}{HPH^T + Q}\right)(Y - \hat{Y}_b) . \tag{4.7}$$

Note in (4.7) that, for $\epsilon = 1$, the estimate of the observation is equal to the observation. Hence, the error in this component of the estimate of state is no worse than the error in the measurement. For $\epsilon = 0$ the optimal filter is obtained.

This modification was suggested by considering the form for the estimate of state using a single measurement. That is, if

$$y = Hx + q , \tag{4.8}$$

then, using the pseudo-inverse to calculate the estimate,

$$\hat{x} = H^{\#}y = \frac{H^T y}{HH^T} . \tag{4.9}$$

As may be noted in Equation (4.6), part of the change in X_a is proportional to Equation (4.9).

The covariance matrix of errors in $\hat{X}_a$ is found by taking

$$E(X - \hat{X}_a)(X - \hat{X}_a)^T = P_a .$$

This operation applied to (4.6) gives

$$P_a = P_b - P_b H^T H P_b / (H P_b H^T + Q) + \underline{\epsilon^2 Q^2 H^T H / [(H P_b H^T + Q)(H H^T)^2]} \tag{4.10}$$

The underlined term of (4.10) shows the additive error (above optimal filtering) caused by the control gain ϵ.

Consider the following simple example, using this non-optimal filter.

Actual Model	*Assumed Model*
$\dot{x} = 0.1$ = a scalar	$\dot{\hat{x}} = 0$
$y = x$	$\hat{y} = \hat{x}$.

Let $Q = 1$.

$P(0+) = Q$ (after first measurement)
$x(0+) = 10$ (after first measurement)
$\hat{x}(0+) = x(0+)$.

Figure 3 shows the estimate of state after the observations at two second intervals are included. The estimate between observations remains constant.

The error in the estimate for the Kalman filter grows indefinitely with time. The growth arises because $x(t)$ was assumed to be a constant, while in reality $x(t)$ is equal to a constant plus a time-dependent term. The modified filter ($\epsilon \neq 0$) has an error growth between measurements. However, if $\epsilon = 1$, the error in $\hat{x}$ after the measurement is included is no greater than the measurement error. For $\epsilon < 1$, the error in estimate, after each measurement is processed, has a bias offset from the measured value. Hence the use of this non-optimal filter tends to look the estimate of state to the recent observations.

Equivalent results to those given in Figure 3 can be obtained by introducing a pseudo-random forcing function. This can readily be seen, since random forcing functions cause an added growth to P between measurements. Prior to including the n^{th} measurement for this example,

$$P_n(-) = P_{n-1} + R ,$$

if R were made extremely large compared to Q, then the results corresponding to $\epsilon = 1$ of Figure 3 would be obtained.

4.2.3 *Discussion*

Two techniques for dynamic model error compensation have been briefly described. In the applications with which the author is familiar either approach will prevent the so-called divergence[3]. The bad features of requiring "out and try" are common to both techniques.

One should note that both techniques can be used to prevent loss of positive definiteness of the covariance matrix. For this application the overweighting-of-most-recent-data approach has been easier (from the author's experience) to use.

Other alternatives for model compensation, such as described in References 7 and 8, have not been covered. These alternatives, although applicable, lead to reasonably complex calculations.

5. CONCLUDING REMARKS

The over-all development of "software" for real time onboard applications of Kalman filtering theory is a very complex subject. This chapter described approaches and algorithms for solving several of the problems which occur in these type applications as well as applications where large general purpose computers are available. There are, however, many other problem areas which must be addressed and resolved in the development of "appropriate" software.

One of the most difficult problems is the definition of the appropriate mathematical models. These models must be sufficiently complex to satisfy the estimation accuracy specifications yet simple enough for the onboard computer. Cleverness in the selection of coordinate systems and fundamental state variables can lead to simple formulations. Approaches which lead to these simplifications are however more of an "art" than a science.

The filter algorithms considered in this chapter were of the "discrete" type. There are instances where continuous filter theory may lead to simpler formulations. Propagation of the covariance matrix between observations in problems where the transition matrix is not required is one possibility[17].

For high data rate problems the filter equations derived from either the discrete or continuous theory may be too complex for the real time computer. Data averaging or data compression techniques appear to offer a good compromise for handling such problems[10]. In applying such techniques some information is lost. In addition, improper formulation or inadequate validation of approximations can lead to unwanted bias type errors in the estimate[18]. The "software" designer should therefore exercise a good deal of caution in the application of such techniques.

From the material presented it is obvious that many practical solutions of filtering problems exist. It is also obvious that the theory available will not provide a uniquely optimal solution when all factors are considered.

REFERENCES

1. Smith, G.L. et al. — *Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity On Board a Circumlunar Vehicle.* NASA TR R-135, 1962.

2. Schmidt, S.F. — *State Space Techniques Applied to the Design of a Space Navigation System.* JACC, New York University, 1962.

3. Schlee, F.H. et al. — *Divergence in the Kalman Filter.* AIAA Journal, Vol.5 1967, pp.1114-1120.

4. Schmidt, S.F. — *Estimation of State with Acceptable Accuracy Constraints.* Report 67-4, Analytical Mechanics Associates, January 1967.

5. Schmidt, S.F. — *Compensation for Modelling Errors in Orbit Determination Problems.* Report 67-16, Analytical Mechanics Associates, November, 1967.

6. Pines, S. et al. — *Modifications of the Minimum Variance Program for the Processing of Real Data.* Analytical Mechanics Associates Technical Report, October 1964.

7. Jazwinski, A.H. — *Limited Memory Optimal Filtering.* Report 67-11, Analytical Mechanics Associates, August, 1967.

8. Jazwinski, A.H. Bailie, A.E. — *Adaptive Filtering.* Report 67-6, Analytical Mechanics Associates, March 1967.

9. Kalman, R.E. — *A New Approach to Linear Filtering and Prediction Problems.* Journal of Basic Engineering, March 1960, pp.33-45.

10. Schmidt, S.F. et al. — *Case Study of Kalman Filtering in the C-5 Aircraft Navigation System.* Institute of Electrical and Electronic Engineers Case Studies Seminar, Ann Arbor, Michigan. June 1968.

11. Faddeev, D.K. Fadeeva, V.N. — *Computational Methods of Linear Algebra.* Freeman, San Francisco, 1963, pp.144-147.

12. Battin, R.H. — *Astronautical Guidance.* McGraw-Hill, 1964, pp.339-340.

13. Bellantoni, J.F. Dodge, K.W. — *A Square Root Formulation of the Kalman-Schmidt Filter.* AIAA Journal, Vol.5 1967, pp.1309-1314.

14. Dyer, P. McReynolds, S.R. — *Computational Accuracy of Square Root Filtering.* Space Programs Summary 37-54, Vol.III, pp 15-19, for period October 1 to November 30, 1968, Jet Propulsion Laboratory, December 31, 1968.

15. Schmidt, S.F. — *Final Report for Mission Analysis Guidance Study.* Analytical Mechanics Associates Report 68-23, January 1969.

16. Schmidt, S.F. — *Application of State Space Methods to Navigation Problems.* "Advances in Control Systems", Volume 3, Academic Press, 1966.

17. Gunckel, T.L. — *Guidance and Control of Reentry and Aerospace Vehicles.* "Advances in Control Systems", Volume 3, Academic Press, 1966, pp.20-21.

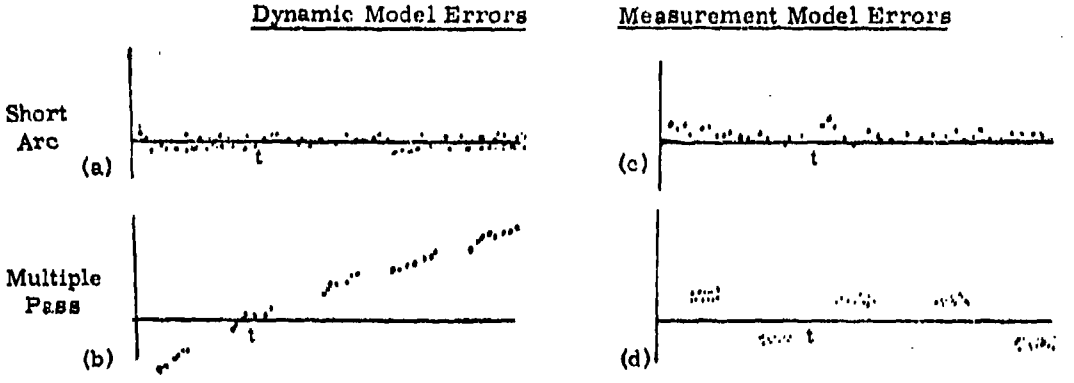


Fig. 1 Characteristics of residuals resulting from model errors

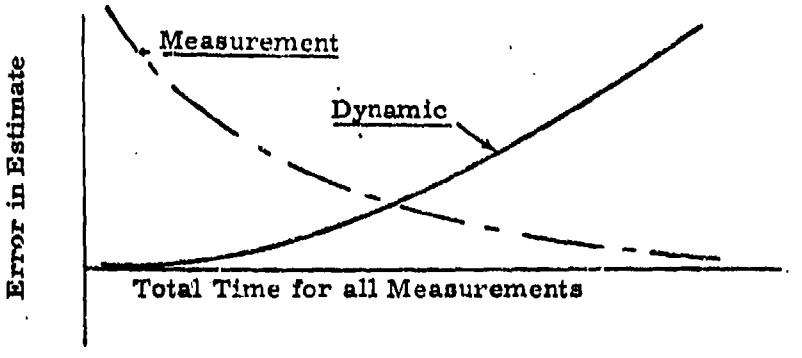


Fig. 2 State estimate error trends resulting from model errors

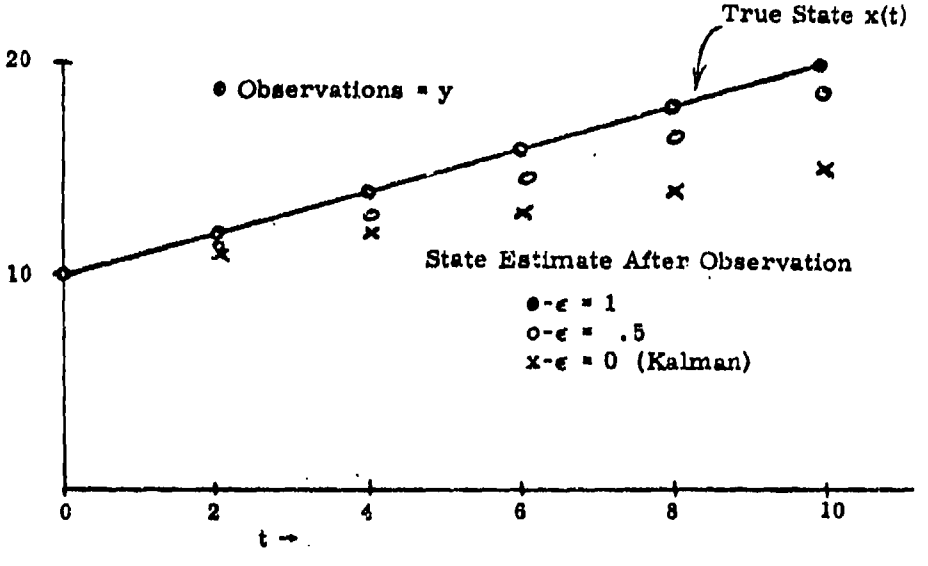


Fig. 3 Example results using model compensation in a Kalman filter

CHAPTER 4 - MODELING ERRORS IN KALMAN FILTERS

by

T. Nishimura

Jet Propulsion Laboratory, California Institute of Technology,
Pasadena, California, USA

CHAPTER 4 - MODELING ERRORS IN KALMAN FILTERS

T.Nishimura

1. INTRODUCTION

In recent years the Kalman filter[1,2] has been extensively used in such applications as the tracking of missiles or planes and the orbit determination of spacecraft. One of the problems arising in these applications is that very often a precise knowledge of the *a priori* statistics of initial conditions and of the noise model (process noise and observation noise), as well as those of system models, are not available, while it is essential for the design of optimal filters.

For example, in the orbit determination problem of spacecraft in deep space, the observation is usually supplied in the form of Doppler, counted Doppler or range data. These data are subject to oscillator instability, disturbances in the ionosphere, receiver noise, and quantization noise of the counter, which together constitute the observation noise.

Also the spacecraft undergoes, during its long journey to the planet (e.g., about 200 days for a typical Mars mission), various unknown disturbances originating from solar pressure, impacts of meteorites, and fuel leakage from valves. It is a rather difficult task to determine the statistics of any one of these noise sources. Besides the uncertainty of the injection conditions of the spacecraft, the coordinates after the midcourse maneuver may enter into the filter design and influence the gain of the filter strongly during the initial period of estimation.

Errors are inevitable in assigning *a priori* covariance matrices of large dimensions, because of lack of sufficient experience or of incapability of analyzing complex correlations among parameters.

Furthermore lack of precise knowledge of system models is a problem which practising engineers frequently encounter in designing filters. This matter is also closely related to the problem of identification, which is another major topic in control theory and its applications.

In this chapter attention is focused on the analysis of the effect of errors in these *a priori* statistics and system models on the performance of the resulting suboptimal filters. Both continuous systems and discrete systems are analyzed and an effort is made to find the upper and lower bound for the error covariances of these suboptimal filters.

Throughout this chapter it is assumed that the systems are linear and that the stochastic variables have gaussian distributions.

2. ANALYSIS FOR CONTINUOUS SYSTEMS

The basic process is described by a first-order differential equation in vector form,

$$\frac{dx(t)}{dt} = F(t)x(t) + G(t)w(t) . \tag{1}$$

The observation is

$$y(t) = H(t)x(t) + n(t) , \tag{2}$$

where $x(t)$ is an n_x vector of states, with

$$E[x(0)] = 0 .$$

$y(t)$ is an n_y vector of observations.

$w(t)$ is an n_w vector of stochastic inputs to the process with

$$E[w(t)] = 0 \tag{3}$$

$$E[w(t)w'(\tau)] = Q(t)\delta(t-\tau) . \tag{4}$$

where $\delta(t)$ is the Dirac delta function.

$n(t)$ is an n_n vector of the observation noise with

$$E[n(t)] = 0 \tag{5}$$

$$E[n(t)n'(\tau)] = R(t)\delta(t-\tau) . \tag{6}$$

$F(t)$, $G(t)$, $H(t)$ are $n_x \times n_x$, $n_x \times n_w$, $n_y \times n_x$ matrices respectively.

$E[\,]$ is an expected value operator on stochastic variables. It is also assumed that the process noise w and observation noise n have no correlation to each other.

$$E[w(t)n'(\tau)] = 0 . \tag{7}$$

Then the optimal estimator $x^*(t)$ of $x(t)$ which minimizes

$$E[\|x^* - x\|^2] ,$$

having the observation $y(t)$ from $t = 0$ to t, is described by the differential equation[7]

$$\frac{dx^*(t)}{dt} = F(t)x^*(t) + K(t)[y(t) - H(t)x^*(t)] , \tag{8}$$

where

$$K(t) = P(t)H'(t)R^{-1}(t) . \tag{9}$$

It is assumed that $R(t)$ is positive definite for $t \geq 0$.

Then the covariance matrix $P(t)$ is defined by

$$P(t) \triangleq E[(x^*(t) - x(t))(x^*(t) - x(t))'] \tag{10}$$

and it is obtained as a solution of a matrix Riccati equation

$$\frac{dP(t)}{dt} = F(t)P(t) + P(t)F'(t) - P(t)H'(t)R^{-1}(t)H(t)P(t) + G(t)Q(t)G'(t) . \tag{11}$$

The initial conditions for Equations (8) and (11) are respectively

$$x^*(0) = 0 \tag{12}$$

$$P(0) = E[x(0)x'(0)] . \tag{13}$$

3. ASSUMPTIONS AND MATHEMATICAL DERIVATIONS

The optimal estimator described in the previous section is based on correct information for initial conditions, noise covariances and coefficient matrices. Suppose the estimator design is based on incorrect information with respect to these quantities, namely,

(i) the incorrect $P_c(0)$ rather than the correct $P(0)$ (*a priori* covariance of states),

(ii) the incorrect $Q_c(t)$ rather than the correct $Q(t)$ (covariance of the process noise),

(iii) the incorrect $R_c(t)$ rather than the correct $R(t)$ (covariance of the observation noise),

(iv) the incorrect $F_c(t)$ rather than the correct $F(t)$ (process matrix)

(v) the incorrect $G_c(t)$ rather than the correct $G(t)$ (coefficient matrix of the process noise),

(vi) the incorrect $H_c(t)$ rather than the correct $H(t)$ (observation matrix). The resulting matrix is no longer then an optimal one, but becomes suboptimal.

This suboptimal estimator is denoted $x_s^*(t)$ and it is described by

$$\frac{dx_s^*(t)}{dt} = F_c(t)x_s^*(t) + K_c(t)[y(t) - H_c(t)x_s^*(t)] , \tag{14}$$

where

$$x_s^*(0) = 0 \tag{15}$$

$$K_c(t) = P_c(t)H_c'(t)R_c^{-1}(t) , \tag{16}$$

and the calculated covariance $P_c(t)$ is computed by the same Riccati equation as Equation (11), but using the incorrect model specified by (i) - (vi).

$$\frac{dP_c(t)}{dt} = F_c(t)P_c(t) + P_c(t)F_c'(t) - P_c(t)H_c'(t)R_c^{-1}(t)H_c(t)P_c(t) + G_c(t)Q_c(t)G_c'(t) . \tag{17}$$

The actual covariance $P_a(t)$ is defined as the error covariance associated with the suboptimal estimator of Equation (14).

Hence
$$P_a(t) \triangleq E[(x_a^*(t) - x(t))(x_a^*(t) - x(t))'] . \tag{18}$$

This is the covariance expected on the estimator when there is insufficient information on the design parameters. It is the main objective of this section to derive equations describing this $P_a(t)$.

For this purpose it is easier to derive a differential equation for $P_a(t)$.

Thus differentiation of $P_a(t)$ of Equation (18), and a change in the order of the differentiating operator and the expected value operator, yield

$$\dot{P}_a(t) = E[(\dot{x}_a^*(t) - \dot{x}(t))(x^*(t) - x(t))'] + E[(x_a^*(t) - x(t))(\dot{x}_a^*(t) - \dot{x}(t))'] . \tag{19}$$

However, from Equations (1) and (14),

$$\dot{x}_a^*(t) - \dot{x}^*(t) = (F_c(t) - K_c(t)H_c(t))(x_a^*(t) - x(t)) + \Delta F(t)x(t) - K_c(t)\Delta H(t)x(t) + K_c(t)n(t) - G(t)w(t) , \tag{20}$$

where
$$\Delta F(t) = F_c(t) - F(t) \tag{21}$$

$$\Delta H(t) = H_c(t) - H(t) . \tag{22}$$

Also, $x(t)$ is obtained from Equation (1) as

$$x(t) = U(t,0)x(0) + \int_0^t U(t,s)G(s)w(s)\,ds , \tag{23}$$

where $U(t,s)$ is defined by

$$\frac{\partial U(t,s)}{\partial t} = F(t)U(t,s) , \tag{24}$$

with
$$U(s,s) = I , \qquad t \geq s \geq 0 , \tag{25}$$

and I is an identity matrix.

Furthermore $x_a^*(t)$ is derived from Equation (14) as

$$x_a^*(t) = \int_0^t V_c(t,s)K_c(s)y(s)\,ds , \tag{26}$$

where $V_c(t,s)$ is defined by

$$\frac{\partial V_c(t,s)}{\partial t} = (F_c(t) - K_c(t)H_c(t))V_c(t,s) , \qquad t \geq s \geq 0 . \tag{27}$$

When $\dot{x}_a^*(t)$ and $x_a^*(t)$ are substituted into Equation (19), together with $\dot{x}(t)$ of Equation (20) and its solution $x(t)$ in Equation (23), paying attention to the fact that $w(t)$ and $n(t)$ are uncorrelated white noises, the following three differential equations are derived[8]:

$$\frac{dP_a(t)}{dt} = (F_c(t) - K_c(t)H_c(t))P_a(t) + P_a(t)(F_c(t) - K_c(t)H_c(t))' + (\Delta F(t) - K_c(t)\Delta H(t))\Lambda(t) + \Lambda'(t)(\Delta F(t) - K_c(t)\Delta H(t))' + K_c(t)R(t)K_c'(t) + G(t)Q(t)G'(t) \tag{28}$$

$$\frac{d\Lambda(t)}{dt} = F(t)\Lambda(t) + \Lambda(t)(F_c(t) - K_c(t)H_c(t))' + P_z(t)(\Delta F(t) - K_c(t)\Delta H(t))' - G(t)Q(t)G'(t) \tag{29}$$

$$\frac{dP_z(t)}{dt} = F(t)P_z(t) + P_z(t)F'(t) + G(t)Q(t)G'(t) , \tag{30}$$

where $\Lambda(t)$ and $P_s(t)$ are defined by

$$\Lambda(t) \triangleq E[x(t)(x_s^*(t) - x(t))'] \tag{31}$$

$$P_s(t) \triangleq E[x(t)x'(t)] . \tag{32}$$

The initial conditions for these differential equations are respectively given by

$$P_s(0) = P(0) \tag{33}$$

$$\Lambda(0) = -P(0) \tag{34}$$

$$P_s(0) = P(0) . \tag{35}$$

4. ERROR BOUNDS OF SUBOPTIMAL FILTERS (CONTINUOUS CASE)

When both process matrix F and observation matrix H are known correctly, the results in the previous section can be simplified considerably. In fact, only the first differential equation needs to be solved[7].

$$\frac{dP_s(t)}{dt} = (F(t) - K_c(t)H(t))P_s(t) + P_s(t)(F(t) - K_c(t)H(t))' + K_c(t)R(t)K_c'(t) + G(t)Q(t)G'(t) , \tag{36}$$

with

$$K_c(t) = P_c(t)H'(t)R_c^{-1}(t) . \tag{37}$$

$$P_s(0) = P(0) , \tag{38}$$

Let $E_{cs}(t)$ be the difference between the computed covariance $P_c(t)$ and the actual covariance $P_s(t)$, then

$$E_{cs}(t) = P_c(t) - P_s(t) . \tag{39}$$

Substituting this into Equation (36) yields the following differential equation of $E_{cs}(t)$, with the aid of Equation (17):

$$\dot{E}_{cs}(t) = (F(t) - K_c(t)H(t))E_{cs}(t) + E_{cs}(t)(F(t) - K_c(t)H(t))' + K_c(t)\Delta R(t)K_c'(t) + G(t)\Delta Q(t)G'(t) , \tag{40}$$

where $\Delta R(t)$ and $\Delta Q(t)$ are the differences between the incorrect noise covariances and the correct ones, namely

$$\Delta R(t) = R_c(t) - R(t) \tag{41}$$

$$\Delta Q(t) = Q_c(t) - Q(t) . \tag{42}$$

Since Equation (39) is a linear differential equation, an explicit analytic solution can be derived.

$$E_{cs}(t) = V_c(t,0)E_{cs}(0)V_c'(t,0) + \int_0^t V_c(t,s)K_c(s)\Delta R(s)K_c'(s)V_c'(t,s)\,ds + \int_0^t V_c(t,s)G(s)\Delta Q(s)G'(s)V_c'(t,s)\,ds . \tag{43}$$

As observed from the Equation (43), $E_{cs}(t)$ is a sum of real symmetric matrices, so that if is semi-positive definite provided every term in the right-hand side of Equation (43) is semi-positive definite. Because of the specific (symmetric) configuration of these terms, every one of them will be respectively semi-positive definite if every matrix at the center of the respective term, namely $E_{cs}(0)$, $\Delta R(s)$, and $\Delta Q(s)$ for $t \geq s \geq 0$, is semi-positive definite.

Based on this discussion, the following theorem can be derived[7].

Theorem 1

$E_{cs}(t) \geq 0$; hence $P_c(t) \geq P_s(t)$ for $t \geq 0$ if the condition C-I is satisfied.

C-I : $E_{cs}(0) \geq 0$, $\Delta Q(t) \geq 0$, and $\Delta R(t) \geq 0$, or equivalently $P_c(0) \geq P_s(0)$, $Q_c(t) \geq Q(t)$, and $R_c(t) \geq R(t)$ for $t \geq 0$.

The implication of the inequality sign (plus equality sign), e.g., $P_c(t) \geq P_s(t)$, is that the difference matrix $P_c(t) - P_s(t)$ is semi-positive definite.

Therefore, an upper bound for the variances of the suboptimal estimator $x_a^*(t)$ can be set, and it is equal to the diagonal components of the calculated covariance $P_c(t)$ when the condition C-I is satisfied. The lower bound of those variances is, of course, zero. So let $p_{cii}(t)$ and $p_{aii}(t)$ be respective diagonal components of $P_c(t)$ and $P_a(t)$; then

$$p_{cii}(t) \geq p_{aii}(t) \geq 0 . \tag{44}$$

Even though the designer does not know exactly the *a priori* statistics, he can expect that his suboptimal estimator will properly behave within the specified range provided the conservative condition C-I is satisfied.

Though it is of less practical importance, the following Corollary is derived from Equation (43).

Corollary 1

$E_{ca}(t) \leq 0$; hence $P_c(t) \leq P_a(t)$ for $t \geq 0$ if the condition C-II is satisfied.

C-II: $E_{ca}(0) \leq 0$, $\Delta Q(t) \leq 0$, and $\Delta R(t) \leq 0$ or, equivalently, $P_c(0) \leq P_a(0)$, $Q_c(t) \leq Q(t)$, and $R_c(t) \leq R(t)$ for $t \geq 0$.

Furthermore, differential equations associated with the other two error matrices, $E_{a0}(t)$ and $E_{c0}(t)$, which are defined as

$$E_{a0}(t) = P_a(t) - P(t) \tag{45}$$

$$E_{c0}(t) = P_c(t) - P(t) , \tag{46}$$

will be derived.

First $\dot{E}_{a0}(t)$ can be obtained as a difference between $\dot{P}_a(t)$ and $\dot{P}(t)$ given by Equations (36) and (11), respectively:

$$\dot{E}_{a0}(t) = (F(t) - K_c(t)H(t))E_{a0}(t) + E_{a0}(t)(F(t) - K_c(t)H(t))' + (K_c(t)R(t) - P(t)H'(t))R(t)^{-1}(K_c(t)R(t) - P(t)H'(t))' . \tag{47}$$

When a similar discussion leading to Theorem 1 is applied to Equation (37) it may be concluded that $E_{a0}(t)$ is always semi-positive definite for all $t \geq 0$, because $R(t)$ is positive definite by assumption and $E_{a0}(0)$ is semi-positive definite, as deduced from the definition of $P(0)$. This is described in the following theorem.

Theorem 2

$$E_{a0}(t) \geq 0 ; \quad \text{hence} \quad P_a(t) \geq P(t) \quad \text{for} \quad t \geq 0 .$$

This result can naturally be expected because $P(t)$ is the minimum variance by definition.

Similarly the differential equation for $E_{c0}(t)$ is derived by subtracting $\dot{P}(t)$ of Equation (11) from $\dot{P}_c(t)$ of Equation (17).

$$\dot{E}_{c0}(t) = (F(t) - K_c(t)H(t))E_{c0}(t) + E_{c0}(t)(F(t) - K_c(t)H(t))' + E_{c0}(t)H'(t)R_c^{-1}(t)H(t)E_{c0}(t) + G(t)\Delta Q(t)G'(t) + P(t)H'(t)(R(t)(\Delta R(t))^{-1}R(t) + R(t)^{-1}H(t)P(t) \quad \text{for} \quad \Delta R > 0 . \tag{48}$$

Again, it is clear from this equation that $E_{c0}(t)$ is semi-positive definite if C-I is satisfied. When $\Delta R = 0$ the same conclusion can be proved by taking a limit $\Delta R \to 0$.

Thus the following theorem is obtained.

Theorem 3

$$E_{c0} \geq 0 ; \quad \text{hence} \quad P_c(t) \geq P(t) \quad \text{for} \quad t \geq 0 , \text{ if C-I is satisfied.}$$

5. ANALYSIS FOR DISCRETE SYSTEMS

In this section the same technique is applied to discrete systems and similar results are derived. Symbols are defined in the same manner as in the continuous systems, and similar assumptions are made concerning modeling errors and noise statistics.

The process equation and observation equations are respectively

$$x(k+1) = \Phi(k)x(k) + G(k)w(k) \tag{49}$$

$$y(k) = H(k)x(k) + n(k) . \tag{50}$$

Then the optimal estimate $x^*(k+1)$ having the information $Y(k) = [y(0), y(1), \ldots, y(k)]$ is given by[1]

$$x^*(k+1) = \Phi(k)x^*(k) + K(k)(y(k) - H(k)x^*(k)) , \tag{51}$$

where

$$K(k) = \Phi(k)P(k)H'(k)[H(k)P(k)H'(k) + R(k)]^{-1} \tag{52}$$

$$x^*(0) = 0 . \tag{53}$$

The covariance matrix $P(k)$ is dufined by

$$P(k) \triangleq E[(x^*(k) - x(k))(x^*(k) - x(k))'] \tag{54}$$

and it is governed by the nonlinear difference equation

$$P(k+1) = (\Phi(k) - K(k)H(k))P(k)(\Phi(k) - K(k)H(k))' + K(k)R(k)K(k)' + G(k)Q(k)G'(k) , \tag{55}$$

with

$$P(0) = E[x(0)x'(0)] . \tag{56}$$

When the incorrect models which are the counterparts in discrete systems of those described in (i) - (vi) are used, the resulting suboptimal estimator $x_a^*(k)$ is computed by

$$x_a^*(k+1) = \Phi_c(k)x_a^*(k) + K_c(k)(y(k) - H_c(k)x_a^*(k)) , \tag{57}$$

with

$$K_c(k) = \Phi_c(k)P_c(k)H_c'(k)[H_c(k)P_c(k)H_c'(k) + R_c(k)]^{-1} \tag{58}$$

$$x_a^*(0) = 0 . \tag{59}$$

The calculated covariance $P_c(k)$ is

$$P_c(k+1) = (\Phi_c(k) - K_c(k)H_c(k))P_c(k)(\Phi_c(k) - K_c(k)H_c(k))' + K_c(k)R_c(k)K'(k) + G_c(k)Q_c(k)G_c'(k) . \tag{60}$$

The actual covariance associated with this suboptimal estimator $x_a^*(k)$ is defined as

$$P_a(k) \triangleq E[(x_a^*(k) - x(k))(x_a^*(k) - x(k))'] . \tag{61}$$

The recurrence equations describing this $P_a(k)$ are derived in a similar manner to the continuous case[10].

$$\begin{aligned} P_a(k+1) = {} & (\Phi_c(k) - K_c(k)H_c(k))P_a(k)(\Phi_c(k) - K_c(k)H_c(k))' + \\ & + (\Delta\Phi(k) - K_c(k)\Delta H(k))\Lambda(k)(\Phi_c(k) - K_c(k)H_c(k))' + \\ & + (\Phi_c(k) - K_c(k)H_c(k))\Lambda'(k)(\Delta\Phi(k) - K_c(k)\Delta H(k))' + \\ & + (\Delta\Phi(k) - K_c(k)\Delta H(k))P_x(k)(\Delta\Phi(k) - K_c(k)\Delta H(k))' + \\ & + K_c(k)R(k)K_c'(k) + G(k)Q(k)G'(k) . \end{aligned} \tag{62}$$

$$\Lambda(k+1) = \Phi(k)\Lambda(k)(\Phi_c(k) - K_c(k)H_c(k))' + \Phi(k)P_x(k)(\Delta\Phi(k) - K_c(k)\Delta H(k))' - G(k)Q(k)G'(k) \tag{63}$$

$$P_x(k+1) = \Phi(k)P_x(k)\Phi'(k) + G(k)Q(k)G'(k) . \tag{64}$$

where $\Lambda(k)$ and $P_x(k)$ are defined by

$$\Lambda(k) \triangleq E[x(k)(x_a^*(k) - x(k))'] \tag{65}$$

$$P_x(k) \triangleq E[x(k)x'(k)] . \tag{66}$$

The initial conditions for these recurrence equations are respectively given by

$$P_a(0) = P(0) \tag{67}$$

$$\Lambda(0) = -P(0) \tag{68}$$

$$P_x(0) = P(0) . \tag{69}$$

6. ERROR BOUNDS OF SUBOPTIMAL FILTERS (DISCRETE CASE)

When the process transition matrix $\Phi(k)$ and the observation matrix $H(k)$ are perfectly known, only the first recurrence equation, Equation (8), needs to be solved in order to find $P_a(k)$. Namely[4-6,8]

$$P_a(k+1) = (\Phi(k) - K_c(k)H(k))P_a(k)(\Phi(k) - K_c(k)H(k))' + K_c(k)R(k)K_c'(k) + G(k)Q(k)G'(k) . \quad (70)$$

Then the difference matrix $E_{ca}(k+1)$ between $P_a(k+1)$ and $P_c(k+1)$ of Equations (53) and (58) respectively becomes

$$E_{ca}(k+1) = (\Phi(k) - K_c(k)H(k))E_{ca}(k)(\Phi(k) - K_c(k)H(k))' + K_c(k)\Delta R(k)K_c'(k) + G(k)Q(k)G'(k) , \quad (71)$$

where
$$E_{ca}(k) = P_c(k) - P_a(k) . \quad (72)$$

Following the same discussions used in the continuous case as well as the induction, the following theorem can be derived for discrete systems.

Theorem 4

$E_{ca}(k) \geq 0$; hence $P_c(k) \geq P_a(k)$ for $k \geq 0$ if the condition C-III is satisfied.

C-III: $E_{ca}(0) \geq 0$, $\Delta Q(k) \geq 0$ and $\Delta R(k) \geq 0$ or equivalently $P_c(0) \geq P_a(0)$, $Q_c(k) \geq Q(k)$ and $R_c(k) \geq R(k)$ for $k \geq 0$.

Also the counterpart of Corollary 1 is derived, which yields the lower bound of $P_a(k)$.

Corollary 2

$E_{ca}(k) \leq 0$; hence $P_c(k) \leq P_a(k)$ if the condition C-III is satisfied.

In the case of the other two differences,

$$E_{a0}(k) = P_a(k) - P(k) \quad (73)$$

$$E_{c0}(k) = P_c(k) - P(k) , \quad (74)$$

results similar to the continuous case can be proved.

First Equations (60) and (70) are substituted into Equation (73) and after certain manipulation of matrices the following matrix form can be derived.

$$E_{a0}(k+1) = (\Phi(k) - K_c(k)H(k))E_{c0}(k)(\Phi(k) - K_c(k)H(k))' + (K_c(k) - K(k))S(k)(K_c(k) - K(k))' , \quad (75)$$

where
$$S(k) = H(k)P(k)H'(k) + R(k) \quad (76)$$

$$S_c(k) = H_c(k)P_c(k)H_c'(k) + R_c(k) . \quad (77)$$

In this derivation the following relation is useful:

$$K_c(k) = K(k) + (\Phi(k) - K(k)H(k))E_{c0}(k)H'(k)S_c^{-1}(k) - K(k)\Delta R(k)S_c^{-1}(k) . \quad (78)$$

In view of Equation (73), the following theorem is derived.

Theorem 5

$$E_{a0}(k) \geq 0 ; \text{ hence } P_a(k) \geq P(k) \text{ for } k \geq 0 .$$

This is a natural conclusion because $P(k)$ is the optimum covariance by definition.

For the third difference matrix $E_{c0}(k)$ the following relation is obtained:

$$E_{c0}(k+1) = (\Phi(k) - K_c(k)H(k))E_{c0}(k)(\Phi(k) - K_c(k)H(k))' + \\ + (K_c(k) - K(k))S(k)(K_c(k) - K(k))' + K_c(k)\Delta R(k)K_c'(k) + G(k)\Delta Q(k)G'(k) . \quad (79)$$

Then the following theorem is derived.

Theorem 6

$E_{c0}(k) \geq 0$, hence $P_c(k) \geq P(k)$ for $k \geq 0$, if C-III is satisfied.

7. EXAMPLES

Two examples demonstrate the theoretical analysis of this chapter. The first example is concerned with the modeling errors in the *a priori* statistics and the second example with the system modeling errors.

Example 1

A spacecraft is cruising with a constant speed along a straight line and information is supplied by the range data, which are contaminated by white noise having the spectral density Φ_r and zero mean.

Let x_1 and x_2 be deviations of speed and position of the spacecraft from the standard trajectory, respectively. Then the process equation becomes

$$\dot{x}_1(t) = 0 \tag{80}$$

$$\dot{x}_2(t) = x_1(t) . \tag{81}$$

The observation equation is

$$y(t) = x_2(t) + n(t) . \tag{82}$$

Therefore

$$F(t) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \tag{83}$$

$$H(t) = [0, 1] \tag{84}$$

$$Q(t) = 0 \tag{85}$$

$$R(t) = \Phi_r . \tag{86}$$

The *a priori* covariance is chosen as

$$P(0) = \begin{bmatrix} p_{11}(0) & 0 \\ 0 & p_{22}(0) \end{bmatrix} . \tag{87}$$

Then the covariance $P(t)$ of the optimal estimator is derived from Equation (11).

$$P(t) = \frac{1}{Z(t)} \begin{bmatrix} p_{11}(t)(1 + p_{22}(0)t/\Phi_r) , & p_{11}(0)t(1 + p_{22}(0)t/2\Phi_r) \\ p_{11}(0)t(1 + p_{22}(0)t/2\Phi_r) , & p_{22}(0) + p_{11}(0)t^2 + \dfrac{p_{11}(0)p_{22}(0)t^3}{3\Phi_r} \end{bmatrix} \tag{88}$$

where

$$Z(t) = 1 + \left[p_{22}(0) + p_{11}(0)t^2/3 + p_{11}(0)p_{22}(0)t^3/12\Phi_r\right]t/\Phi_r . \tag{89}$$

Suppose that the incorrect model actually used in the design of the suboptimal estimator is given by

$$P_c(0) = \begin{bmatrix} p_{c11}(0) & 0 \\ 0 & p_{c22}(0) \end{bmatrix}$$

$$= \begin{bmatrix} p_{11}(0) + e_{11}(0) & 0 \\ 0 & p_{22}(0) + e_{22}(0) \end{bmatrix} \tag{90}$$

and

$$\Phi_{rc} = \Phi_r + \Delta R . \tag{91}$$

Then the diagonal components of $E_{ca}(t)$ are computed by Equation (36),

$$e_{call}(t) = e^0_{11}(t) + e^R_{11}(t) , \tag{92}$$

where

$$e^0_{11}(t) = \frac{1}{Z_c(t)^2}\left[e_{11}(0)(1 + p_{c22}(0)t/\Phi_{rc})^2 + e_{22}(0)p^2_{c11}(0)t^4/4\Phi^2_{rc}\right] \tag{93}$$

$$e^R_{11}(t) = \frac{p^2_{c11}(0)\Delta R t^3}{12\Phi^4_{rc} Z_c(t)^2}\left[4\Phi^2_{rc} + 2p_{c22}(0)\Phi_{rc}t + p^2_{c22}(0)t^2\right] \tag{94}$$

and

$$Z_c(t) = 1 + \left[p_{c22}(0) + p_{c11}(0)t^2/3 + p_{c11}(0)p_{c22}(0)t^3/12\Phi_{rc}\right]t/\Phi_{rc} . \tag{95}$$

Also

$$e_{cm22}(t) = e^0_{22}(t) + e^R_{22}(t) , \tag{96}$$

where

$$e^0_{22}(t) = \frac{1}{Z_c(t)^2}\left[e_{11}(0)(1 + p_{c22}(0)t/2\Phi_{rc})^2 t^2 + e_{22}(0)(1 - p_{c11}(0)t^3/6\Phi_{rc})^2\right] \tag{97}$$

$$e^R_{22}(t) = \frac{\Delta R t}{\Phi^2_{rc} Z_c(t)^2}\left[(p_{c22}(0) + p_{c11}(0)t^2/2 + p_{c11}(0)p_{c22}(0)t^3/12\Phi_{rc})^2 + \frac{p_{c11}(0)^2 t^4}{12}\left(1 + \frac{p_{c22}(0)t}{2\Phi_{rc}}\right)^2\right] . \tag{98}$$

In Figure 1 the optimal variance $p_{22}(t)$ and actual variance $p_{c22}(t)$ of position of the spacecraft are depicted using $p_{c11}(0)$ as a parameter. The computation of the optimal variance $p_{22}(t)$ is based on the true model.

The suboptimal filter is designed in such a way that

$$p_{c22}(0) > p_{22}(0)$$

$$\Phi_{rc} > \Phi_r$$

and using $p_{c11}(0)$ as a parameter.

The figure indicates that the variance of the suboptimal filter is quite sensitive to variation of $p_{c11}(0)$, i.e., the incorrect initial speed variance. Case (a) is when an excessively large *a priori* uncertainty of speed was employed, i.e., $p_{c11}(0) = 10p_{11}(0)$. On the other hand, case (e) is when the *a priori* value taken was smaller than the true value, namely, $p_{c11}(0) = \frac{1}{10}p_{11}(0)$. For both cases significant overshoots of the variance are observed. This is because the gain $K_c(t)$ was ill-conditioned for both extreme cases. In other words, sufficient weight had not been assigned to the information during the initial period so that the station did not track the spacecraft properly. Case (e) especially demonstrates how the estimator can behave poorly when an optimistic selection is made on the *a priori* covariance.

In Figure 2 the calculated variances $p_{c11}(t)$ used in this case are plotted for the same parameters.

Figure 3 is the case when the variance of initial position $p_{c22}(0)$ is changed as a parameter. It can be observed that the suboptimal filter is not so sensitive to the initial uncertainty of position as to that of speed. However, case (d) reveals a degraded performance of the filter when a smaller value is picked up for the positional uncertainty than the true value.

$$p_{c22}(0) = \tfrac{1}{10}p_{22}(0) .$$

In Figure 4 the incorrect information Φ_{rc} of the power spectral density of the observation noise is employed as a parameter.

The suboptimal filter behaves very poorly for Φ_{rc}, which is either very large (case (a); $\Phi_{rc} = 10\ \Phi_r$) or very small (case (d); $\Phi_{rc} = \frac{1}{10}\Phi_r$) compared to the true Φ_r.

Figure 5 is one example of variances of speed of the spacecraft, corresponding to case (e) of Figure 3.

Example 2

An exponentially correlated signal is processed by a sequential continuous detector which is contaminated by additive white noise. Let x be the signal, then

$$\dot{x}(t) = -\beta x(t) + w \tag{99}$$

$$y(t) = x(t) + n , \tag{100}$$

where w and n are uncorrelated white noise with zero mean, having power spectral density q and r respectively.

It is assumed that the variance σ^2_x of signal is known, but its correlation time $\tau(= 1/\beta)$ is not exactly known *a priori*. Hence an approximate value $\tau_c(= 1/\beta_c)$ is employed in designing the filter.

Thus

$$\beta_c = \beta + \Delta\beta . \tag{101}$$

Also it is assumed that a sufficient time is assigned for tracking compared with the correlation time of the signal, so that only steady-state solutions of Riccati equations which are the limiting values as $t \to \infty$ are investigated.

The purpose of the subsequent analysis is to find the effect on the filter performance of deviation of the assigned correlation time from its true value. The Riccati equation of the optimal variance $p(t)$ associated with the estimator $x^*(t)$ becomes

$$\dot{p}(t) = -2\beta p(t) - \frac{p^2(t)}{r} + q . \tag{102}$$

At the steady state $\dot{p} \to 0$; hence the steady state solution p_s is computed

$$p_s = -\beta r + (\beta^2 r^2 + rq)^{1/2} \tag{103}$$

where

$$q = 2\beta\sigma_x^2 . \tag{104}$$

On the other hand, the actual filter design is based on incorrect knowledge of the correlation time (or β_c). Thus the steady-state value p_{cs} of the calculated variance is given by

$$p_{cs} = -\beta_s r + (\beta_s^2 r^2 + r q_s)^{1/2} \tag{105}$$

$$q_s = 2\beta_s \sigma_x^2 . \tag{106}$$

The actual variance $p_a(t)$ associated with the suboptimal estimator $x_a^*(t)$ which is designed using this $p_c(t)$ is determined from the following three differential equations:

$$\dot{p}_a(t) = -2[\beta_c + p_c(t)/r]p_a(t) - 2\Delta\beta\lambda(t) + p_c^2(t)/r + q \tag{107}$$

$$\dot{\lambda}(t) = -[\beta + \beta_c + p_c(t)/r]\lambda(t) - \Delta\beta p_x(t) - q \tag{108}$$

$$\dot{p}_x(t) = -2\beta p_x(t) + q . \tag{109}$$

The steady-state solutions of these equations are computed and, after some manipulation, the steady state value p_{as} is found as

$$p_{as} = p_{cs} - \frac{r p_{cs} \sigma_x^2 \Delta\beta}{(p_{cs} + r\beta_c)[p_{cs} + r(\beta + \beta_c)]} . \tag{110}$$

This p_{as} becomes a minimum at $\Delta\beta = 0$ and it is equal to the optimal value p_s . It is noticed, however, that as the assigned correlation time becomes shorter p_{as} approaches σ_x^2 .

$$\beta_c \to \infty , \quad p_{cs} \to \sigma_x^2 , \quad p_{as} \to \sigma_x^2 .$$

Also as the assigned correlation time becomes longer, again p_{as} approaches σ_x^2 .

$$\beta_c \to 0 , \quad p_{cs} \to 0 , \quad p_{as} \to \sigma_x^2 .$$

In both extreme cases, the filter is so ill-conditioned that the estimation of the signal is not performed at all. Hence any deviation of the assigned correlation time from its true value in either direction will result in an increase of the actual variance. In Figure 6, p_{as} is plotted against $\Delta\beta/\beta$ for unit value of r and σ_x^2 respectively.

The ratio between the maximum deviation of p_{as} and the optimal value p_s is computed as

$$\alpha = \frac{\sigma_x^2 - p_s}{p_s} = \frac{1}{2}\left(-1 + \left[1 + 2\left(\frac{\sigma_x^2}{\beta r}\right)\right]^{1/2}\right). \tag{111}$$

This α provides the range of the probable deviations of the suboptimal filter performance from that of the optimal one.

8. CONCLUSION

The algorithms for evaluating the effect of errors due to modeling errors in the Kalman filter have been presented in this chapter for both continuous and discrete systems

Also, the error bound of the Kalman filter has been studied when the incorrect *a priori* statistics on initial conditions as well as on noise models are employed. The conservative design criterion expressed in Theorems 1 and 4 guarantees that the suboptimal filter satisfying it remains within the specified range over the estimation period. Also, the formulae of Equations (28) - (31) for continuous systems and of Equations (62) - (64) for discrete systems supply the necessary information to evaluate the effect of errors qualitatively for parametric studies. Such parametric investigations are very important to find out to what extent conservative assignment of *a priori* statistics and noise models can be made, because large covariances of initial coordinates and noises tend to increase the covariance of estimates and eventually to slow down its convergence, thus deteriorating the sensitivity of the filter.

The first example of this chapter well demonstrates the importance of preflight parametric studies when estimations are to be carried out in rather short periods. An optimistic selection of the *a priori* statistics (smaller values of $P_c(0)$, Q_c, and R_c than true values) is especially dangerous because it prevents the estimator from having a proper gain $K(t)$ during the initial period of estimation (case (e) of Figure 1, case (d) of Figure 3 and case (d) of Figure 4). On the other hand, it has been observed in examples that an excessively conservative choice may also be harmful, because it frequently results in a large offset of suboptimal covariances from the optimal ones at the end of the estimation period (case (a) of Figures 1, 3, and 4).

The second example is intended to study the influence of correlation time on the suboptimal filter performance. This is, again, an important problem in space missions because it is often difficult to obtain the exact values of correlation time of stochastic variables such as fluctuations of solar pressure or of low-thrust-engine power. Therefore it is essential to carry out a sensitivity study of the filter as in this example.

ACKNOWLEDGMENT

This chapter represents one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under NASA Contract NAS 7-100.

REFERENCES

1. Kalman, R.E. — *A New Approach to Linear Filtering and Prediction Theory.* Journal of Basic Engineering, Transactions of the American Society of Mechanical Engineers, Vol. 82D, March 1960, pp. 35-45.

2. Kalman, R.E.; Bucy, R.S. — *New Results in Linear Filtering and Prediction Theory.* Journal of Basic Engineering, Transactions of the American Society of Mechanical Engineers, Vol. 83D, March 1961, pp. 95-108.

3. Soong, T.E. — *On a priori Statistics in Minimum Variance Estimation Problems.* Journal of Basic Engineering, Transactions of the American Society of Mechanical Engineers, Vol. 87D, March 1965, pp. 109-112.

4. Nishimura, T. — *On the a priori Information in Sequential Estimation Problems.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol. AC-11, April 1966, pp. 197-204.

5. — *Correction to and Extension of "On the a priori Information in Sequential Estimation Problems".* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control (Correspondence), Vol. AC-12, February 1967, p. 123.

6. Heffes, J. — *The Effect of Erroneous Models on the Kalman Filter Response.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol. AC-11, July 1966, pp. 541-543.

7. Nishimura, T. — *Error Bounds of Continuous Kalman Filters and the Application to Orbit Determination Problems.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol. AC-12, June 1967, pp. 268-275.

8. Griffin, R.E.; Sage, A.P. — *Large and Small Scale Sensitivity Analysis of Optimum Estimation Algorithms.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol. AC-13, August 1968, pp. 320-329.

9. Sawaragi, Y.
Katayama, T.

On the Performance Loss in Sequential Estimation and the Related Problem of Duality. Transactions of the Society of Instrument and Control Engineers of Japan, Vol. 4, No. 3, September 1968, pp. 248-255.

10. Nishimura, T.

Sensitivity Analysis of Noise Correlation Time. Space Program Summary No. 37-57, Vol. II, April 1969, Jet Propulsion Laboratory, California Institute of Technology.

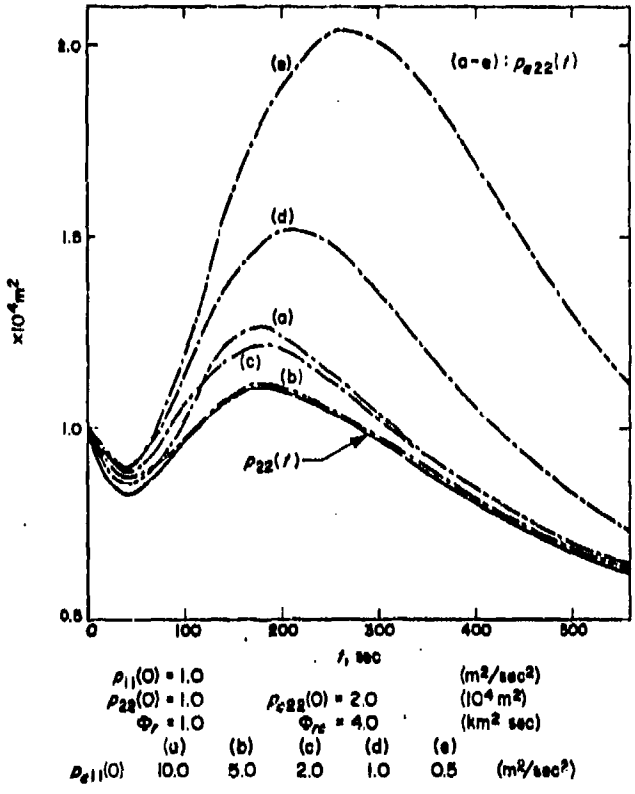


Fig. 1 Actual and optimal variances of position of a spacecraft with range data for various initial suboptimal speed variances

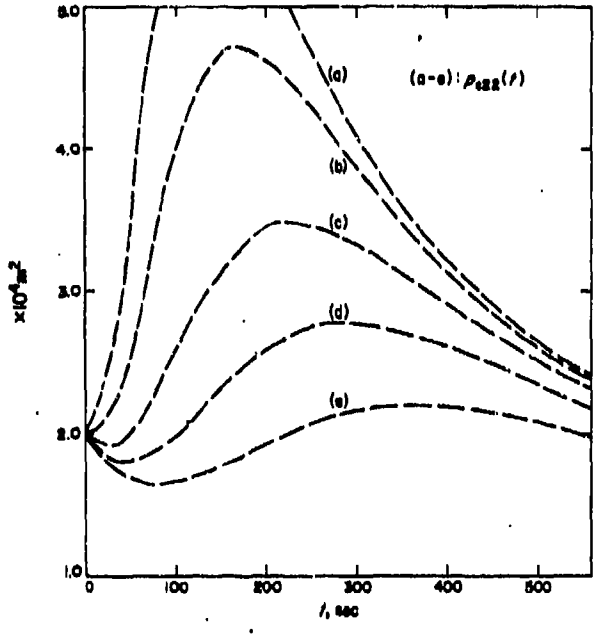
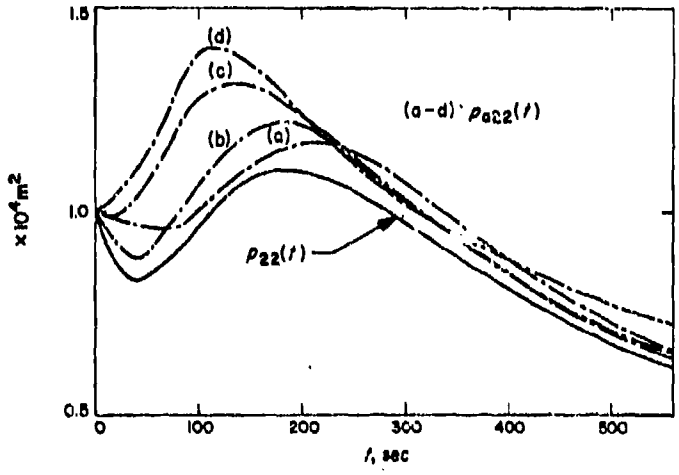
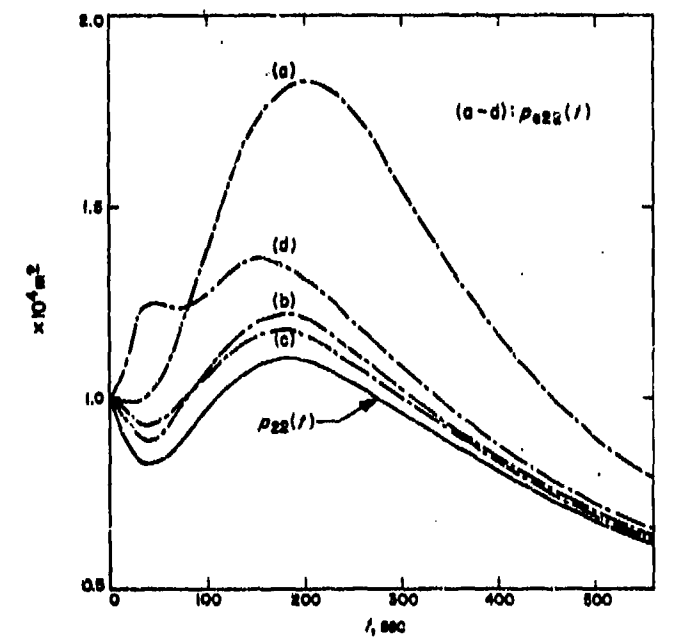


Fig. 2 Computed variance of position of a spacecraft with range data for various initial suboptimal speed variances corresponding to Figure 1



$\rho_{11}(0) = 1.0$		$\rho_{c11}(0) = 2.0$			(m^2/sec^2)
$\rho_{22}(0) = 1.0$					$(10^4 m^2)$
$\Phi_r = 1.0$		$\Phi_{rc} = 4.0$			$(km^2\ sec)$
	(a)	(b)	(c)	(d)	
$\rho_{a22}(0)$	10.0	2.0	0.5	0.1	$(10^4 m^2)$

Fig. 3 Actual and optimal variances of position of a spacecraft with range data for various initial suboptimal position variances



$\rho_{11}(0) = 1.0$		$\rho_{c11}(0) = 2.0$			(m^2/sec^2)
$\rho_{22}(0) = 1.0$		$\rho_{a22}(0) = 2.0$			$(10^4 m^2)$
$\Phi_r = 1.0$					$(km^2\ sec)$
	(a)	(b)	(c)	(d)	
Φ_{rc}	10.0	4.0	1.0	0.5	$(km^2\ sec)$

Fig. 4 Actual and optimal variances of position of a spacecraft with range data for various suboptimal noise spectral densities

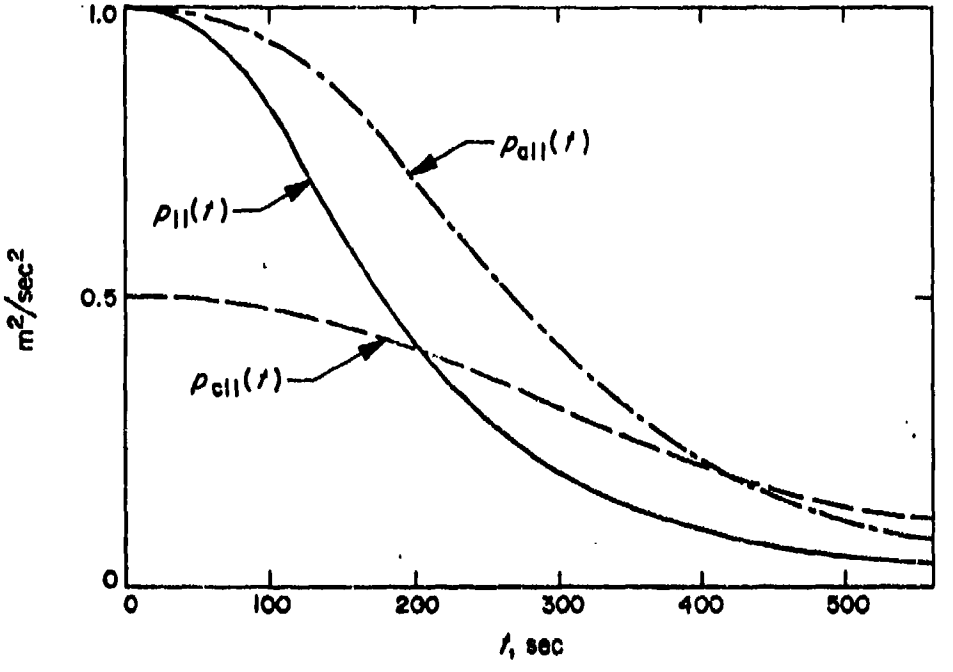


Fig. 5 Computed, actual, and optimal variances of speed of a spacecraft with range data corresponding with case (e) of Figure 1

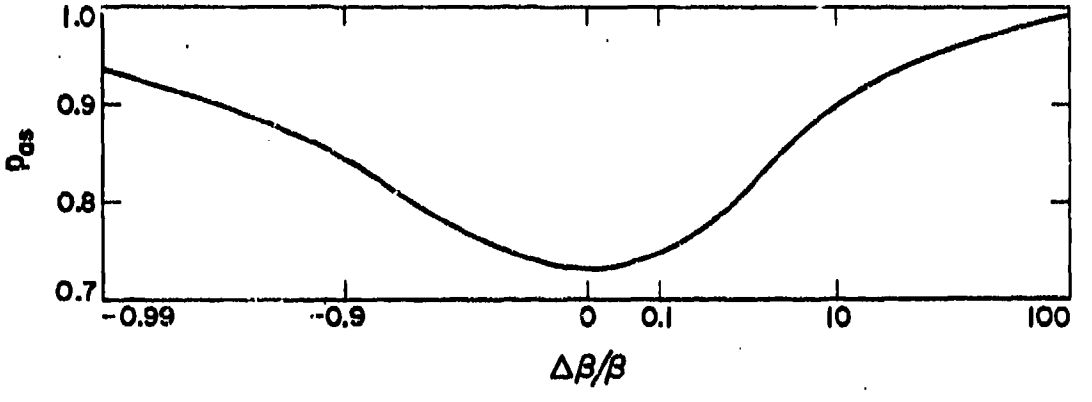


Fig. 6 Actual variance of suboptimal filter

CHAPTER 5 – SUBOPTIMAL KALMAN FILTER TECHNIQUES

by

A. R. Stubberud* and D. A. Wismer†

*School of Engineering
University of California,
Irvine, California, USA

†School of Engineering and Applied Science
University of California, Los Angeles, USA

CHAPTER 5 - SUBOPTIMAL KALMAN FILTER TECHNIQUES

A. R. Stubberud and D. A. Wismer

1. INTRODUCTION

The application of Kalman filter theory[1] requires the definition of a linear mathematical model describing the system for which the application is intended. In many cases a highly complex model must be used to accurately describe the system. Usually in these cases only a few of the state variables are of primary interest, the rest merely enhancing the description of the system. Since the computational burden associated with the implementation of a Kalman filter increases significantly with the dimension of the system model, designers are motivated to seek simplifications to the required filtering equations which do not result in a large performance degradation in the estimation of state vector components of principal interest. The resulting filters are called suboptimal filters.

One successful technique investigated by Joseph[2], Meditch[3] and Pentecost[4] for simplifying filter computations involves partitioning the state vector into strongly coupled subsystems, employing Kalman's optimal filtering algorithm for the state vectors of these lower-dimensional systems and reconstructing an estimate of the original system from the lower-dimensional estimates. Another approach, introduced by Aoki[6] and Huddle[8] and related to observer theory of Luenberger[7], employs the concept of linear aggregation of states to achieve a concise representation of data, thereby reducing the dimension of the estimator. Important work on the effects on the Kalman filter performance of error in the knowledge of the covariance matrices describing the initial conditions of the system state vector, and the white observation and disturbance noise vectors has been contributed by Heffes[9] and Nishimura[13].

The techniques for generating suboptimal filters can be roughly divided into two classes. In the first, the number of equations defining the filter is the same as for the optimal filter, but the number of equations necessary for computing the filter gain is reduced by partitioning the system. In the second, the number of state variables defining the filter is reduced by aggregation. This also reduces the dimensionality of the covariance matrix and hence reduces the computation required to generate the filter gain.

These two classes of suboptimal techniques are discussed in the next two sections.

2. DISCRETE TIME SUBOPTIMAL FILTERS EMPLOYING REDUCED FILTER GAIN COMPUTATION

The computational burden associated with a minimum variance, unbiased linear filter is strongly dependent on the number of simultaneous difference equations which define the filter and the filter gain. If the state of an n-dimensional system is being estimated, a total of n difference equations define the filter and a total of $n(n+1)/2$ difference equations must be solved to determine the optimal filter gain. In this section a particular class of suboptimal filters is discussed. The property of these filters is that the order of the equations defining the filter is the same as the order of the optimal filter but the number of equations used to calculate the filter gain is less than the number used in the optimal filter. Hopefully, the total number of equations can be reduced considerably without the performance of the filter being significantly different from that of the optimal filter.

2.1 Problem Formulation

Consider a dynamic system defined by the linear stochastic difference equation

$$x_{k+1} = \Phi_k x_k + u_k , \qquad k = 0, 1, 2, \ldots, \tag{2.1}$$

where x_k is the n-dimensional state vector

u_k is an n-dimensional zero mean white noise sequence with covariance matrix $Q_k \delta_{jk}$

Φ_k is the $n \times n$ dimensional state transition matrix

x_0, the initial condition, is a zero mean random variable, independent of u_k, with covariance matrix P^*.

The system is assumed to have the observable output

$$y_k = M_k x_k + v_k . \tag{2.2}$$

where y_k is the m-dimensional observation vector

v_k is an m-dimensional zero mean white noise sequence, independent of both u_k and x_0, with covariance matrix $R_k\delta_{jk}$

M_k is an $m \times n$ matrix defining m linearly independent observable combinations of the state variables.

The problem is to generate a linear estimator for the state x_k which makes use of the observations y_j, $j = 0, 1, \ldots k$, to generate the estimate.

Consider the class of linear unbiased filters defined by the stochastic difference equation

$$\hat{x}_{k+1} = \tilde{x}_{k+1} + K_{k+1}[y_{k+1} - M_{k+1}\tilde{x}_{k+1}] , \tag{2.3}$$

where $\tilde{x}_{k+1}$ is defined by

$$\tilde{x}_{k+1} = \phi_k \hat{x}_k , \tag{2.4}$$

with initial condition $\hat{x}_0 = 0$. Here $\hat{x}_{k+1}$ is an unbiased estimate of x_{k+1} and K_{k+1} is an arbitrary $n \times m$ gain matrix. The estimation error associated with this estimator is defined as

$$e_k = x_k - \hat{x}_k \tag{2.5}$$

and the covariance matrix of this error is defined by

$$\bar{P}_k = E[\hat{e}_k \hat{e}_k^T] . \tag{2.6}$$

This matrix (specifically the elements on the main diagonal) is a measure of the effectiveness of the filter and it satisfies the matrix equations

$$P_{k+1} = \phi_k \bar{P}_k \phi_k^T + Q_k \tag{2.7}$$

$$\bar{P}_k = [I - K_k M_k] P_k [I - K_k M_k]^T + K_k R_k K_k^T , \tag{2.8}$$

with the initial condition $\bar{P}_0 = P^*$, where

$$P_k = E[(x_k - \tilde{x}_k)(x_k - \tilde{x}_k)^T] = E[\tilde{e}_k \tilde{e}_k^T] . \tag{2.9}$$

Note that, since $\tilde{x}_k$ is the best estimate of x_k based on $y_0, y_1 \ldots y_{k-1}$, it is generally a poorer estimate than $\hat{x}_k$, which includes y_k also. $\bar{P}_k$ is thus the true measure of the filter performance and P_k is simply an intermediate matrix which it is convenient to carry along. Equations (2.3) and (2.4) which define the class of unbiased linear filters are valid for any gain K_{k+1}, and Equations (2.7) and (2.8) define the variances of the errors for any filter having this form. Thus the effectiveness of any suboptimal gain can be evaluated from these general variance equations.

A special member of this class of filters results when the gain is chosen as

$$K_{k+1} = P_{k+1} M_{k+1}^T [M_{k+1} P_{k+1} M_{k+1}^T + R_{k+1}]^{-1} . \tag{2.10}$$

The filter is then a minimum variance unbiased filter. In this case Equation (2.8) reduces to

$$\bar{P}_k = [I - K_k M_k] P_k . \tag{2.8a}$$

The computational burden of this optimal filter includes the solution of the n difference equations defined by Equations (2.3) and (2.4) plus the solution of the $n(n+1)/2$ difference equations defined by Equations (2.7) and (2.8) to calculate the optimal gain K_k. (The symmetry of the n^2 difference equations requires that only $n(n+1)/2$ need to be solved.)

In the following sub-sections some methods for choosing suboptimal gains K_{k+1} are discussed.

2.2 Constant Gains

In terms of a reduction in computation, an obvious choice for K_k is a constant; that is,

$$K_k = K \quad \text{for all } k .$$

Probably, the most common choice of a constant gain is the steady-state value (if one exists) of the optimal gain matrix. Consider the case of a time-invariant system, that is, the case where

$$\phi_k = \phi , \quad M_k = M , \quad Q_k = Q , \quad R_k = R \quad \text{for all } k .$$

The steady-state gain for this system (if it exists) is determined from Equations (2.7), (2.8), and (2.10) by setting

$$P_k = P , \quad \bar{P}_k = \bar{P} , \quad K_k = K \quad \text{for all } k .$$

The result is a set of algebraic equations

$$P = \phi \bar{P} \phi^T + Q \tag{2.11}$$

$$\bar{P} = [I - KM] P [I - KM]^T + KRK^T \tag{2.12}$$

$$K = PM^T [MPM^T + R]^{-1} . \tag{2.13}$$

The error variance equation for this case becomes

$$P_{k+1} = [I - KM]\phi P_k \phi^T [I - KM]^T + [I - KM]Q[I - KM]^T + KRK^T . \tag{2.14}$$

It should be noted that the filter generated in this way is the classical Wiener filter for discrete time systems, which is suboptimal until steady-state conditions are reached.

2.3 A Suboptimal Technique Based On Partitioning

In this section a method is demonstrated which is based on partitioning of the system state vector into several sub-state vectors. The states of the subsystems thus generated are estimated individually and the resulting estimates are combined linearly to generate the total estimate. Since the number of variance equations varies approximately as the square of the dimension of the system state the number of variance equations used to compute the optimal gains is significantly reduced.

Consider a dynamic system defined by Equation (2.1). Joseph[7] described a general method for partitioning the state vector of the system into a set of i subsystem state vectors. The i subsystem state vectors are defined by

$$\xi_k^j = D_j x_k , \qquad j = 1, 2, \ldots, i , \tag{2.15}$$

where D_j is an $n_j \times n$ matrix of rank n_j. The state vector x_k is recovered from the subsystem state vectors by the relationship

$$x_k = \sum_{j=1}^{i} A_j \xi_k^j = \sum_{j=1}^{i} A_j D_j x_k , \tag{2.16}$$

where A_j is an $n \times n_j$ matrix.

Suppose the state vector x_k has been partitioned into i sub-state vectors

$\xi_k^1 , \quad \xi_k^2 , \ldots, \xi_k^i$ with respective dimensions

$n_1 , \quad n_2 , \ldots, n_i$ and suppose that individual estimates

$\hat{\xi}_k^1 , \quad \hat{\xi}_k^2 , \ldots, \hat{\xi}_k^i$ have been generated from the observations

$y_l , \; l = 0, 1, \ldots, k.$

Then the estimate of the state vector is given by

$$\hat{x}_k = \sum_{j=1}^{i} A_j \hat{\xi}_k^j . \tag{2.17}$$

This estimate is then propagated by the state transition matrix, thus forming

$$\bar{x}_{k+1} = \phi_k \hat{x}_k . \tag{2.18}$$

The propagated sub-state estimate of ξ_{k+1}^j is then formed by

$$\bar{\xi}_{k+1}^j = D_j \bar{x}_{k+1} . \tag{2.19}$$

Consider now a partitioning of the observation vector defined in Equation (2.2). The vector y_k is partitioned into i observation vectors

$$\eta_k^j = E_j y_k , \qquad j = 1, 2, \ldots, i , \tag{2.20}$$

where E_j is an $m_j \times n$ matrix. The estimate of ξ_{k+1}^j is now updated by the equation

$$\hat{\xi}_{k+1}^j = \bar{\xi}_{k+1}^j + K_{k+1}^j [\eta_{k+1}^j - \bar{\eta}_{k+1}^j] , \tag{2.21}$$

where

$$\begin{aligned} \eta_{k+1}^j - \bar{\eta}_{k+1}^j &= E_j y_{k+1} - E_j \bar{y}_{k+1} \\ &= E_j y_{k+1} - E_j M_{k+1} \bar{x}_{k+1} \\ &= E_j [y_{k+1} - M_{k+1} \bar{x}_{k+1}] . \end{aligned} \tag{2.22}$$

The updated state estimate is then given by

$$\hat{x}_{k+1} = \sum_{j=1}^{L} A_j \hat{\xi}_{k+1}$$

$$= \sum_{j=1}^{L} A_j \tilde{\xi}^j_{k+1} + \sum_{j=1}^{L} A_j K_j [\eta^j_{k+1} - \bar{H}^j_{k+1}]$$

$$= \bar{x}_{k+1} + \sum_{j=1}^{L} A_j K^j_{k+1} E_j [y_{k+1} - H_{k+1} \bar{x}_{k+1}] . \tag{2.23}$$

Comparing Equations (2.3) and (2.23), it is apparent that this represents a member of the class of unbiased filters defined by Equation (2.3). Here the gain is

$$K_{k+1} = \sum_{j=1}^{L} A_j K^j_{k+1} E_j . \tag{2.24}$$

Note that D_j, A_j, and E_j can all be made a function of the time index k so that

$$K_{k+1} = \sum_{j=1}^{L} A^j_{k+1} K^j_{k+1} E^j_{k+1} . \tag{2.24a}$$

This generalization allows the subsystems to be changed if the system has several modes of operation which make it mandatory, convenient, or more effective to change the system partitions for the several modes of operation.

Equation (2.23) indicates that the individual estimates of the subsystem states need not be generated. It is seen that, for given A_j and E_j, only the K^j_{k+1} need to be generated and then K_{k+1} can be generated by Equation (2.24). This K_{k+1} can be used directly in Equation (2.23) to generate the estimate $\hat{x}_{k+1}$.

The next step in this suboptimal method is the generation of the set of equations used to compute the gains K^j_k. First, an error equation is generated. Subtracting Equation (2.4) from Equation (2.1) generates the error equation

$$\tilde{e}_{k+1} = \phi_k \hat{e}_k + u_k . \tag{2.25}$$

It follows immediately that

$$\tilde{e}^j_{k+1} = \xi^j_{k+1} - \tilde{\xi}^j_{k+1} = D_j x_{k+1} - D_j \bar{x}_{k+1}$$

$$= D_j \tilde{e}_{k+1} = D_j \phi_k \hat{e}_k + D_j u_k . \tag{2.26}$$

Now

$$\hat{e}^j_k = D_j \hat{e}_k . \tag{2.27}$$

and therefore

$$\hat{e}_k = D^\dagger_j \hat{e}^j_k + z , \tag{2.28}$$

where z is a vector in the null space of D_j and $D^\dagger_j$ is the pseudo-inverse of D_j.

If D_j has maximal rank, then

$$D^\dagger_j = D^T_j (D_j D^T_j)^{-1} \tag{2.29}$$

is a suitable pseudo-inverse.

Substituting Equation (2.28) into Equation (2.26) yields

$$\tilde{e}^j_{k+1} = D_j \phi_k D^\dagger_j \hat{e}^j_k + D_j \phi_k z + D_j u_k$$

$$= \phi^j_k \hat{e}^j_k + D_j \phi_k z + D_j u_k , \tag{2.30}$$

where

$$\phi^j_k = D_j \phi_k D^\dagger_j . \tag{2.31}$$

At this point an approximation is made. It is assumed that the subsystems are uncoupled. This allows the gains for each subsystem to be calculated individually. The absence of coupling between the subsystems implies that $z = 0$. In general this is not true and the resulting gains are suboptimal. Under this assumption Equation (2.30) becomes

$$\epsilon^j_{k+1} = \phi^j_k \epsilon^j_k + D_j u_k . \tag{2.32}$$

A variance equation is now generated from Equation (2.32) as

$$E[\tilde{e}^j_{k+1} (\tilde{e}^j_{k+1})^T] = \phi^j_k E[\tilde{e}^j_k (\hat{e}^j_k)^T] (\phi^j_k)^T + D_j E[u_k u_k^T] D^T_j . \tag{2.33}$$

Now define

$$P_k^j = E[\hat{z}_k^j(\hat{z}_k^j)^T] \tag{2.34}$$

$$P_k^j = E[\bar{z}_k^j(\hat{z}_k^j)^T] \tag{2.35}$$

and Equation (2.33) becomes

$$P_{k+1}^j = \phi_k^j P_k^j(\phi_k^j)^T + Q_k^j . \tag{2.36}$$

where

$$Q_k^j = D_j Q_k D_j^T . \tag{2.37}$$

Now consider the error equation

$$e_{k+1}^j = \hat{z}_{k+1}^j - \bar{z}_{k+1}^j . \tag{2.38}$$

Substituting from Equation (2.21)

$$\begin{aligned} e_{k+1}^j &= \hat{z}_{k+1}^j - \bar{z}_{k+1}^j - K_{k+1}^j[\eta_{k+1}^j - \bar{\eta}_{k+1}^j] \\ &= e_{k+1}^j - K_{k+1}^j[E_j y_{k+1} - E_j \bar{y}_{k+1}] \\ &= e_{k+1}^j - K_{k+1}^j[E_j M_{k+1}x_{k+1} + E_j v_{k+1} - E_j M_{k+1}\bar{x}_{k+1}] \\ &= e_{k+1}^j - K_{k+1}^j[E_j M_{k+1}e_{k+1} + E_j v_{k+1}] . \end{aligned} \tag{2.39}$$

Since

$$\bar{z}_{k+1}^j = D_j \bar{z}_{k+1} , \tag{2.40}$$

then

$$\bar{z}_{k+1} = D_j^T \bar{z}_{k+1}^j + z_1 , \tag{2.41}$$

where z_1 is in the null space of D_j. Again the coupling is ignored by assuming $z_1 = 0$. Equation (2.39) becomes

$$\hat{e}_{k+1}^j = [I - K_{k+1}^j E_j M_{k+1} D_j^T]\bar{e}_{k+1}^j - K_{k+1}^j E_j v_{k+1} . \tag{2.42}$$

The corresponding variance equation is given by

$$\begin{aligned} E[\hat{e}_{k+1}^j(\hat{e}_{k+1}^j)^T] &= P_{k+1}^j \\ &= [I - K_{k+1}^j E_j M_{k+1} D_j^T]P_{k+1}^j[I - K_{k+1}^j E_j M_{k+1} D_j^T]^T + K_{k+1}^j E_j R_{k+1} E_j^T (K_{k+1}^j)^T , \end{aligned} \tag{2.43}$$

with initial condition $\bar{P}_0^j = D_j P_0 D_j^T$ when $P_0 = E[x_0 x_0^T]$. The gain which minimizes P_{k+1}^j is then given by

$$K_{k+1}^j = P_{k+1}^j(M_{k+1}^j)^T[M_{k+1}^j P_{k+1}^j(M_{k+1}^j)^T + R_{k+1}^j]^{-1} , \tag{2.44}$$

where

$$M_{k+1}^j = E_j M_{k+1} D_j^T \tag{2.45}$$

$$R_{k+1}^j = E_j R_{k+1} E_j^T . \tag{2.46}$$

The basic equations for generating K_{k+1}^j are Equations (2.36), (2.43) and (2.44). Having calculated K_{k+1}^j, $j = 1, 2, \ldots, i$, the total gain is calculated from Equation (2.24) and the estimate of x_k is given by the difference equation (2.23).

As a simple example which illustrates only the computational advantages of this method, consider the case of a system state with nine state variables ($n = 9$) and suppose it has been partitioned into three subsystems, each with three state variables. The optimal filter requires the solution of a total of $(9)(10)/2 = 45$ simultaneous variance equations. The suboptimal filter requires the solution of $3[(3)(4)/2] = 18$ simultaneous equations.

Obviously there is no unique choice for the partition of a specific system. Pentecost[4] discusses a rationale for partitioning a system.

One general rule for partitioning a system is that each state variable should be assigned, through a choice of the matrices D_j, to a subsystem which assures it is strongly correlated to an observation. Heuristically, using a suboptimal filter is equivalent to making the system less observable. The correlation between a state and the observations is a measure of the observability of the state and should be kept as large as possible when choosing a partition. Several observations will generally be associated with a single subsystem.

A second partitioning rule is that strongly correlated states should be assigned to the same subsystem since, in a general partition, the correlation between states in different subsystems is not taken into account. Any partition should be chosen to minimize this loss of correlation. One measure of the correlations between state variables can be obtained from the error covariance $\tilde{P}_k$ for the optimal filter. The correlation matrix C_k with elements

$$c_{ij}^k = \frac{p_{ij}^k}{\sqrt{p_{ii}^k p_{jj}^k}} , \tag{2.47}$$

where p_{ij}^k is an element of P_k. The terms c_{ij}^k, $i \neq j$, being a measure of state variable correlation, can be used to group state variables into the same subsystem. Note that it is possible that one state variable will belong to more than one subsystem.

In order to properly design a suboptimal filter as described in this section, it is appropriate to perform a full scale simulation of the optimal filter covariance equations and suboptimal filter covariance equations, iterating on the choice of subsystems should lead to a suitable filter design. In most aerospace applications extensive simulations are usually the rule; therefore this technique is well suited to these applications.

3. SUBOPTIMAL FILTERS EMPLOYING STATE REDUCTION

An alternate approach to that cited in Section 2 assumes that unavoidable errors will occur in the formulation of a mathematical model of the system. With this premise, a primary system model is formulated including only those state vector components of dominant special interest to the designer. The other components and their interactions are considered as part of a secondary subsystem. When this is done, the state vector of interest has reduced dimension and allows a reduction in the computational burden required for implementing a Kalman filter based on the primary system model. In many practical problems[9-11] the structure of the plant is such that the state vector components which are ignored by such a procedure are independent sources of correlated disturbance or additive observation noise. Thus, the secondary system may be considered to be a model of colored noise in addition to the usual white noise sources assumed in the primary subsystem. This section considers the performance degradation incurred by basing the filter design on the primary system only. The matrix differential equation governing the error in the covariance matrix of the primary system is expressed explicitly in terms of the neglected state components, i.e. the secondary system. The expression for performance loss incurred with any proposed filter allows the designer to determine directly the utility of simple models in filter synthesis. Error models of this type have been discussed extensively by Larson[12].

3.1 Problem Formulation

Consider a dynamic system which can be represented by two linear stochastic difference equations of the form

$$x_{k+1}^1 = \Phi_k x_k^1 + C_k x_k^2 + u_k^1 \tag{3.1}$$

$$x_{k+1}^2 = \Lambda_k x_k^2 + u_k^2 , \tag{3.2}$$

where

- x_k^1 is an n_1-dimensional state vector composed of those components of a refined linear mathematical model which are dominant or of special interest (the state of the primary subsystem).
- x_k^2 is an n_2-dimensional vector composed of those system components which appear as time-correlated disturbances or additive observation noise in a refined model (the state of the secondary subsystem)
- u_k^1 is an n_1-dimensional zero-mean white noise sequence with covariance matrix $Q_k^1 \delta_{jk}$
- u_k^2 is an n_2-dimensional zero-mean white noise sequence with covariance matrix $Q_k^2 \delta_{jk}$
- Φ_k is the state transition matrix for the primary subsystem, an $n_1 \times n_1$ dimension matrix
- C_k is a matrix which couples the effect of the secondary state vector into the primary system, an $n_1 \times n_2$ dimensional matrix
- Λ_k is the state transition matrix for the secondary subsystem, an $n_2 \times n_2$ dimensional matrix.

The system is assumed to have the observable outputs

$$y_k = M_k x_k + v_k . \tag{3.3}$$

where

- y_k is an m-dimensional observation vector
- v_k is an m-dimensional zero-mean white noise sequence which is independent of u_k^1 and u_k^2 and has covariance matrix $R_k \delta_{jk}$
- x_k is an $n = (n_1 + n_2)$-dimensional vector composed of x_k^1 and x_k^2
- $M_k = [M_{1k} : M_{2k}]$ is an $m \times n$ dimensional matrix defining m linearly independent observable combinations of system state variables. M_{1k} is an $m \times n_1$ dimensional matrix and M_{2k} is an $m \times n_2$ dimensional matrix.

To reduce computational requirements for implementing an estimate of the primary subsystem state vector x_k^1, the filter design is based on the simplified model

$$X_{k+1} = \Phi_k X_k + U_k \tag{3.4}$$

$$y_k = M_{1k} X_k + v_k . \tag{3.5}$$

where

$\bar{x}_k$ is an n_1-dimensional state vector

$\bar{y}_k$ is an m-dimensional observation vector

$\bar{u}_k$ is an n_1-dimensional zero-mean white noise sequence with covariance $\bar{Q}_k\delta_{jk}$

$\bar{v}_k$ is an m-dimensional zero-mean white noise sequence with covariance matrix $\bar{R}\delta_{jk}$

It is now of interest to determine the error of the estimate of the state x_k^1 where the Kalman filtering algorithm is applied to the simplified model defined by Equations (3.4) and (3.5).

In order to calculate the performance loss from optimal incurred by this simplification, we could compute the optimal estimate and compare it with the estimate based on the simplified model. However the optimal estimate requires the solution of an $n \times n$ dimensional matrix Riccatti equation which we wanted to avoid by using the suboptimal approach. It is possible to determine a portion of the performance degradation with little additional computation. This portion is the effect of the secondary state variables on the estimate when the suboptimal Kalman gain is used. Although this is not the total performance loss due to the suboptimal estimation it does give an indication of this effect.

The linear filter for the entire system is given by

$$\hat{x}_{k+1} = A_k\hat{x}_k + K_k[y_k - M_k\hat{x}_k] , \tag{3.6}$$

where

$$A_k = \left[\begin{array}{c|c} \Phi_k & C_k \\ \hline 0 & \Lambda_k \end{array}\right]$$

$K_k = \begin{bmatrix} K_{1k} \\ \hline K_{2k} \end{bmatrix}$ is an $n \times m$ arbitrary gain matrix, K_{1k}

being $n_1 \times m$ dimensional and K_{2k}

being $n_2 \times m$ dimensional.

If the designer assumes the simplified form of the plant as a basis for synthesizing a filter for the state vector components that are of primary interest, the gain matrix becomes

$$K_k = \begin{bmatrix} K_{1k} \\ \hline 0 \end{bmatrix} .$$

Now defining the error of the estimate as

$$\left.\begin{aligned} \tilde{x}_k^1 &= x_k^1 - \hat{x}_k^1 \\ \text{and} \qquad \tilde{x}_k^2 &= x_k^2 - \hat{x}_k^2 . \end{aligned}\right\} \tag{3.7}$$

we can determine the covariance matrix for the filter. However, since the filter gain is based on the primary system only, $\hat{x}_k^2 = 0$ and $\tilde{x}_k^2 = x_k^2$. Thus, by appropriate substitution into (3.7), we get

$$\tilde{x}_{k+1}^1 = [\Phi_k - K_{1k}M_{1k}]\tilde{x}_k^1 + \bar{C}_k + \tilde{x}_k^2 + u_k^1 - K_{1k}v_k , \tag{3.8}$$

where

$$\bar{C}_k = C_k - K_{1k}M_{2k} .$$

The covariance matrix for the overall system can now be written in terms of the subsystem elements. Defining

$$\rho_{ij}^k = E[(\tilde{x}_k^i)(\tilde{x}_k^j)^T] ,$$

where the dimensions of ρ_{11}^k, ρ_{12}^k, ρ_{22}^k are respectively $n_1 \times n_1$, $n_1 \times n_2$, and $n_2 \times n_1$ and where $\rho_{12}^k = (\rho_{21}^k)^T$ Performing the indicated operations yields

$$\rho_{11}^{k+1} = B_k\rho_{11}^kB_k^T + B_k\rho_{12}^k\bar{C}_k^T + \bar{C}_k\rho_{21}^kB_k + \bar{C}_k\rho_{22}^k\bar{C}_k^T + Q_k^1 + K_{1k}R_kK_{1k}^T \tag{3.9a}$$

$$\rho_{12}^{k+1} = B_k\rho_{12}^k\Lambda_k^T + \bar{C}_k\rho_{22}^k\Lambda_k^T \tag{3.9b}$$

$$\rho_{22}^{k+1} = \Lambda_k\rho_{22}^k\Lambda_k^T + Q_k^2 , \tag{3.9c}$$

where

$$B_k = \Phi_k - K_{1k}M_{1k} .$$

Note that the nature of the coupling between ρ_{22}^{k+1}, ρ_{12}^{k+1}, and ρ_{11}^{k+1} permits the explicit solution for each of these matrices in a serial manner.

In order to evaluate the degradation in performance due to the effect of the secondary subsystem, it is necessary to determine the covariance matrix based on the simplified model. This is easily determined from ρ_{11}^{k+1} by setting $\bar{C}_k$ to zero. Defining

$$\rho^k = E[(\tilde{x}_k)(\tilde{x}_k)] ,$$

we obtain the recurrence relation for the simplified system as

$$\rho^{k+1} = B_k \rho^k B_k^T + \bar{Q}_k + K_{1k} \bar{R}_k K_{1k}^T . \tag{3.10}$$

Now defining the error introduced into the primary subsystem by the secondary subsystem as

$$E_k = \rho_{11}^k - \rho^k ,$$

yields a recurrence calculation for the error matrix; namely

$$E_{k+1} = B_k E_k B_k^T + B_k \rho_{12}^k \bar{C}_k^T + \bar{C}_k \rho_{21}^k B_k^T + \bar{C}_k \rho_{22}^k \bar{C}_k^T + \Delta Q_k + K_{1k} \Delta R_k K_{1k}^T , \tag{3.11}$$

where

$$\Delta Q_k = Q_k^1 - \bar{Q}_k$$

$$\Delta R_k = R_k - \bar{R}_k ,$$

and where the matrices ρ_{12}^k and ρ_{22}^k are obtained recursively as above. The ΔQ_k and ΔR_k matrices represent the error in modeling the white noise sequences for the simplified system. Thus the error matrix E^k can be computed simultaneously with the simplified filter equations in order to indicate the error incurred by omitting the secondary subsystem variables from the primary model.

The above analysis holds for any arbitrary gain matrix K_{1k}. However if the optimal Kalman filter is employed, the gain matrix is chosen as

$$K_{1k} = \rho^k M_{1k}^T [M_{1k} \rho^k M_{1k}^T + \bar{R}_k]^{-1} . \tag{3.12}$$

3.2 Results for the Continuous Case

In the case where the system is described by linear differential equations instead of difference equations similar results are obtained. Thus if the system is given as

$$\left.\begin{aligned} \dot{x}_1 &= A(t)x_1 + C(t)x_2 + u_1 \\ \dot{x}_2 &= \Lambda(t)x_2 + u_2 \end{aligned}\right\} \tag{3.13}$$

and

$$y(t) = M(t)x(t) + v(t) ,$$

where the vectors and matrices are defined to be analogous to the discrete case and if the simplified system is given by

$$\left.\begin{aligned} \dot{\bar{x}}(t) &= A_1(t)\bar{x}(t) + u_1 \\ \bar{y}(t) &= M_1(t)\bar{x}(t) + v , \end{aligned}\right\} \tag{3.14}$$

the corresponding covariance equations are

$$\dot{\rho}_{11} = B\rho_{11} + \rho_{11}B^T + \bar{C}\rho_{21} + \rho_{12}\bar{C}^T + Q_1 + K_1 R K_1^T , \tag{3.15a}$$

$$\dot{\rho}_{12} = B\rho_{12} + \bar{C}\rho_{22} + \rho_{12}\Lambda^T \tag{3.15b}$$

$$\dot{\rho}_{22} = \Lambda\rho_{22} + \rho_{22}\Lambda^T + Q_2 , \tag{3.15c}$$

where

$$B = A_1(t) - K_1(t)M_1(t)$$

$$\bar{C} = C(t) - K_1(t)M_2(t) .$$

Similarly the covariance matrix for the simplified system is given by

$$\dot{\rho} = B\rho + \rho B^T + \bar{Q} + K_1 \bar{R} K_1^T . \tag{3.16}$$

In this case the differential equation describing the degradation in the system performance due to the existence of unmodeled state variables is given by

$$\dot{E} = BE + EB^T + \bar{C}\rho_{21} + \rho_{12}\bar{C}^T + \Delta Q + K_1 \Delta R K_1^T , \tag{3.17}$$

where

$$E(0) = \rho_{12}(0) - \rho(0) ,$$

which equals zero if the initial conditions for these covariance matrices coincide. In the continuous case the optimal estimate based on the simplified model is implemented when the gain matrix is chosen as

$$K_1 = \rho M_1^T \bar{R}^{-1} .$$

3.3 The Steady-State Case

An analytic closed form expression for the steady-state values of performance loss can be obtained directly from the above differential equations. Temporarily assuming that the design corresponding to the simplified model

is stable, three matrix equations are obtained by setting

$$\dot{E} = \dot{\rho}_{12} = \dot{\rho}_{22} = 0 .$$

Employing the well-known solution of these equations[12], the following sequence of expressions permits the determination of the steady-state performance loss E due to the modeling errors.

$$\rho_{22} = \int_0^\infty \exp[\Lambda t] Q_2 \exp[\Lambda^T t]\, dt \tag{3.18a}$$

$$\rho_{12} = \int_0^\infty \exp[Bt] \bar{C} \rho_{22} \exp[\Lambda^T t]\, dt \tag{3.18b}$$

$$E = \int_0^\infty \exp[Bt] [\bar{C} \rho_{12}^T + K_1 \Delta R K_1^T + \Delta Q + \rho_{12} \bar{C}^T] \exp[B^T t]\, dt . \tag{3.18c}$$

3.4 The Case of Random Sampling

In many applications of Kalman filtering, the filter states are updated randomly in time. This situation may be caused by computational allocation in the system computer, subsystem malfunctions, or when measurements are made aperiodically[8-10]. The "optimal" gain for this type of filter is zero except at the update time when it is given by

$$K_1 = \rho M_1^T [M_1 \rho M_1^T + \bar{R}]^{-1} .$$

This is nearly the same optimal gain as the discrete time filter, except that in this case the covariance matrices are propagated by the following differential equations obtained by setting $K_1 = 0$ in (3.15b) (3.15c) and (3.17) for the continuous case. Thus, between the random update times, we have

$$\dot{E} = BE + EB^T + C\rho_{21} + \rho_{12} C^T + \Delta Q \tag{3.19a}$$

$$\dot{\rho}_{12} = B\rho_{12} + \rho_{12}\Lambda^T + C\rho_{22} \tag{3.19b}$$

$$\dot{\rho}_{22} = \Lambda\rho_{22} + \rho_{22}\Lambda^T + Q_2 . \tag{3.19c}$$

At the time an update in the estimated state of the plant occurs, the old value $\hat{x}_1^-$ (*a priori* estimate) is replaced by a new value $\hat{x}_1^+$ (*a posteriori* estimate). The *a posteriori* estimate is obtained as

$$\hat{x}_1^+ = \hat{x}_1^- + K_1[M_1 x_1 + v + M_2 x_2 - M_1 \hat{x}_1^-] . \tag{3.20}$$

Since no estimate of the state vector x_2 is generated,

$$\hat{x}_2^+ = \hat{x}_2^- = 0 .$$

Defining the instantaneous estimation errors as

$$X_1^+ = x_1 - \hat{x}_1^+$$

$$X_1^- = x_1 - \hat{x}_1^-$$

$$X_2 = x_2 ,$$

we obtain

$$X_1^+ = X_1^- + K_1[M_1 X_1^- + M_2 X_2 + v] \tag{3.21}$$

$$X_2^+ = X_2^- . \tag{3.22}$$

Using these expressions we can obtain expressions for updating the covariance matrices corresponding to the *a priori* and *a posteriori* estimates in a straightforward manner. Thus, defining

$$\rho_{ij}^+ = E[(X_i^+)(X_j^+)^T]$$

and

$$\rho_{ij}^- = E[(X_i^-)(X_j^-)^T]$$

yields

$$\rho_{11}^+ = [I - K_1 M_1]\rho_{11}^-[I - K_1 M_1]^T + K_1 R K_1^T + [K_1 M_2]\rho_{21}^-[I - K_1 M_1]^T + [I - K_1 M_1]\rho_{12}^-[K_1 M_2]^T + [K_1 M_2]\rho_{22}^-[K_1 M_2]^T \tag{3.23a}$$

$$\rho_{12}^+ = [I - K_1 M_1]\rho_{12}^- + K_1 M_2 \rho_{22}^- \tag{3.23b}$$

$$\rho_{22}^+ = \rho_{22}^- . \tag{3.23c}$$

where I is the identity matrix of dimension n_1.

The corresponding equation for updating the covariance matrix of the simplified system is

$$\rho^+ = [I - K_1 M_2]\rho^-[I - K_1 M_1]^T + K_1 R K_1^T . \tag{3.24}$$

Finally the equation for updating the performance loss defined as

$$E = \rho_{11} - \rho$$

is obtained from (3.23a) and (3.24) as

$$E^+ = (I - K_1M_2)E^-(I - K_1M_1)^T + K_1\Delta RK_1^T + [K_1M_2]\rho_{21}^-[I - K_1M_1]^T + [I - K_1M_1]\rho_{12}^-[K_1M_2]^T + [K_1M_2]\rho_{22}^-[K_1M_2]^T . \quad (3.25)$$

4. CONCLUSIONS

In this chapter the suboptimal filtering problem is discussed. Roughly, the suboptimal filtering techniques can be broken into two general categories. The first category has the property that the order of the suboptimal filter is the same as that of the optimal filter, but the number of equations necessary to generate the filter gain is reduced. One way of performing this reduction is by partitioning the system and ignoring any correlation which exists between the various subsystems thus formed. Generally an extensive simulation is necessary to minimize the degradation in filter performance caused by the suboptimality. The savings in computation can, however, be very substantial if the system can be finely partitioned.

The second category of suboptimal filters has the property that the filter design is based on a lower dimensional state vector than the higher dimensional model which defines the system state. As a result the number of covariance equations is also reduced. An estimate of the most important elements of the higher dimensional model can be obtained along with a differential equation describing the performance loss incurred by the model simplification. The filtering and performance loss equations can be solved simultaneously to obtain an indication of actual performance.

REFERENCES

1. Kalman, R.E. — *A New Approach to Linear Filtering and Prediction Problems.* Transactions, American Society of Mechanical Engineers, Journal of Basic Engineering, Vol.82, 1960, pp.35-45.

2. Joseph, P.D. — *Suboptimal Linear Filtering.* Space Technology Laboratory Inc., Technical Report, December 1963.

3. Meditch, J.S. — *Suboptimal Linear Filtering for Continuous Dynamic Process.* Aerospace Corporation, Technical Report, July 1964.

4. Pentecost, E.E. — *Synthesis of Computationally Effective Sequential Linear Estimation* PhD dissertation, Department of Engineering, University of California, Los Angeles, May 1965.

5. Heffes, H. — *The Effect of Erroneous Models on the Kalman Filter Response.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control (Short Papers), Vol.AC-11, July 1966, pp.541-543.

6. Aoki, M. — *Control of Large Scale Dynamic Systems by Aggregation,* Preprints, Joint Automatic Control Conference (Philadelphia, Pa., 1967), pp.178-186.

7. Luenberger, D.G. — *Observing the State of a Linear System.* Institute of Electrical and Electronic Engineers, Transactions on Military Electronics, Vol.MIL-8, April 1964, pp.74-80.

8. Aoki, M.; Huddle, J.R. — *Estimation of the State Vector of a Linear Stochastic System with a Constrained Estimator.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control (Short Papers), Vol.AC-12, August 1967, pp.432-433.

9. Hanbrook, H.W. et al. — *Final Engineering Report, LORAN C/D Navigation System Integration.* Chapter 5, Contract AF-33(657)-14389. January 1966.

10. Huddle, J.R. et al. — *Final Engineering Report, Doppler-Inertail LORAN C/D Navigation System Integration.* Appendix I, Contract AF-33(615)67C-1184, October 1967.

11. Richman, J.; Friedland, B. — *Design of Optimum Mixer Filter for Aircraft Navigation Systems.* Presented at the ION Conference, Dayton, Ohio, May 1967.

12. Bellman, R. — *Introduction to Matrix Analysis,* McGraw-Hill, New York, 1960.

13. Nishimura, T. — *Error Bounds of Continuous Kalman Filters and the Application to Orbit Determination Problems.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-12, June 1967, pp.268-275.

14. Larson, R.E. — *Optimum Combined Control and Estimation for Partially Controlled Systems.* Proceedings, Fourth Allerton Conference on Circuit and System Theory, University of Illinois, Urgana, October 1966.

15. Widnall, W. — *On the Design of Nearly Optimal Linear Time-Varying Sampled Data Stochastic Controllers.* PhD, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, Massachusetts, September 1966.

16. Martin, F.H. — *Closed Loop Near-Optimum Steering for a Class of Space Missions.* ScD, Report T-413, Instrumentation Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts, 1965.

17. Johansen, D.E. — *Optimal Control of Linear Systems with Complexity Constraints.* PhD, Harvard University, Cambridge, Massachusetts, 1964.

18. Aoki, M. — *Determination of Tuning Factor in a Discrete Observation Data Incorporation Scheme.* Litton Industries Inc., Guidance and Control Division Publication No.7651, Woodland Hills, California, April 1968.

19. Pearson, J.D. — *Dynamic Decomposition Techniques.* In "Optimization Methods for Large Scale Systems", D.A. Wismer (Editor), McGraw-Hill, New York, (in press).

20. Huddle, J.R.; Wismer, D.A. — *Degradation of Linear Filter Performance Due to Modeling Error.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-13, No.4, August 1968.

21. Schmidt, S.F. — *Estimation of State with Acceptable Accuracy Constraints* Analytical Mechanics Associates, Inc., Report 67-4, Westbury, Long Island, New York, January 1967.

CHAPTER 6 - COMPARISON OF KALMAN, BAYESIAN AND MAXIMUM LIKELIHOOD ESTIMATION TECHNIQUES

by

H. W. Sorenson

University of California at San Diego
La Jolla, California 92037, USA

CHAPTER 6 - COMPARISON OF KALMAN, BAYESIAN AND MAXIMUM LIKELIHOOD ESTIMATION TECHNIQUES

H.W. Sorenson

1. PROBABILISTIC APPROACHES TO ESTIMATION THEORY

There are many different approaches to the inferential problem of estimating parameters or states from observed data. Some of these have been discussed in other chapters (e.g., Chapter 1). In this chapter the problem is approached by dealing with the probability density functions describing the state and measurement variables. This general probabilistic formulation is referred to as the *Bayesian approach*[1-3] and it provides a framework within which many other approaches can be subsumed. Using the Bayesian formulation, one arrives naturally at the maximum *a posteriori* and maximum likelihood estimation procedures[4]. It is also seen that deterministic least-squares can be reinterpreted as maximum *a posteriori* estimation theory. In Chapter 1 the Kalman filter equations have been derived as the solution of the unbiased, minimum variance estimation problem. It is pointed out here that the minimum variance criterion is just one of many that could be chosen which all yield the same estimate for linear, gaussian systems.

The estimation problem for time-discrete, nonlinear, stochastic systems is formulated in Section 1.2. Within the Bayesian framework, a general treatment of nonlinear systems can be developed in a straightforward manner. As indicated in Sections 2 and 3, it is necessary to introduce additional restrictions in order to obtain practical solutions, but it is felt that the general formulation and subsequent simplifications provide desirable insights into the nonlinear estimation problem.

1.1 Deterministic Least-squares Estimation

As a preliminary to the probabilistic discussion of estimation theory that is found in Sections 2 and 3, the problem of estimating parameters or states from measurement data is first treated using deterministic least-squares. This formulation leads to a minimization problem whose solution is nontrivial. It is pointed out in Section 3 that the deterministic least-squares approach is equivalent to the maximum *a posteriori* approach when the appropriate assumptions are introduced.

To simplify the discussion, first consider the problem of estimating an unknown parameter $\underline{x}$ from measurement data $\underline{z}_i (i = 1, 2, \ldots, N)$, where the collection of data will be denoted as Z_N. Suppose that the data Z_N are obtained at discrete instants of time $t_i (i = 1, 2, \ldots, N)$ and that the $\underline{z}_i$ is an m-dimensional vector representing m independent measurements. These data contain unknown errors $\underline{v}_i$ and are related to the parameters according to

$$\underline{z}_i = \underline{h}_i(\underline{x}) + \underline{v}_i ; \qquad i = 1, 2, \ldots, N . \tag{1.1}$$

where the $\underline{h}_i$ are known functions of the n-dimensional vector of parameter $\underline{x}$.

In the classical *least-squares* procedure, one treats (1.1) in a deterministic fashion and finds the value of $\underline{x}$ that minimizes the sum of the squares of the error. For example, combine the N measurement vectors into one mN vector (or, suppose N = 1) and denote it as

$$Z_N = \underline{h}(\underline{x}) + \underline{V} , \tag{1.2}$$

where Z_N, $\underline{h}$ and $\underline{V}$ have obvious definitions. Then, choose the $\underline{x}$ so that

$$L = [Z_N - \underline{h}(\underline{x})]^T [Z_N - \underline{h}(\underline{x})] \tag{1.3}$$

is minimized. Since there are no other constraints, a necessary condition for L to be minimized is*

$$-\left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)^T Z_N + \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)^T \underline{h}(\underline{x}) = 0 . \tag{1.4}$$

* The following convention is used. The partial derivative of a scalar function h of an n-dimensional vector variable $\underline{x}$ is denoted by $\partial h/\partial \underline{x}$ and is a $1 \times n$ matrix whose elements are $\partial h/\partial x_1$, $\partial h/\partial x_2$, ..., $\partial h/\partial x_n$. If an m-dimensional vector function $\underline{h}$ is differentiated, then $\partial \underline{h}/\partial \underline{x}$ is an $m \times n$ matrix with the ij^{th} element equal to $\partial h_i/\partial x_j$. Also, the second derivative of a scalar function h is written as $\partial/(\partial h/\partial \underline{x})^T/\partial \underline{x}$, and is an $n \times n$ matrix whose ij^{th} element is $\partial^2 h/\partial x_j \partial x_i$.

It is not generally possible to solve (1.4) explicitly for $\underline{x}$. However, if (1.2) is linear so that

$$\underline{Z}_N = H\underline{x} + \underline{V} , \tag{1.5}$$

where H is an $mN \times n$ matrix, then

$$\frac{\partial \underline{h}}{\partial \underline{x}} = H^T ,$$

so that (1.4) becomes

$$(H^TH)\underline{x} = H^T\underline{Z}_N . \tag{1.6}$$

Denoting the solution of (1.6) by $\hat{\underline{x}}$, the least-squares estimate, supposing that (H^TH) has an inverse, is seen to be

$$\hat{\underline{x}} = (H^TH)^{-1} H^T\underline{Z}_N . \tag{1.7}$$

Observe in (1.7) that an $n \times n$ matrix must be inverted and that *all* of the measurement data are processed. The positive-definiteness of (H^TH) is equivalent to the observability of the system (1.5) as introduced in Chapter 1. It is possible to generate least-squares estimates recursively as new data are obtained and thereby avoid reprocessing old data[5].

Suppose $\hat{\underline{x}}_{k-1}$ is the estimate of $\underline{x}$ based on the data $\underline{z}_1, \underline{z}_2, \ldots, \underline{z}_{k-1}$. Let $\underline{Z}_{k-1}$ denote the set of all these data and suppose that new data $\underline{z}_k$ are obtained. Let

$$L_k = \sum_{i=1}^{k} [\underline{z}_i - H_i\underline{x}]^T[\underline{z}_i - H_i\underline{x}] . \tag{1.8}$$

Note that (1.8) is identical with (1.3), except that the individual samples are written explicitly and the system is assumed to be described by a linear version of (1.1).

$$L_k = \sum_{i=1}^{k-1} [\underline{z}_i - H_i\underline{x}]^T[\underline{z}_i - H_i\underline{x}] + [\underline{z}_k - H_k\underline{x}]^T[\underline{z}_k - H_k\underline{x}]$$

$$= L_{k-1} + [\underline{z}_k - H_k\underline{x}]^T[\underline{z}_k - H_k\underline{x}] .$$

To determine the estimate that minimizes L_k, form

$$\frac{\partial L_k}{\partial \underline{x}} = 0 = \frac{\partial K_{k-1}}{\partial \underline{x}} - 2\,\underline{z}_k^TH_k + 2\,\underline{x}^TH_k^TH_k$$

or, letting

$$H^{k-1} \triangleq \begin{bmatrix} H_1 \\ H_2 \\ \cdot \\ \cdot \\ \cdot \\ H_{k-1} \end{bmatrix} ,$$

this condition becomes

$$-2(\underline{Z}_{k-1})^TH^{k-1} + 2\hat{\underline{x}}_k^T(H^{k-1})^TH^{k-1} - 2\underline{z}_k^TH_k + 2\hat{\underline{x}}_k^TH_k^TH_k = 0 .$$

Thus, one obtains

$$\hat{\underline{x}}_k^T[(H^{k-1})^TH^{k-1} + H_k^TH_k] = (\underline{Z}_{k-1})^TH^{k-1} + \underline{z}_k^TH_k$$

$$= \hat{\underline{x}}_{k-1}^T(H^{k-1})^TH^{k-1} + \underline{z}_k^TH_k . \tag{1.9}$$

Guided by the intuition provided by Chapter 1, let

$$P_k^{-1} \triangleq (H^k)^TH^k . \tag{1.10}$$

Then (1.9) becomes

$$P_k^{-1}\hat{\underline{x}}_k = P_{k-1}^{-1}\hat{\underline{x}}_{k-1} + H_k^T\underline{z}_k .$$

Assuming that P_k^{-1} has an inverse, one obtains

$$\hat{x}_k = P_k P_{k-1}^{-1} \hat{x}_{k-1} + P_k H_k^T z_k \quad (1.11)$$

But it follows, using the Matrix Inversion Lemma (see Section 2.1.4 of Chapter 1), that

$$\begin{aligned} P_k P_{k-1}^{-1} &= (P_{k-1}^{-1} + H_k^T H_k)^{-1} P_{k-1}^{-1} \\ &= I - P_{k-1} H_k^T [H_k P_{k-1} H_k^T + I]^{-1} H_k \\ &\triangleq I - K_k H_k \end{aligned}$$

where

$$K_k = P_{k-1} H_k^T [H_k P_{k-1} H_k^T + I]^{-1} \quad (1.12a)$$

As shown in Chapter 1, it follows that

$$K_k = P_k H_k^T \quad (1.12b)$$

so (1.11) reduces to

$$\hat{x}_k = \hat{x}_{k-1} + K_k [z_k - H_k \hat{x}_{k-1}] \quad (1.13)$$

There is a generalization that is particularly useful. Suppose that certain measurement errors are to be given more weight than others. Then, the cost function can be written as a weighted least-squares sum,

$$L_k = \sum_{i=1}^{k} [z_i - H_i x]^T R_i^{-1} [z_i - H_i x] \quad (1.14)$$

It follows, by forming $\partial L_k / \partial x$, that

$$\frac{\partial L_k}{\partial x_k} = 0 = \frac{\partial L_{k-1}}{\partial x} - 2 z_k^T H_k R_k^{-1} + 2 x^T H_k^T R_k^{-1} H_k$$

Thus, the weighting matrix R_k^{-1} modifies only the matrices P_k and K_k,

$$P_k = [P_{k-1}^{-1} + H_k^T R_k^{-1} H_k]^{-1} \quad (1.15)$$

and

$$K_k = P_{k-1} H_k^T [H_k P_{k-1} H_k^T + R_k]^{-1} \quad (1.16a)$$

$$= P_k H_k^T R_k^{-1} . \quad (1.16b)$$

Clearly, deterministic least-squares estimates are closely related to the Kalman filter estimates. However, no probabilistic interpretations have been attempted. As will be seen below, the problem can be couched in a framework that provides a probabilistic interpretation of least-squares as the *maximum a posteriori* or *most probable estimate*.

Thus far, the parameters x have been treated as constants. Another way of considering these parameters is to view them as *state variables* described by the plant equation

$$x_k = x_{k-1} .$$

The extension to a more general model for the plant is not especially difficult. Suppose that the plant is described by

$$x_k = f_k(x_{k-1}) + w_{k-1} ; \qquad k = 1, 2, \ldots \quad (1.17)$$

and the measurements involve only the current state

$$z_k = h_k(x_k) + v_k ; \qquad k = 1, 2, \ldots . \quad (1.18)$$

Equation (1.17) represents a constraint on the system. The w_{k-1} represent uncertainties or errors in the model that cannot otherwise be accounted for. w_{k-1} will be considered as parameters that are chosen to minimize the estimation error.

Consider the problem of choosing the sequences x_k and w_{k-1} to minimize the mean-square error

$$L_k = [x_0 - a_0]^T P_0^{-1} [x_0 - a_0] + \sum_{i=1}^{k} ([z_i - h_i(x_i)]^T R_i^{-1} [z_i - h_i(x_i)] + w_{i-1}^T Q_{i-1}^{-1} w_{i-1}) \quad (1.19)$$

subject to the constraint (1.17). The first term in (1.19), is included to account for some initial estimate of the state $\underline{x}_0$. Weighting matrices P_0^{-1}, R_i^{-1}, Q_i^{-1} are included for generality.

This discussion has been included to set the stage for later discussion. As should be anticipated, the solution of this general problem is difficult. However, when (1.17) and (1.18) are linear, one obtains the Kalman filter equations. This aspect is dealt with again in Section 3.3.

1.8 General Estimation Problem for Time-Discrete Stochastic Systems

It is possible to approach the linear estimation problem in a variety of ways. Too often, however, the treatment of the linear problem does not indicate the manner in which one can apply the technique to nonlinear problems, nor are the difficulties presented by nonlinear systems made apparent. It is the intent in this chapter to describe a nonlinear estimation problem and to discuss two of the methods used to attack this problem. After the general character of the approaches is defined, the discussion is specialized to linear systems to obtain the Kalman filter equations. Time-discrete systems are considered in order to simplify the mathematics and thereby avoid the difficulties that have led to the development of the stochastic calculuses of Ito[6] and of Stratonovich[7] for time-continuous systems.

Consider a dynamical system described by a nonlinear difference equation in which the state evolves according to

$$\underline{x}_k = f_k(\underline{x}_{k-1}, \underline{w}_{k-1}) . \tag{1.20}$$

The state has an initial state $\underline{x}_0$ which is a random variable with a known probability density function*, say $p(\underline{x}_0)$. The $\underline{w}_{k-1}$ represents a sample of a random sequence with known probability density function. Throughout this discussion, it is assumed that the $\underline{w}_i$ are independent between the sampling times and, thereby, constitute a *white-noise sequence*. Thus

$$p(\underline{w}_0, \underline{w}_1, \ldots, \underline{w}_k) = p(\underline{w}_0)p(\underline{w}_1)\ldots p(\underline{w}_k) .$$

The state is observed through related measurement data $\underline{z}_k$ described by

$$\underline{z}_k = h_k(\underline{x}_k, \underline{v}_k) ; \qquad k = 1, 2, \ldots , \tag{1.21}$$

The $\underline{v}_k$ represent a sample from a white-noise sequence for which the probability density function is known and given by

$$p(\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_k) = p(\underline{v}_1)p(\underline{v}_2)\ldots p(\underline{v}_k) .$$

The plant and measurement noise sequences are assumed to be independent of each other and of the initial state. This assumption could be eliminated without significant conceptual difficulty but the notation becomes more complicated.

The filtering problem† of estimating $\underline{x}_k$ from measurement data Z_k will be considered almost exclusively. The solution of the prediction problem is a trivial result. As shown in Chapter 1, the solution of the smoothing problem is obtained from the filtering solution, so it is not discussed here.

The problem has been cast in a probabilistic mold that reflects the uncertainty in the dynamic model and the measurements that commonly exist in physical systems. Admittedly, this formulation might be considered to be artificial, since one is unlikely to have a good knowledge of the density functions involved, so the uncertainties appear to have been compounded by their introduction. Setting this objection aside, it will be seen that the classical least-squares approach is imbedded in this formulation, so that one does not obscure the problem by these considerations. In fact, one is led to the conclusion that the least-squares philosophy is enriched and deepened by recasting the problem in this manner and that new insights are obtained thereby.

The Bayesian approach provides the theoretical structure within which a variety of approaches can be considered. In Section 2 the general results of this method are presented and the Kalman filter equations are derived. The maximum likelihood and maximum *a posteriori* estimation procedures are discussed in Section 3. The Cramer-Rao inequality is discussed and the Fisher information matrix is related to the error covariance matrix of the Kalman filter.

---

* The notation that is used follows Fel'dbaum[1] and has the disadvantage that the argument of the function serves a dual purpose. It is used to name the function, as here, and is also treated as a variable name (e.g., it is treated as the variable of integration). The meaning should be clear from the context.

† The reader is directed to Chapter 1 for a definition of the terms relating to the different aspects of the estimation problem.

2. THE BAYESIAN APPROACH

In the "Bayesian approach" to the filtering problem one is concerned first of all with the determination of the *a posteriori* density function $p(\underline{x}_k|\underline{Z}_k)$. This density function provides all the information required for the solution of this problem. This statement will be discussed in more detail below. The following four aspects of conditional density functions will be used frequently*.

(i) For random variables $\underline{a}$ and $\underline{b}$ with joint probability density function $p(\underline{a},\underline{b})$, the conditional density of $\underline{a}$, given $\underline{b}$, is defined as

$$p(\underline{a}|\underline{b}) \triangleq \frac{p(\underline{a},\underline{b})}{p(\underline{b})} . \tag{2.1}$$

(ii) For random variables $\underline{a}$, $\underline{b}$, and $\underline{c}$, it follows from (2.1) that

$$p(\underline{a},\underline{b}|\underline{c}) = p(\underline{b}|\underline{c})p(\underline{a}|\underline{b},\underline{c}) . \tag{2.2}$$

This is known as the chain rule.

(iii) Further, it can be seen from (2.2) and the properties of density functions that*

$$p(\underline{a}|\underline{b}) = \int p(\underline{a}|\underline{b},\underline{c})p(\underline{b}|\underline{c})\, d\underline{b} . \tag{2.3}$$

This is an integrated form of the chain rule and is essentially the Chapman-Kolmogorov equation.

(iv) Equation (2.1) also implies that

$$p(\underline{a}|\underline{b}) = \frac{p(\underline{b}|\underline{a})p(\underline{a})}{p(\underline{b})} . \tag{2.4}$$

This relation is known as Bayes's rule and is the source for the term used to describe the approach in this section.

2.1 Performance Criteria Considerations

Equations (1.20) and (1.21) and the concomitant density functions provide the information required for the Bayesian formulation of the problem. Before going to the development of the *a posteriori* density relations, first note that the solution of the unbiased, minimum variance estimation problem, as posed in Chapter 1, is obtained if the *a posteriori* density of the state conditioned on all available measurement data is known. This is demonstrated in the following statement.

Theorem 2.1

Suppose that a random variable $\underline{x}$ is to be estimated from measurement data $\underline{Z}$ and suppose that $\underline{x}$ and $\underline{Z}$ have the joint probability density function $p(\underline{x},\underline{Z})$. The unbiased estimate of $\underline{x}$ based on the data $\underline{Z}$ that yields the minimum error variance

$$E[(\underline{x}-\hat{\underline{x}})^T(\underline{x}-\hat{\underline{x}})] = \text{minimum}$$

is given by

$$\hat{\underline{x}} = E[\underline{x}|\underline{Z}] . \tag{2.5}$$

Proof: First, write the error variance in terms of the conditional density, using the identity

$$E[(\underline{x}-\hat{\underline{x}})^T(\underline{x}-\hat{\underline{x}})] = E(E[(\underline{x}-\hat{\underline{x}})^T(\underline{x}-\hat{\underline{x}})|\underline{Z}]) .$$

Clearly, to minimize the error variance it is sufficient to minimize the conditional expectation. Thus, consider

$$\begin{aligned} E[(\underline{x}-\hat{\underline{x}})^T(\underline{x}-\hat{\underline{x}})|\underline{Z}] &= \hat{\underline{x}}^T\hat{\underline{x}} - 2\hat{\underline{x}}^T E[\underline{x}|\underline{Z}] + E[\underline{x}^T\underline{x}|\underline{Z}] \\ &= (\hat{\underline{x}}-E[\underline{x}|\underline{Z}])^T(\hat{\underline{x}}-E[\underline{x}|\underline{Z}]) + \\ &\quad + E[\underline{x}^T\underline{x}|\underline{Z}] - (E[\underline{x}|\underline{Z}])^T E[\underline{x}|\underline{Z}] . \end{aligned} \tag{2.6}$$

Only the first term in (2.6) involves $\hat{\underline{x}}$. It is quadratic, so that the smallest value it can assume is zero. This obtains when

$$\hat{\underline{x}} = E[\underline{x}|\underline{Z}] .$$

The $\hat{\underline{x}}$ given by (2.5) is unbiased, as is seen easily by observing that

$$E[\hat{\underline{x}}] = E(E[\underline{x}|\underline{Z}]) = E[\underline{x}] .$$

---

* The single integral sign in (2.3) and the differential $d\underline{b}$ are used to indicate integrations involving vector variables. When more than one vector is involved, the differential will be written as $d(\underline{a},\underline{b},\underline{c},\ldots)$.

This completes the proof and demonstrates that the first moment of $p(\underline{x}|\underline{Z})$ provides the unbiased, minimum variance estimate of $\underline{x}$.

The minimum error variance criterion is only one of many criteria that could be selected. It has been frequently stated that it is used because it is more tractable analytically and not always because it is the most appropriate for the problem. This implies that the minimum variance estimates represent a very specialized class. However, Sherman[9] demonstrated that the results obtained for this criterion actually apply to a broader class of cost functions that includes many other meaningful criteria.

Consider a general error criterion $L(\tilde{\underline{x}})$ where $\tilde{\underline{x}} \triangleq (\underline{x} - \hat{\underline{x}})$ and the L has the following characteristics:

(i) L is symmetric so that

$$L(\tilde{\underline{x}}) = L(-\tilde{\underline{x}}) . \tag{2.7}$$

(ii) L is convex so that

$$L[\lambda\tilde{\underline{x}}_1 + (1-\lambda)\tilde{\underline{x}}_2] \leq \lambda L(\tilde{\underline{x}}_1) + (1-\lambda)L(\tilde{\underline{x}}_2) , \qquad \lambda \geq 0 . \tag{2.8}$$

This class certainly includes the minimum variance criterion

$$L_{mv} = \tilde{\underline{x}}^T\tilde{\underline{x}}$$

and the absolute error criterion

$$L_{abs.} = |\tilde{\underline{x}}| .$$

Suppose that the $\underline{x}$ is to be estimated from data $\underline{Z}$. Then the following result is applicable.

Theorem 2.2:

If the conditional density $p(\underline{x}|\underline{Z})$ is symmetric about its mean value, then the estimate $\hat{\underline{x}}$ that minimizes any cost function L in this class is identical with the estimate $\hat{\underline{x}}_{mv}$ obtained with the minimum variance criterion.

Proof: It is desired that the estimate $\hat{\underline{x}}$ that minimizes the expected value of $L(\tilde{\underline{x}})$ be determined. First, note that

$$E[L(\tilde{\underline{x}})|\underline{Z}] = E[L(\underline{x} - \hat{\underline{x}})|\underline{Z}] = E[L(\hat{\underline{x}} - \underline{x})|\underline{Z}] ,$$

where the symmetry of L has been used.

Let

$$\underline{\varepsilon} \triangleq \underline{x} - \hat{\underline{x}}_{mv}$$
$$= \underline{x} - E[\underline{x}|\underline{Z}] .$$

Then, observing that

$$E[L(\underline{x} - \hat{\underline{x}})|\underline{Z}] = \int_{-\infty}^{\infty} L(\underline{x} - \hat{\underline{x}})p(\underline{x}|\underline{Z})\, d\underline{x} ,$$

and using the definition of $\underline{\varepsilon}$, this becomes

$$E[L(\underline{x} - \hat{\underline{x}})|\underline{Z}] = \int_{-\infty}^{\infty} L(\underline{\varepsilon} + \hat{\underline{x}}_{mv} - \hat{\underline{x}})p(\underline{\varepsilon}|\underline{Z})\, d\underline{\varepsilon} .$$

But it has been assumed that $p(\underline{x}|\underline{Z})$ is symmetric about the mean, so this implies that

$$p(\underline{\varepsilon}|\underline{Z}) = p(-\underline{\varepsilon}|\underline{Z}) .$$

Thus one obtains

$$E[L(\underline{x} - \hat{\underline{x}})|\underline{Z}] = \int_{-\infty}^{\infty} L(-\underline{\varepsilon} + \hat{\underline{x}}_{mv} - \hat{\underline{x}})p(\underline{\varepsilon}|\underline{Z})\, d\underline{\varepsilon}$$

$$= \int_{-\infty}^{\infty} L(\hat{\underline{x}} - \hat{\underline{x}}_{mv} + \underline{\varepsilon})p(\underline{\varepsilon}|\underline{Z})\, d\underline{\varepsilon}$$

$$= \int_{-\infty}^{\infty} L(\hat{\underline{x}} - \hat{\underline{x}}_{mv} - \underline{\varepsilon})p(\underline{\varepsilon}|\underline{Z})\, d\underline{\varepsilon} .$$

Using these identities, it follows that

$$E[L(\underline{x} - \hat{\underline{x}})|\underline{Z}] = \tfrac{1}{2} E[L(\underline{\varepsilon} + (\hat{\underline{x}} - \hat{\underline{x}}_{mv}))|\underline{Z}] + \tfrac{1}{2} E[L(\underline{\varepsilon} - (\hat{\underline{x}} - \hat{\underline{x}}_{mv}))|\underline{Z}] .$$

But L is convex so

$$E[L(\underline{x}-\hat{\underline{x}})|\underline{Z}] \geq E(L[\tfrac{1}{2}(\underline{\varepsilon}+(\hat{\underline{x}}-\hat{\underline{x}}_{mv})) + \tfrac{1}{2}(\underline{\varepsilon}-(\hat{\underline{x}}-\hat{\underline{x}}_{mv}))]|\underline{Z}) = E(L(\underline{\varepsilon})|\underline{Z}) .$$

From the definition of $\underline{\varepsilon}$, it can be seen that equality occurs if $\hat{\underline{x}} = \hat{\underline{x}}_{mv}$.

Observe that, if L is required to be strictly convex, the $\hat{\underline{x}}$ is unique and must equal $\hat{\underline{x}}_{mv}$.

This result can be extended to non-convex cost functions by considering the following statement.

Theorem 2.3

Suppose that L is symmetric and non-decreasing and that $p(\underline{x}|\underline{Z})$ is symmetric about the mean. Suppose also that $p(\underline{x}|\underline{Z})$ is unimodal and satisfies the conditions

$$\lim_{\underline{x}\to\infty} L(\underline{x})p(\underline{x}|\underline{Z}) = 0 .$$

Then the $\hat{\underline{x}}$ that minimizes a cost function L in this class is identical to the minimum variance estimate $\hat{\underline{x}}_{mv}$.

The proof is similar to that of Theorem 2.2 and is omitted.

This theorem permits a variety of non-convex cost functions to be considered. For example it is sometimes more meaningful to weigh equally all errors larger than a certain magnitude, since any errors larger than the prescribed limit may be undesirable. Thus, one could consider a modified minimum variance criterion in which

$$L(\underline{\tilde{x}}) = \begin{cases} [\underline{\tilde{x}}^T\underline{\tilde{x}}] \\ k^2 \end{cases} , \qquad \begin{matrix} |\underline{\tilde{x}}| \leq k \\ |\underline{\tilde{x}}| > k . \end{matrix}$$

Another example would be the uniform cost function in which errors within a certain magnitude are accepted without cost and all other errors are weighted equally. This is described by the following function

$$L(\underline{\tilde{x}}) = \begin{cases} 0 , & |\underline{\tilde{x}}| \leq \dfrac{k}{2} \\ 1 , & |\underline{\tilde{x}}| > \dfrac{k}{2} . \end{cases}$$

This cost criterion is closely connected to the maximum *a posteriori* density. Consider the expected cost for this criterion,

$$E[L(\underline{\tilde{x}})] = E(E[L(\underline{\tilde{x}})|\underline{Z}])$$

$$= \int_{-\infty}^{\infty} \left\{ \int_{-\infty}^{\infty} L(\underline{\tilde{x}})p(\underline{x}|\underline{Z})\, d\underline{x} \right\} p(\underline{Z})\, d\underline{Z}$$

$$= \int_{-\infty}^{\infty} \left\{ 1 - \int_{\hat{\underline{x}}_{unf}-k/2}^{\hat{\underline{x}}_{unf}+k/2} p(\underline{x}|\underline{Z})\, d\underline{x} \right\} p(\underline{Z})\, d\underline{Z} ,$$

where $\hat{\underline{x}}_{unf}$ is the estimate associated with this criterion. The cost is minimized by maximizing the inner integral. For small enough values of k, the best choice is essentially the maximum value of the *a posteriori* density. Estimates of this type are discussed in Section 3. The theorem implies that the maximum *a posteriori* estimates are equivalent to the minimum variance estimates when the appropriate conditions are satisfied.

2.2 The *a posteriori* Density Function

The *a posteriori* density function can be seen from the preceding discussion to provide all of the information required to determine estimates for any cost functions. One of the principal advantages of the Kalman filter equations is their recursive character, which enables new measurements to be processed without reprocessing older data. Thus, consider the problem of determining the filtering density in a recursive fashion. The following theorem provides the desired relations.

Theorem 2.4

For the system (1.20) - (1.21), the *a posteriori* density function $p(\underline{x}_k|\underline{Z}_k)$ evolves according to

$$p(\underline{x}_k|\underline{Z}_k) = \frac{p(\underline{x}_k|\underline{Z}_{k-1})p(\underline{z}_k|\underline{x}_k)}{p(\underline{z}_k|\underline{Z}_{k-1})} , \tag{2.9}$$

where the normalizing constant is

$$p(\underline{Z}_k|\underline{Z}_{k-1}) = \int p(\underline{x}_k|\underline{Z}_{k-1})p(\underline{z}_k|\underline{x}_k)\, d\underline{x}_k .$$

The prediction density $p(\underline{x}_k|\underline{Z}_{k-1})$ is described by

$$p(\underline{x}_k|\underline{Z}_{k-1}) = \int p(\underline{x}_k|\underline{x}_{k-1})p(\underline{x}_{k-1}|\underline{Z}_{k-1})\, d\underline{x}_{k-1} \tag{2.10}$$

The initial condition $p(\underline{x}_0|\underline{Z}_0)$ is given by

$$p(\underline{x}_0|\underline{Z}_0) = \frac{p(\underline{Z}_0|\underline{x}_0)p(\underline{x}_0)}{p(\underline{Z}_0)} .$$

Proof: First, note that the initial condition follows immediately from Bayes's rule (2.4).

Consider arbitrary k. From the chain rule (2.2), one sees that

$$p(\underline{x}_k, \underline{z}_k|\underline{Z}_{k-1}) = p(\underline{x}_k|\underline{Z}_k)p(\underline{z}_k|\underline{Z}_{k-1}) ;$$

so that

$$p(\underline{x}_k|\underline{Z}_k) = \frac{p(\underline{x}_k, \underline{z}_k|\underline{Z}_{k-1})}{p(\underline{z}_k|\underline{Z}_{k-1})} .$$

The chain rule also enables one to write

$$p(\underline{x}_k, \underline{z}_k|\underline{Z}_{k-1}) = p(\underline{z}_k|\underline{x}_k, \underline{Z}_{k-1})p(\underline{x}_k|\underline{Z}_{k-1}) ,$$

which can be simplified to

$$p(\underline{x}_k, \underline{z}_k|\underline{Z}_{k-1}) = p(\underline{z}_k|\underline{x}_k)p(\underline{x}_k|\underline{Z}_{k-1}) ,$$

since $\underline{z}_k$, given $\underline{x}_k$, is independent of $\underline{Z}_{k-1}$. Equating the two relations for $p(\underline{x}_k, \underline{z}_k|\underline{Z}_{k-1})$, one obtains

$$p(\underline{x}_k|\underline{Z}_k) = \frac{p(\underline{x}_k|\underline{Z}_{k-1})p(\underline{z}_k|\underline{x}_k)}{p(\underline{z}_k|\underline{Z}_{k-1})} ,$$

which proves (2.9).

The normalizing constant is determined immediately by using the condition that

$$\int p(\underline{x}_k|\underline{Z}_k)\, d\underline{x}_k = 1 .$$

The proof of (2.10) follows from the integrated chain rule (2.3). Note that

$$p(\underline{x}_k|\underline{Z}_{k-1}) = \int p(\underline{x}_k|\underline{x}_{k-1}, \underline{Z}_{k-1})p(\underline{x}_{k-1}|\underline{Z}_{k-1})\, d\underline{x}_{k-1} ,$$

which reduces to

$$p(\underline{x}_k|\underline{Z}_{k-1}) = \int p(\underline{x}_k|\underline{x}_{k-1})p(\underline{x}_{k-1}|\underline{Z}_{k-1})\, d\underline{x}_{k-1} ,$$

thereby proving (2.10) and completing the proof of the lemma.

Several characteristics of (2.9) and (2.10) require discussion. First, note that the density $p(\underline{z}_k|\underline{x}_k)$ in (2.9) is defined by the measurement model (1.21) and the prescribed density function for the measurement noise. Similarly, the density $p(\underline{x}_k|\underline{x}_{k-1})$ appearing in (2.10) is defined by the plant Equation (1.20) and the density function prescribed for the plant noise. Thus, theoretically, one knows these two functions. The initial *a posteriori* density $p(\underline{x}_0|\underline{z}_0)$ is known from *a priori* information so it is possible to determine the *a posteriori* density $p(\underline{x}_k|\underline{Z}_k)$ for any subsequent value of k.

In a practical sense several difficulties hinder greatly or prevent $p(\underline{x}_k|\underline{Z}_k)$ from being determined.

(1) The plant and measurement equations, $\underline{f}_k$ and $\underline{h}_k$, may inhibit the determination of the $p(\underline{z}_k|\underline{x}_k)$ and $p(\underline{x}_k|\underline{x}_{k-1})$. To avoid this difficulty, it is frequently assumed that the plant and measurement noise are additive. Then the plant and measurement equations take the form

$$\underline{x}_k = \underline{f}_k(\underline{x}_{k-1}) + \underline{u}_{k-1} \tag{2.11}$$

and

$$\underline{z}_k = \underline{h}_k(\underline{x}_k) + \underline{v}_k . \tag{2.12}$$

It is also common to assume that these noise sequences are gaussian, with densities

$$p(\underline{w}_k) = k_w \exp(-\tfrac{1}{2}\underline{w}_k^T Q_k^{-1}\underline{w}_k) \tag{2.13}$$

$$p(\underline{v}_k) = k_v \exp(-\tfrac{1}{2}\underline{v}_k^T R_k^{-1}\underline{v}_k) . \tag{2.14}$$

Using (2.11) - (2.14), it follows that

$$p(\underline{x}_k|\underline{x}_{k-1}) = k_w \exp\left\{-\tfrac{1}{2}[\underline{x}_k - \underline{f}_k(\underline{x}_{k-1})]^T Q_{k-1}^{-1}[\underline{x}_k - f_k(\underline{x}_{k-1})]\right\} \tag{2.15}$$

$$p(\underline{z}_k|\underline{x}_k) = k_v \exp\left\{-\tfrac{1}{2}[\underline{z}_k - \underline{h}(\underline{x}_k)]^T R_k^{-1}[\underline{z}_k - h_k(\underline{x}_k)]\right\}. \tag{2.16}$$

This shows that the difficulty in determining these conditional densities is avoided by assuming additive noise.

(ii) The integration required in (2.10) cannot generally be accomplished in a closed form. The principal exception occurs when the plant and measurement equations are linear and the initial state and the noise sequences are gaussian.

(iii) Equation (2.9) requires the multiplication of two functions so, if they are known, the $p(\underline{x}_k|\underline{Z}_k)$ can at least be determined to within a multiplicative constant. It is impossible in most instances to compute moments or expected values of particular quantities (e.g. cost functions) in a closed form, so that the estimation problem can not actually be solved through knowledge of the *a posteriori* density function. Again the major exception occurs when the system is linear and gaussian. This case is treated later in this section where it is seen that the Kalman filter equations describe the mean and covariance of the gaussian *a posteriori* density function.

To circumvent the difficulties described in items (ii) and (iii), methods for approximating the density function have been proposed[3, 10]. These aspects are beyond the scope of this chapter.

The recursive *a posteriori* density function relations for time-discrete systems have their analog for time-continuous systems. Stratonovich[11] first derived a partial differential equation to describe the evolution of $p(\underline{x}(t)|\underline{Z}(t))$. Subsequently, Kushner[12] and, then Bucy[13] modified these results consistent with the Ito stochastic calculus. These equations for the *a posteriori* density are very difficult to solve. Linear, gaussian systems are again the principal case for which solutions can be obtained. Fisher[14] has attempted to obtain approximate solutions.

2.3 Linear, Gaussian Systems

Suppose that the plant and measurement data are described by linear equations so that the system is

$$\underline{x}_k = \Phi_{k,k-1}\underline{x}_{k-1} + \underline{w}_{k-1} \tag{2.17}$$

$$\underline{z}_k = H_k\underline{x}_k + \underline{v}_k , \tag{2.18}$$

where the initial state is a gaussian random variable with density function

$$p(\underline{x}_0) = k_0 \exp(-\tfrac{1}{2}(\underline{x}_0 - \underline{\bar{x}})^T M_0^{-1}(\underline{x}_0 - \underline{\bar{x}})) \tag{2.19}$$

and the plant and measurement noise sequences are gaussian white-noise with density functions defined by (2.11) and (2.14).

For this linear, gaussian system the *a posteriori* density function is characterized by the following result.

Theorem 2.5

The *a posteriori* density $p(\underline{x}_k|\underline{Z}_k)$ for the system (2.17), (2.18) is gaussian

$$p(\underline{x}_k|\underline{Z}_k) = [(2\pi)^n|P_k|]^{-1/2} \exp(-\tfrac{1}{2}(\underline{x}_k - \hat{\underline{x}}_k)^T P_k^{-1}(\underline{x}_k - \hat{\underline{x}}_k)) , \tag{2.20}$$

with mean value given by

$$\hat{\underline{x}}_k = \hat{\underline{x}}_k' + K_k(\underline{z}_k - H_k\hat{\underline{x}}_k') . \tag{2.21}$$

The $\hat{\underline{x}}_k'$ represents the mean value of the *a posteriori* density $p(\underline{x}_k|\underline{Z}_{k-1})$ and is

$$\hat{\underline{x}}_k' = E[\underline{x}_k|\underline{Z}_{k-1}] = \Phi_{k,k-1}\hat{\underline{x}}_{k-1} . \tag{2.22}$$

The matrix P_k' is the covariance of $p(\underline{x}_k|\underline{Z}_{k-1})$ and is

$$P_k' = E[(\underline{x}_k - \hat{\underline{x}}_k')(\underline{x}_k - \hat{\underline{x}}_k')^T|\underline{Z}_{k-1}] = \Phi_{k,k-1}P_{k-1}\Phi_{k,k-1}^T + Q_{k-1} . \tag{2.23}$$

K_k is a gain matrix defined by

$$K_k = P_k' H_k^T (H_k P_k' H_k^T + R_k)^{-1} . \tag{2.24}$$

The covariance matrix of $p(\underline{x}_k|\underline{Z}_k)$ is defined as P_k and is given by

$$P_k = E[(\underline{x}_k - \hat{\underline{x}}_k)(\underline{x}_k - \hat{\underline{x}}_k)^T|\underline{Z}_k] = P_k' - K_k H_k P_k' . \tag{2.25}$$

Assuming the first measurement occurs at t_0, the initial density is $p(\underline{x}_0)$; so

$$\hat{\underline{x}}_0 = \underline{a} + K_0(\underline{z}_0 - H_0\underline{a}) , \tag{2.26}$$

where $$K_0 = M_0 H_0^T (H_0 M_0 H_0^T + R_0)^{-1}$$

and $$P_0 = M_0 - K_0 H_0 M_0 . \tag{2.27}$$

Note: The equations describing the conditional mean and covariance are identical with the Kalman filter equations. This is not unexpected, since it has been shown that the conditional mean provides the minimum variance estimate. Thus in the linear case the Bayesian approach yields the results presented in Chapter 1. It should also be noted that the covariance of the conditional density is independent of the measurement data. As a result the conditional covariance is identical with the covariance of the error in the estimate. That is,

$$P_k = E[(\underline{x}_k - \hat{\underline{x}}_k)(\underline{x}_k - \hat{\underline{x}}_k)^T|\underline{Z}_k] = E[(\underline{x}_k - \hat{\underline{x}}_k)(\underline{x}_k - \hat{\underline{x}}_k)^T] .$$

It is interesting that the proof[3] of the theorem is most easily accomplished by resorting to the use of characteristic functions. This approach is taken here. Certainly, direct evaluation of the general recursion relations (2.9) and (2.10) will provide an equivalent result, but it is interesting to note that it is the result stated in Section 2.1.4 of Chapter 1 as an alternative form. This occurs because the density functions involve the inverse of the covariance matrix; so one obtains P_k^{-1} rather than P_k when proceeding directly from (2.9) and (2.10). On the other hand characteristic functions require the covariance matrices and not their inverses, so this appears to have several advantages. Kalman pointed out that it is more satisfying to deal with the characteristic function formulation since then M_0, Q_k, R_k need not be assumed to be positive-definite.

Before proceeding with the proof some characteristic function relations shall be stated.

(i) The characteristic function ϕ and the probability density function p associated with a random variable $\underline{x}$ form a Fourier transform pair

$$\phi(\underline{s}) \triangleq E[\exp(i\underline{s}^T\underline{x})] = \int \exp(i\underline{s}^T\underline{x}) p(\underline{x})\, d\underline{x} \tag{2.28}$$

$$p(\underline{x}) \triangleq (2\pi)^{-n} \int \exp(-i\underline{s}^T\underline{x})\phi(\underline{s})\, d\underline{s} . \tag{2.29}$$

(ii) It is also useful to recognize that

$$(2\pi)^{-n} \int \exp[i\underline{s}^T\underline{x}]\, d\underline{x} = \delta(\underline{s}) , \tag{2.30}$$

where $\delta(\cdot)$ is the Dirac delta function.

(iii) It can be proven that

$$\int_{-\infty}^{\infty} \exp[\underline{u}^T\underline{x} - \underline{x}^T A\underline{x}]\, d\underline{x} = \pi^{n/2}|A|^{-1/2} \exp[\tfrac{1}{4}\underline{u}^T A^{-1}\underline{u}] \tag{2.31}$$

for any complex $\underline{u}$ and positive-definite A.

From the definitions it follows that the characteristic function for $\underline{x}_0$ is

$$\phi(\underline{s}_m) = \exp(i\underline{s}_m^T \underline{a} - \tfrac{1}{2}\underline{s}_m^T M_0 \underline{s}_m) . \tag{2.32}$$

For $p(\underline{z}_k|\underline{x}_k)$ and $p(\underline{x}_{k+1}|\underline{x}_k)$ the characteristic functions are

$$\phi(\underline{s}_v) = \exp(i\underline{s}_v^T H_k \underline{x}_k - \tfrac{1}{2}\underline{s}_v^T R_k \underline{s}_v) \tag{2.33}$$

$$\phi(\underline{s}_w) = \exp(i\underline{s}_w^T \Phi_{k+1,k}\underline{x}_k - \tfrac{1}{2}\underline{s}_w^T Q_k \underline{s}_w) . \tag{2.34}$$

To accomplish the proof, first observe that the characteristic function of $p(\underline{x}_k|\underline{Z}_k)$ is

$$\phi(\underline{s}_k) = \int \exp(i\underline{s}_k^T \underline{x}_k) p(\underline{x}_k|\underline{Z}_k)\, d\underline{x}_k$$

$$= \frac{1}{(2\pi)^{n+m} p(\underline{z}_k|\underline{Z}_{k-1})} \int \exp[-i(\underline{s}_{k/k-1} - \underline{s}_k)^T \underline{x}_k - i\underline{u}_v^T \underline{z}_k]\, p(\underline{x}_{k/k-1})\phi(\underline{s}_v)\, d(\underline{x}_k, \underline{z}_v, \underline{x}_{k/k-1}) . \tag{2.35}$$

Note also that the characteristic function of $p(x_k|Z_{k-1})$ is

$$\Phi(\underline{s}_{y/k-1}) = \int \exp[i\underline{s}_{k/k-1}^T x_k] p(x_k|Z_{k-1})\, dx_k$$

$$= \frac{1}{(2\pi)^{2n}} \int \exp[-i(s_w - s_{k/k-1})^T x_k - i\underline{s}_{k-1}^T x_{k-1}] \times$$

$$\times \Phi(s_w)\Phi(s_{k-1}) d(x_{k-1}, x_k, s_w, s_{k-1}) . \tag{2.36}$$

Using the above relations, the proof follows in a straightforward manner.

Proof: The initial conditions (2.26) and (2.27) will be established. Use (2.32) and (2.33) in (2.35). Then

$$\Phi(\underline{s}_0) = \frac{1}{(2\pi)^{n+m} p(Z_0)} \int \exp[-i(s_m - s_0 - H_0^T s_v)^T x_0 -$$

$$- i\underline{s}_v^T \underline{x}_0 + i\underline{s}_m^T \underline{e} - \tfrac{1}{2}\underline{s}_m^T M_0 s_m - \tfrac{1}{2}\underline{s}_v^T R_0 s_v] \, d(s_m, s_v, x_0) .$$

Integrate with respect to x_0 and use (2.30). Then one obtains

$$\Phi(s_0) = \frac{1}{(2\pi)^m p(Z_0)} \int \delta(H_0^T s_v + s_0 - s_m) \exp[-i\underline{s}_v^T x_0 + i\underline{s}_m^T e -$$

$$- \tfrac{1}{2}s_m^T M_0 s_m - \tfrac{1}{2}s_v^T R_0 s_v] d(s_m, s_v) .$$

Because of the delta function integration with respect to s_m is trivial, so this reduces to

$$\Phi(s_0) = \frac{1}{(2\pi)^m p(Z_0)} \exp[i\underline{s}_0^T e - \tfrac{1}{2}\underline{s}_0^T M_0 s_0] \exp\{\underline{s}_v^T[-i(z_0 - H_0 e) - H_0 M_0 s_0] -$$

$$- \tfrac{1}{2}s_v^T (H_0 M_0 H_0^T + R_0) s_v\} \, ds_v .$$

Using (2.31) and evaluating $p(z_0)$, it follows that

$$\Phi(s_0) = \exp\{i\underline{s}_0^T[e + K_0(z_0 - H_0 e)] - \tfrac{1}{2}s_0^T[M_0 - K_0 H_0 M_0] s_0\} . \tag{2.37}$$

But this is the characteristic function of a gaussian variable with mean and covariance described by (2.26) and (2.27).

To verify (2.22) and (2.23), assume that the lemma is true for t_{k-1} and form $\Phi(s_{k/k-1})$, using (2.36). This follows in a straightforward manner as does the proof of (2.21), (2.24), and (2.25), using (2.35). The details are omitted.

The use of characteristic functions has eliminated the requirement that the *a priori* distributions have nonsingular covariance matrices. In addition, the development leads directly to the equations presented in the theorem which have come to be known as the Kalman filter equations. As an alternative, it is possible to derive the density relations directly from (2.9) and (2.10). The direct application of these equations leads to an equivalent set of equations under the assumption that all of the *a priori* covariance matrices are nonsingular but which have some important differences. The density is, of course, still described by (2.20) and the mean values given by (2.21) and (2.22) are unchanged, as is the covariance P_k' described by (2.23). However, the covariance and gain take the following form

$$P_k^{-1} = (P_k')^{-1} + H_k^T R_k^{-1} H_k \tag{2.38}$$

$$K_k = P_k H_k^T R_k^{-1} . \tag{2.39}$$

Of importance here is the observation that the $n \times n$ matrix P_k must be inverted, rather than the $m \times m$ matrix $(H_k P_k' H_k^T + R_k)$ appearing in (2.24). Since m may be significantly smaller than n, this is an important computational consideration. It is shown in Section 2.4 of Chapter I that the two representations are related through the use of a matrix inversion lemma.

3. MAXIMUM *A POSTERIORI* AND MAXIMUM LIKELIHOOD ESTIMATES

In the preceding section, recursive relations for the *a posteriori* density function were developed which led in a natural way to the recursive Kalman filter equations when the system was assumed to be linear and gaussian. In this section the density functions will not be written in a recursive fashion and the considerations will lead to the maximum *a posteriori* estimation procedure. When the system is assumed to be linear, the Kalman equations are obtained again, as must be the case because of the Theorem 2.4. This discussion can still be considered to be Bayesian since the results are obtained using Bayes's rule to describe the *a posteriori* density. Again, the density relations will be non-recursive and this aspect constitutes the primary difference from the discussion of Section 2.

3.1 The Maximum *a posteriori* Estimation Problem

Consider the collection of all states X_k (i.e. $x_0, x_1, \ldots, x_k$) and all measurement data Z_k which are described in general by (1.20) and (1.21) and suppose that X_k and Z_k have the joint density $p(X_k, Z_k)$. Then, by Bayes's rule (2.4), it follows that

$$p(X_k|Z_k) = \frac{p(X_k, Z_k)}{p(Z_k)} \tag{3.1}$$

$$= \frac{p(Z_k|X_k)p(X_k)}{p(Z_k)} . \tag{3.2}$$

But the measurement noise is independent between samples, so that

$$p(Z_k|X_k) = p(z_0|x_0)p(z_1|x_1) \cdots p(z_k|x_k) = \prod_{i=0}^{k} p(z_i|x_i) \tag{3.3}$$

and (3.2) becomes

$$p(X_k|Z_k) = \frac{p(X_k) \prod_{i=0}^{k} p(z_i|x_i)}{p(Z_k)} . \tag{3.4}$$

Also, the definition of the plant and plant noise allows one to write

$$\begin{aligned} p(X_k) &= p(x_k|X_{k-1})p(X_{k-1}) \\ &= p(x_k|x_{k-1})p(x_{k-1}|x_{k-2}) \cdots p(x_1|x_0)p(x_0) \\ &= p(x_0) \prod_{i=1}^{k} p(x_i|x_{i-1}) , \end{aligned} \tag{3.5}$$

so that (3.4) can be rewritten as

$$p(X_k|Z_k) = \frac{p(x_0)\left[\prod_{i=1}^{k} p(x_i|x_{i-1}) \prod_{j=0}^{k} p(z_j|x_j)\right]}{p(Z_k)} . \tag{3.6}$$

As pointed out in Section 2, the densities $p(x_i|x_{i-1})$ and $p(z_i|x_i)$ are defined by the plant and measurement noise characteristics and the equations describing the two systems. If the noise sequences enter nonlinearly, it can be very difficult to define these conditional densities in a tractable form. To circumvent this problem, suppose that the noise enters additively and is gaussian, as described by (2.11) – (2.14). Then Equation (2.15) and Equation (2.16) apply and (3.6) can be written more explicitly. In particular, it is seen that

$$\begin{aligned} p(Z_k|X_k) &= C_v \exp\left\{-\tfrac{1}{2} \sum_{i=0}^{k} \left[z_i - h_i(x_i)\right]^T R_i^{-1} \left[z_i - h_i(x_i)\right]\right\} \\ &\geq C_v \exp\left\{-\tfrac{1}{2} \sum_{i=0}^{k} \| z_i - h_i(x_i) \|^2_{R_i^{-1}}\right\} , \end{aligned} \tag{3.7}$$

$$p(X_k) = p(x_0)\, C_w \exp\left\{-\tfrac{1}{2} \sum_{i=1}^{k} \| x_i - f_i(x_{i-1}) \|^2_{Q_{i-1}^{-1}}\right\} . \tag{3.8}$$

Assuming that x_0 is gaussian, (3.6) becomes

$$p(X_k|Z_k) = C \exp\left\{-\frac{1}{2}\left[\|x_0 - \bar{x}\|^2_{M_0^{-1}} + \sum_{i=1}^{k}\|x_i - f_{i-1}(x_{i-1})\|^2_{Q_{i-1}^{-1}} + \sum_{i=0}^{k}\|z_i - h_i(x_i)\|^2_{R_i^{-1}}\right]\right\}, \quad (3.9)$$

where C is a normalizing constant.

The *a posteriori* density $p(X_k|Z_k)$ contains the information required to obtain estimates of any or all of the states x_k. As discussed earlier, estimates for many criteria cannot actually be obtained because of the complications introduced by the nonlinearities. However, a reasonable estimate to try to obtain is the maximum *a posteriori* estimate described in Section 2. In this case one attempts to determine the mode (i.e. the maximum) of $p(X_k|Z_k)$. This is equivalent to choosing the estimates of x_k as those values which minimize the negative of the exponent of the *a posteriori* density. Let

$$L_k = \frac{1}{2}\|x_0 - \bar{x}\|^2_{M_0^{-1}} + \frac{1}{2}\sum_{i=1}^{k}\|x_i - f_{i-1}(x_{i-1})\|^2_{Q_{i-1}^{-1}} + \frac{1}{2}\sum_{i=0}^{k}\|z_i - h_i(x_i)\|^2_{R_i^{-1}}; \quad (3.10)$$

so

$$L_k = \log_e C - \log_e p(X_k|Z_k).$$

The L_k is to be minimized through the selection of the estimates. Equation (3.10) can be put into a form that is similar to that for optimal control problems so that the theory developed to solve those problems can be applied.

As stated, the estimation problem has been reduced to a deterministic minimization problem. Define variables u_{i-1} so that the plant can be considered as

$$x_k = f_k(x_{k-1}) + u_{k-1} \quad (3.11)$$

This equation is similar to (2.11) but the u_{k-1} have been introduced to emphasize that, unlike the w_{k-1}, they are not random variables. Now consider the estimation problem as the following.

Choose the sequences x_k and u_{k-1} so that the cost function

$$L_N = \frac{1}{2}\|x_0 - \bar{x}\|^2_{M_0^{-1}} + \frac{1}{2}\sum_{i=0}^{N}\|z_i - h_i(x_i)\|^2_{R_i^{-1}} + \frac{1}{2}\sum_{i=0}^{N-1}\|u_i\|_{Q_i^{-1}} \quad (3.12)$$

is minimized subject to the constraint

$$x_k = f_k(x_{k-1}) + u_{k-1}, \qquad k = 1, 2, \ldots, N-1. \quad (3.13)$$

One can attempt to solve this problem using the mathematical formalism of optimal control theory. It is not generally possible for arbitrary functions f_k and h_k but can be achieved when the system is linear. The general problem is beyond the scope of this chapter but the linear problem will be considered below.

It should be observed that the problem that has been posed is identical to the deterministic least-squares problem formulated in Section 1. This indicates that this deterministic problem has been imbedded in a probabilistic framework in which the errors and uncertainties have been assumed to be gaussian random variables and white-noise sequences. Thus, the deterministic and probabilistic problems are not fundamentally different although the language and analysis procedures are very dissimilar. This formulation has been suggested by Cox[15] for time-discrete systems. Cox[16] also considered time-continuous systems in which the summations of (3.12) are replaced by integrations and the difference Equation (3.13) is replaced by a differential equation. Detchmendy and Sridhar[17] also considered the nonlinear estimation problem in this manner and derived a system of equations that are similar to those of Cox. More recently, Mortensen[18] has considered the time-continuous problem and Gura and Henriksen[19] have considered the generalized least-squares procedure of Chapter 1 and have obtaind and extended the results of Detchmendy and Sridhar.

3.2 Maximum Likelihood Estimation Theory

The maximum likelihood procedure that has been widely used is closely related to the maximum *a posteriori* estimates of the preceding section. To discuss the difference, consider Equation (3.2). The denominator $p(Z_k)$ is a normalizing constant and can be ignored. In the maximum *a posteriori* procedure, the estimate is determined that maximizes $p(X_k|Z_k)$. Suppose that the *a priori* information relating to the x_k, as described by $p(X_k)$, does not suggest that any state is more likely than any other. Then, the $p(X_k)$ will be a uniform density and the maximization of $p(X_k|Z_k)$ will actually be determined by the maximization of $p(Z_k|X_k)$. The values of x_k which are found to maximize $p(Z_k|X_k)$ are referred to as *maximum likelihood estimates* and $p(Z_k|X_k)$ is referred to as the *likelihood function*.

Maximum likelihood estimates have several properties which are desirable and have therefore seen widespread application. For the discussion immediately below, some of these aspects are considered and will be related to some of the properties of the Kalman filter. For a more complete discussion the reader is referred to Cramer[20] or van Trees[4].

The covariance of the error in any estimate can be bounded below through a relatively straightforward application of the Schwarz inequality. This bound is usually referred to as the Cramer-Rao inequality. In the following the Cramer-Rao inequality is proved for unbiased (see van Trees[4] for a treatment for biased estimates) estimates of an unknown vector parameter and then for a random parameter. Suppose that it is desired to estimate an unknown parameter $\underline{x}$ from measurement data $\underline{Z}_N$. The parameter and the data are assumed to be related according to

$$\underline{z}_k = \underline{h}_k(\underline{x}, \underline{y}_k) , \qquad k = 1, 2, \ldots, N , \tag{3.14}$$

where the noise $\underline{y}_k$ has a known distribution. As a result, the conditional density $p(\underline{Z}_N|\underline{x})$ can be determined.

Theorem 3.1 (Cramer-Rao Inequality)

If $\hat{\underline{x}}$ is any unbiased estimate of $\underline{x}$ based on the measurement data $\underline{Z}_N$, then the conditional covariance of the error in the estimate given $\underline{x}$ is bounded below by the inverse, assuming it exists, of the *Fisher information matrix* J,

$$E[(\underline{x} - \hat{\underline{x}})(\underline{x} - \hat{\underline{x}})^T|\underline{x}] \geq J^{-1} , \tag{3.15}$$

where

$$J = E\left\{\left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right]^T \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right] |\underline{x}\right\} \tag{3.16}$$

$$= -E\left\{\frac{\partial}{\partial \underline{x}}\left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right]^T |\underline{x}\right\} . \tag{3.17}$$

Equality holds in (3.15) if, and only if,

$$\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x}) = k(\underline{x})(\hat{\underline{x}} - \underline{x})^T . \tag{3.18}$$

It is also assumed that

$$\frac{\partial p}{\partial \underline{x}}(\underline{Z}_N|\underline{x})$$

and

$$\frac{\partial}{\partial \underline{x}}\left[\frac{\partial}{\partial \underline{x}} p(\underline{Z}_N|\underline{x})\right]^T$$

exist and are absolutely integrable.

Proof: First, consider that

$$E[(\hat{\underline{x}} - \underline{x})|\underline{x}] = \int (\hat{\underline{x}} - \underline{x}) p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N = 0 ,$$

since $\hat{\underline{x}}$ is unbiased. Differentiate both sides with respect to $\underline{x}$ to obtain

$$0 = \int \frac{\partial}{\partial \underline{x}} [(\hat{\underline{x}} - \underline{x}) p(\underline{Z}_N|\underline{x})]\, d\underline{Z}_N$$

$$= \left\{-\int p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N\right\} I + \int (\hat{\underline{x}} - \underline{x}) \frac{\partial p(\underline{Z}_N|\underline{x})}{\partial \underline{x}}\, d\underline{Z}_N .$$

But observe that

$$\frac{\partial p(\underline{Z}_N|\underline{x})}{\partial \underline{x}} = \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right] p(\underline{Z}_N|\underline{x}) ,$$

so one obtains

$$I = \int (\hat{\underline{x}} - \underline{x}) \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right] p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N$$

$$= \int \left\{(\hat{\underline{x}} - \underline{x})[p(\underline{Z}_N|\underline{x})]^{1/2}\right\} \left\{\left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right]\left[p(\underline{Z}_N|\underline{x})\right]^{1/2}\right\} d\underline{Z}_N .$$

Applying the Schwarz inequality, it follows that*

$$I \leq \left(\int (\hat{\underline{x}} - \underline{x})(\hat{\underline{x}} - \underline{x})^T p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N\right) J , \tag{3.19}$$

* The matrix inequality $A \leq B$ means that $(B-A)$ is non-negative-definite.

where J is defined by (3.16). If J^{-1} exists, then (3.15) is proven. It remains to demonstrate the second form for the Fisher information matrix. Equality holds in (3.19) if and only if

$$\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x}) = k(\underline{x})(\underline{\hat{x}} - \underline{x}) .$$

By definition, it is true that

$$\int p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N = 1 .$$

Differentiate with respect to $\underline{x}$. Then one obtains

$$\int \frac{\partial}{\partial \underline{x}} p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N = \int \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right] p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N = 0 .$$

Differentiating a second time with respect to $\underline{x}$, it is found that

$$\frac{\partial}{\partial \underline{x}} \int \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right]^T p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N = \int \frac{\partial}{\partial \underline{x}} \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right]^T p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N +$$

$$+ \int \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right]^T \frac{\partial}{\partial \underline{x}} p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N$$

or

$$0 = \int \frac{\partial}{\partial \underline{x}} \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right]^T p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N + \int \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right]^T \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right] p(\underline{Z}_N|\underline{x})\, d\underline{Z}_N .$$

Equation (3.17) follows immediately and the proof is complete. Note that the conditions on the partial derivatives are required so that the interchanges of differentiation and integration are valid.

This theorem and the Fisher information matrix can be applied to the linear filtering problem and some interesting relationships are obtained. Consider the following examples.

Example 1: Suppose that the measurement data are linearly related to the parameter $\underline{x}$,

$$\underline{Z}_k = H_k \underline{x} + \underline{Y}_k , \tag{3.20}$$

and assume that the noise is gaussian and samples are independent,

$$p(\underline{Y}_k) = k_y \exp\left(-\tfrac{1}{2} \underline{Y}_k^T R_k^{-1} \underline{Y}_k\right) . \tag{3.21}$$

Thus, the conditional density $p(\underline{Z}_N|\underline{x})$ is

$$p(\underline{Z}_N|\underline{x}) = k \exp\left\{-\tfrac{1}{2} \sum_{i=0}^{N} (\underline{Z}_i - H_i\underline{x})^T R_i^{-1} (\underline{Z}_i - H_i\underline{x})\right\} ,$$

where k is the appropriate normalization constant. It follows that

$$\log_e p(\underline{Z}_N|\underline{x}) = \log_e k - \tfrac{1}{2} \sum_{i=0}^{N} (\underline{Z}_i - H_i\underline{x})^T R_i^{-1} (\underline{Z}_i - H_i\underline{x}) ;$$

so

$$\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x}) = \sum_{i=0}^{N} (\underline{Z}_i - H_i\underline{x})^T R_i^{-1} H_i$$

and

$$\frac{\partial}{\partial \underline{x}} \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{Z}_N|\underline{x})\right]^T = -\sum_{i=0}^{N} H_i^T R_i^{-1} H_i .$$

Using (3.17), the Fisher information matrix is found to be

$$J = \sum_{i=0}^{N} H_i^T R_i^{-1} H_i \tag{3.22}$$

and the covariance of the error in any unbiased estimate of $\underline{x}$ is bounded by

$$E[(\underline{x} - \underline{\hat{x}})(\underline{x} - \underline{\hat{x}})^T] \geq \left[\sum_{i=0}^{N} H_i^T R_i^{-1} H_i\right]^{-1} \tag{3.23}$$

if J^{-1} exists.

Example 2: As a minor variation on the preceding example, suppose that the parameter $\underline{x}$ is time-varying but satisfies the constraint that

$$\underline{x}_k = \Phi_{k,k-1}\underline{x}_{k-1}, \qquad k = 1, 2, \ldots, N. \tag{3.24}$$

The measurements are described by

$$\underline{z}_k = H_k\underline{x}_k + \underline{v}_k. \tag{3.25}$$

Solving (3.24) for the initial state $\underline{x}_0$, one has

$$\underline{x}_k = \Phi_{k,0}\underline{x}_0, \tag{3.26}$$

where $$\Phi_{k,0} = \Phi_{k,k-1}\Phi_{k-1,k-2}\cdots\Phi_{1,0}.$$

Assuming that $\Phi_{k,k-1}$ is nonsingular for all k, the measurement equation becomes

$$\underline{z}_k = H_k\Phi_{k,0}\underline{x}_0 + \underline{v}_k. \tag{3.27}$$

Equation (3.27) has essentially been reduced to the same form as (3.20), so it follows that the Fisher information matrix is

$$J = \sum_{i=0}^{N} \Phi_{i,0}^T H_i^T R_i^{-1} H_i \Phi_{i,0} \tag{3.28}$$

and the error covariance is bounded by

$$E[(\underline{x}-\hat{\underline{x}})(\underline{x}-\hat{\underline{x}})^T] \geq \left[\sum_{i=0}^{N} \Phi_{i,0}^T H_i^T R_i^{-1} H_i \Phi_{i,0}\right]^{-1}. \tag{3.29}$$

This example is important because it shows that the Fisher information matrix is essentially identical with the observability matrix that was introduced in Section 3.1 of Chapter 1. Furthermore, the Cramer-Rao inequality (3.29) provides a lower bound that is very similar to the bound presented in Section 3.3 of Chapter 1.

Any estimate for which equality holds in (3.15) is said to be an *efficient* estimator. It has been seen that equality occurs if and only if

$$\frac{\partial}{\partial \underline{x}} \log_e p(Z_N|\underline{x}) = [\hat{\underline{x}} - \underline{x}]k(\underline{x}).$$

It is easily shown that, if this condition can be satisfied, it can be accomplished by a maximum likelihood estimate.

The maximization of $p(Z_N|\underline{x})$ can be replaced by the maximization of $\log_e p(Z_N|\underline{x})$. Then a necessary condition that an estimate $\hat{\underline{x}}_{ml}$ maximize $\log_e p(Z_N|\underline{x})$ is

$$\frac{\partial}{\partial \underline{x}} \log_e p(Z_N|\underline{x})\Big|_{\underline{x}=\hat{\underline{x}}_{ml}} = 0 = (\hat{\underline{x}} - \underline{x})k(\underline{x})\Big|_{\underline{x}=\hat{\underline{x}}_{ml}}. \tag{3.30}$$

Thus, for the last relation to equal zero, either

$$\hat{\underline{x}} = \hat{\underline{x}}_{ml} \tag{3.31}$$

or $$k(\hat{\underline{x}}_{ml}) = 0.$$

Condition (3.31) is selected since it provides an estimate that depends upon the data; so, if an efficient estimate exists, it is a maximum likelihood estimate. If an efficient estimate does not exist, then there is no measure of the accuracy of the estimate. This constitutes a major drawback of this approach or maximum *a posteriori* estimates. Thus there may be unbiased estimates which yield a "smaller" error covariance when there is no efficient estimate.

In stating Theorem 3.1, it was assumed that an unknown parameter $\underline{x}$ was to be treated. If $\underline{x}$ is treated as a random variable, a similar result can be derived which yields insight into the existence of efficient estimators. This development follows van Trees[4].

Theorem 3.2

Suppose $\underline{x}$ and Z_N have joint density $p(\underline{x}, Z_N)$ and let $\hat{\underline{x}}$ be any unbiased estimate of $\underline{x}$ based on the measurement data Z_N. Then, assuming L^{-1} exists, the covariance of the error in the estimate is bounded below by

$$E[(\underline{x}-\hat{\underline{x}})(\underline{x}-\hat{\underline{x}})^T] \geq L^{-1}, \tag{3.32}$$

where
$$L = \left\{\left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{x}, Z_N)\right]^T \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{x}, Z_N)\right]\right\} \tag{3.33}$$

$$= -E\left\{\frac{\partial}{\partial \underline{x}}\left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{x}, Z_N)\right]^T\right\}. \tag{3.34}$$

Equality holds if and only if

$$\frac{\partial}{\partial \underline{x}} \log_e p(\underline{x}, Z_N) = k(\hat{\underline{x}} - \underline{x})^T . \tag{3.35}$$

The first and second partial derivatives

$$\frac{\partial p}{\partial \underline{x}} (\underline{x}, Z_N) \qquad \text{and} \qquad \frac{\partial}{\partial \underline{x}}\left[\frac{\partial}{\partial \underline{x}} p(\underline{x}, Z_N)\right]$$

are assumed to exist and are absolutely integrable with respect to $\underline{x}$ and Z_N. Also, it is assumed that

$$\lim_{\underline{x} \to \infty} \underline{b}(\underline{x}) p(\underline{x}) = 0 \tag{3.36}$$

$$\lim_{\underline{x} \to \infty} \underline{b}(\underline{x}) p(\underline{x}) = 0 , \tag{3.37}$$

where
$$\underline{b}(\underline{x}) \triangleq \int [\hat{\underline{x}} - \underline{x}] p(Z_N|\underline{x})\, dZ_N . \tag{3.38}$$

Proof: The proof is similar to that of Theorem 3.1, although expectations with respect to $\underline{x}$ must be considered also. First form $p(\underline{x})\underline{b}(\underline{x})$ and differentiate with respect to $\underline{x}$. This yields

$$\frac{d}{d\underline{x}} p(\underline{x})\underline{b}(\underline{x}) = -\int p(\underline{x}, Z_N)\, dZ_N I + \int (\hat{\underline{x}} - \underline{x}) \frac{\partial}{\partial \underline{x}} p(\underline{x}, Z_N)\, dZ_N .$$

Integrate with respect to $\underline{x}$ and invoke conditions (3.36) and (3.37) to obtain

$$0 = -I + \int (\hat{\underline{x}} - \underline{x}) \left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{x}, Z_N)\right] p(\underline{x}, Z_N) d(\underline{x}, Z_N) .$$

Application of the Schwarz inequality yields

$$I \leq [E((\hat{\underline{x}} - \underline{x})(\hat{\underline{x}} - \underline{x})^T)]\, L ,$$

where L is given by (3.33) and equality holds if and only if (3.35) is satisfied. Note that the averaging with respect to both $\underline{x}$ and Z_N implies that, unlike the condition of Theorem 3.1, k is a constant independent of both $\underline{x}$ and Z_N.

The remainder of the proof is identical with that of Theorem 3.1.

Two results relating to the effect and consequence of treating $\underline{x}$ as a random parameter are worth noting. First, the matrix L as given by (3.34) can be rewritten as

$$L = J - E\left\{\frac{\partial}{\partial \underline{x}}\left[\frac{\partial}{\partial \underline{x}} \log_e p(\underline{x})\right]^T\right\}, \tag{3.39}$$

where J is defined in Theorem 3.1. Thus, the probabilistic description of $\underline{x}$ enters independently of the noise statistics. This will be discussed further below.

The condition (3.35) for equality in (3.32) involves the constant k. Note that the condition can be written as

$$\frac{\partial}{\partial \underline{x}}[\log_e p(\underline{x}|Z_N)] = k(\hat{\underline{x}} - \underline{x})^T .$$

Integrate with respect to $\underline{x}$ and take the antilog to obtain

$$p(\underline{x}|Z_N) = \exp\left[-\frac{k}{2} \underline{x}^T \underline{x} + k\underline{x}^T \underline{x} + c\right] . \tag{3.40}$$

But this implies that the *a posteriori* density $p(\underline{x}|Z_N)$ must be *gaussian* for an efficient estimate to exist.

It follows easily that a maximum *a posteriori* estimate will be efficient if an efficient estimate exists. However, since $p(\underline{x}|Z_N)$ must be gaussian in this case, the minimum variance estimator yields the same result, so it is also efficient.

Consider these results through the following example.

Example 3: Consider the system described in Example 2 with the additional requirement that $\mathbf{x}_0$ is a gaussian random variable with density function

$$p(\mathbf{x}_0) = k_0 \exp\left(-\tfrac{1}{2}(\mathbf{x}_0 - \bar{\mathbf{x}})^T M_0^{-1}(\mathbf{x}_0 - \bar{\mathbf{x}})\right) . \tag{3.41}$$

Then, the information matrix L, as given by (3.38), is

$$L = J + \tfrac{1}{2} E\left\{\frac{\partial}{\partial \mathbf{x}}\left[\frac{\partial}{\partial \mathbf{x}}(\mathbf{x}_0 - \bar{\mathbf{x}})^T M_0^{-1}(\mathbf{x}_0 - \bar{\mathbf{x}})\right]\right\}$$

$$= \sum_{i=1}^{N} \Phi_{i,0}^T H_i^T R_i^{-1} H_i \Phi_{i,0} + M_0^{-1}$$

and the error covariance is bounded below by

$$E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T] \geq \left[M_0^{-1} + \sum_{i=1}^{N} \Phi_{i,0}^T H_i^T R_i^{-1} H_i \Phi_{i,0}\right]^{-1} .$$

But this is essentially the bound discussed in Section 3.3 of Chapter 1. The right-hand side represents the error covariance given by the Kalman filter for this problem and provides another verification that the minimum variance estimator is efficient for this system. To obtain precise agreement, note that the estimate of the terminal state $\mathbf{x}_N$ is given by

$$\hat{\mathbf{x}}_N = \Phi_{N,0}\hat{\mathbf{x}}_0 .$$

It follows that the error covariance associated with $\hat{\mathbf{x}}_N$ is

$$E[(\mathbf{x}_N - \hat{\mathbf{x}}_N)(\mathbf{x}_N - \hat{\mathbf{x}}_N)^T] = \Phi_{N,0} E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^T] \Phi_{N,0}^T .$$

Using this relation, one obtains the result given in Section 3.3.2 of Chapter 1.

This completes the discussion of error bounds using the Cramer-Rao inequality and its generalizations.

3.3 Maximum *a posteriori* Estimates for Linear Systems

Suppose that the plant and measurement equations are linear, so that the maximum *a posteriori* estimation problem becomes that of choosing the sequences $\mathbf{x}_N$ to minimize the cost function

$$L_N = \tfrac{1}{2}\| \mathbf{x}_0 - \bar{\mathbf{x}} \|^2_{M_0^{-1}} + \tfrac{1}{2}\sum_{i=0}^{N} \| \mathbf{z}_i - H_i\mathbf{x}_i \|^2_{R_i^{-1}} + \tfrac{1}{2}\sum_{i=0}^{N-1} \| \mathbf{w}_i \|^2_{Q_i^{-1}} , \tag{3.42}$$

subject to the constraint

$$\mathbf{x}_k = \Phi_{k,k-1}\mathbf{x}_{k-1} + \mathbf{w}_{k-1} . \tag{3.43}$$

The minimization is accomplished inductively in the following manner. Consider the cost function after the first measurement has been obtained,

$$L_0 = \tfrac{1}{2}\| \mathbf{x}_0 - \bar{\mathbf{x}} \|^2_{M_0^{-1}} + \tfrac{1}{2}\| \mathbf{z}_0 - H_0\mathbf{x}_0 \|^2_{R_0^{-1}} . \tag{3.44}$$

As will be seen, this simple problem contains within it a major portion of the problem, since a recursive solution is to be obtained.

Choose the $\mathbf{x}_0$ that minimizes L_0 and call it $\hat{\mathbf{x}}_0$. A necessary and sufficient condition for this problem is that

$$\left(\frac{\partial L_0}{\partial \mathbf{x}_0}\right)^T = 0 = M_0^{-1}(\hat{\mathbf{x}}_0 - \bar{\mathbf{x}}) - H_0^T R_0^{-1}(\mathbf{z}_0 - H_0\hat{\mathbf{x}}_0) . \tag{3.45}$$

Solve for $\hat{\mathbf{x}}_0$ to obtain

$$(M_0^{-1} + H_0^T R_0^{-1} H_0)\hat{\mathbf{x}}_0 = M_0^{-1}\bar{\mathbf{x}} + H_0^T R_0^{-1}\mathbf{z}_0 \tag{3.46}$$

or

$$\hat{\mathbf{x}}_0 = (M_0^{-1} + H_0^T R_0^{-1} H_0)^{-1}\left[M_0^{-1}\bar{\mathbf{x}} + H_0^T R_0^{-1}\mathbf{z}_0\right] . \tag{3.47}$$

Let

$$P_0 \triangleq (M_0^{-1} + H_0^T R_0^{-1} H_0)^{-1} \tag{3.48a}$$

$$= M_0 - M_0 H_0^T (H_0 M_0 H_0^T + R_0)^{-1} H_0 M_0 , \tag{3.48b}$$

using the matrix inversion lemma.

Using (3.48b), (3.47) becomes

$$\hat{x}_0 = \bar{x} + k_0[z_0 - H_0\bar{x}] , \tag{3.49}$$

where

$$K_0 \triangleq M_0H_0^T(H_0M_0H_0^T + R_0)^{-1} . \tag{3.50}$$

Thus the maximum *a posteriori* estimate of x_0, given $\bar{x}$ and z_0, is described by (3.49). As must be true, this agrees with the estimate derived in Section 2 using the Bayesian recursion relations.

Suppose now that there are two measurements and the cost function L_1 is given by

$$L_1 = \tfrac{1}{2} \| x_0 - \bar{x} \|^2_{M_0^{-1}} + \tfrac{1}{2} \sum_{i=0}^{1} \| z_i - H_i x_i \|^2_{R_i^{-1}} + \tfrac{1}{2} \| w_0 \|^2_{Q_0^{-1}} , \tag{3.51}$$

where the x_0, x_1 and w_0 are related by

$$x_1 = \Phi_{1,0}x_0 + w_0 .$$

The estimate $\hat{x}_0$ that was derived above will be different than that obtained by minimizing (3.51), since two measurements are involved. In fact, the estimate of x_0 that is derived by minimizing (3.51) is a *smoothed* estimate rather than the filtered estimate given by (3.49). To obtain the desired recursion relation, consider L_0 and express it relative to $\hat{x}_0$.

$$L_0 = \tfrac{1}{2} \| x_0 - \bar{x} \|^2_{M_0^{-1}} + \tfrac{1}{2} \| z_0 - H_0x_0 \|^2_{R_0^{-1}} . \tag{3.52}$$

Expand in a Taylor series, which is given exactly by

$$L_0 = \tfrac{1}{2} \| \hat{x}_0 - \bar{x} \|^2_{M_0^{-1}} + \tfrac{1}{2} \| z_0 - H_0\hat{x}_0 \|^2_{R_0^{-1}} + \left.\frac{\partial L_0}{\partial x_0}\right|_{x_0=\hat{x}_0} (x_0 - \hat{x}_0) +$$

$$+ \tfrac{1}{2} (x_0 - \hat{x}_0)^T \left(\left.\frac{\partial^2 L_0}{\partial x_0^2}\right)\right|_{x_0=\hat{x}_0} (x_0 - \hat{x}_0) . \tag{3.53}$$

But

$$\left.\frac{\partial L_0}{\partial x_0}\right|_{x_0=\hat{x}_0} = 0$$

and

$$\frac{\partial^2 L_0}{\partial x_0^2} = \frac{\partial}{\partial x} [M_0^{-1}(x_0 - \bar{x}) - H_0^TR_0^{-1}(z_0 - H_0x_0)]$$

$$= M_0^{-1} + H_0^TR_0^{-1}H_0 ;$$

so

$$L_0 = \tfrac{1}{2} \| \hat{x}_0 - \bar{x} \|^2_{M_0^{-1}} + \tfrac{1}{2} \| z_0 - H_0\hat{x}_0 \|^2_{R_0^{-1}} + \tfrac{1}{2} \| x_0 - \hat{x}_0 \|^2_{P_0^{-1}}$$

$$\triangleq C_0 + \tfrac{1}{2} \| x_0 - \hat{x}_0 \|^2_{P_0^{-1}} , \tag{3.54}$$

where C_0 is independent of x_0 and therefore will not have an influence on subsequent discussion.

The L_1 is now given by

$$L_1 = C_0 + \tfrac{1}{2} \| x_0 - \hat{x}_0 \|^2_{P_0^{-1}} + \| z_1 - H_1x_1 \|^2_{R_1^{-1}} + \tfrac{1}{2} \| w_0 \|^2_{Q_0^{-1}} . \tag{3.55}$$

To eliminate x_0, introduce the plant constraint. Assuming that the inverse of $\Phi_{1,0}$ exists and is $\Phi_{0,1}$, L_1 becomes

$$L_1 = C_0 + \tfrac{1}{2} \| \Phi_{0,1}(x_1 - w_0) - \hat{x}_0 \|^2_{P_0^{-1}} + \| z_1 - H_1x_1 \|^2_{R_1^{-1}} + \tfrac{1}{2} \| w_0 \|^2_{Q_0^{-1}} .$$

Now, choose the w_0 to minimize L_1.

$$\left(\frac{\partial L_1}{\partial w_0}\right)^T = 0 = -\Phi_{0,1}^TP_0^{-1}[\Phi_{0,1}(x_1 - \hat{w}_0) - \hat{x}_0] + Q_0^{-1}\hat{w}_0 . \tag{3.56}$$

Solving for $\hat{w}_0$, one obtains

$$\hat{w}_0 = [I - \Phi_{1,0}P_0\Phi_{1,0}^T(\Phi_{1,0}P_0\Phi_{1,0}^T + Q_0)]^{-1}[x_1 - \Phi_{1,0}\hat{x}_0] . \tag{3.57}$$

Let $$P_1' \triangleq \Phi_{1,0} P_0 \Phi_{1,0}^T + Q_0 \tag{3.58}$$

and note that $$\Phi_{1,0} P_0 \Phi_{1,0}^T P_0'^{-1} = I - Q_0 P_0'^{-1} .$$

It follows from (3.57) and (3.58) that

$$\begin{aligned} \tfrac{1}{2} \| \Phi_{0,1}(\hat{x}_1 - \hat{w}_0) - \hat{x}_0 \|^2_{P_0^{-1}} + \tfrac{1}{2} \| \hat{w}_0 \|^2_{Q_0^{-1}} &= \tfrac{1}{2} \| \hat{x}_1 - \Phi_{1,0}\hat{x}_0 - \hat{w}_0 \|^2_{\Phi_{0,1}^T P_0^{-1} \Phi_{0,1}} + \tfrac{1}{2} \| \hat{w}_0 \|^2_{Q_0^{-1}} \\ &= \tfrac{1}{2} \| \hat{x}_1 - \Phi_{1,0}\hat{x}_0 \|^2_{P_1'^{-1} Q_0 P_1'^{-1}} + \tfrac{1}{2} \| \hat{x}_1 - \Phi_{1,0}\hat{x}_0 \|^2_{P_1'^{-1} Q_0 P_1'^{-1}} \\ &= \tfrac{1}{2} \| \hat{x}_1 - \Phi_{1,0}\hat{x}_0 \|^2_{P_1'^{-1}} . \end{aligned} \tag{3.59}$$

With this relation, the cost function assumes the form

$$L_1 = C_0 + \tfrac{1}{2} \| \hat{x}_1 - \Phi_{1,0}\hat{x}_0 \|^2_{P_1'^{-1}} + \tfrac{1}{2} \| z_1 - H_1 \hat{x}_1 \|^2_{R_1^{-1}} . \tag{3.60}$$

Note that, in the absence of a new measurement, the best predicted estimate $\hat{x}_1'$ is

$$\hat{x}_1' = \Phi_{1,0}\hat{x}_0 . \tag{3.61}$$

which corresponds with the anticipated result.

It can be seen that (3.60) has essentially the same form as the L_0 described by (3.44), since C_0 does not depend upon $\hat{x}_1$. Thus minimization of L_1 with respect to $\hat{x}_1$ will yield the result that was obtained above, except for the notation change. In particular, it follows that

$$\hat{x}_1 = \Phi_{1,0}\hat{x}_0 + K_1 [z_1 - H_1 \Phi_{1,0}\hat{x}_0] , \tag{3.62}$$

where $$K_1 = P_1' H_1^T (H_1 P_1' H_1^T + R_1)^{-1} . \tag{3.63}$$

Also, the cost L_1 is seen to be given by

$$L_1 = C_1 + \| \hat{x}_1 - \hat{x}_1' \|^2_{P_1'^{-1}} , \tag{3.64}$$

where $$P_1 = (P_1'^{-1} + H_1^T R_1^{-1} H_1)^{-1} \tag{3.65}$$

$$= P_1' - P_1' H_1^T (H_1 P_1' H_1^T + R_1)^{-1} H_1 P_1'$$

$$= P_1' - K_1 H_1 P_1' . \tag{3.66}$$

The generalization to an arbitrary stage follows inductively and is accomplished by using the same arguments as above. Thus it follows that the recursive, maximum *a posteriori* estimate is described by the system

$$\hat{x}_k = \Phi_{k,k-1}\hat{x}_{k-1} + K_k [z_k - H_k \Phi_{k,k-1}\hat{x}_{k-1}] , \tag{3.67}$$

where $$K_k = P_k' H_k^T (H_k P_k' H_k^T + R_k)^{-1} \tag{3.68}$$

and $$P_k' = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + Q_{k-1} \tag{3.69}$$

$$P_k = P_k' - K_k H_k P_k' . \tag{3.70}$$

Once again, the Kalman filter equations have been derived.

4. SUMMARY OF RESULTS

The filtering problem for time-discrete, nonlinear, stochastic systems has been considered within the general probabilistic framework provided by the application of Bayes's rule. This Bayesian approach provides insight into the general character of the behavior of the *a posteriori* density function for nonlinear systems. The development has been designed to provide an understanding of the difficulties arising in nonlinear filtering theory and provides a structure within which the nonlinear problem can be attacked. Further, the Bayesian approach leads naturally to the discussion of the maximum *a posteriori* and maximum likelihood estimation procedures. The former procedure is shown to be identical with deterministic least-squares when the plant and measurement noise sequences are assumed to be additive and gaussian. The recursive Bayesian and maximum *a posteriori* approaches are seen to yield the same filtering equations as the unbiased, minumum variance estimates when the plant and measurement systems are linear with gaussian noise.

The principal results of this presentation are stated in the following paragraphs. The general problem is stated in Section 1.2.

Section 2.1: The conditional mean $E[\underline{x}|\underline{Z}]$ is shown in Theorem 2.1 to provide the unbiased, minimum variance estimate of a random variable $\underline{x}$ from measurement data $\underline{Z}$. Thus, knowledge of the *a posteriori* density $p(\underline{x}|\underline{Z})$ would enable this estimate to be determined, as well as virtually any other type of estimate that might be desired.

It is shown in Theorem 2.2 and in Theorem 2.3 that a significant number of estimation criteria will lead to the same estimate as the minimum variance criterion. This is important because it indicates that the choice of the minimum variance criterion because of its analytical tractability has broader application than might be otherwise appreciated.

Section 2.2: Recursion relations which describe the manner in which the *a posteriori* density function changes as new data become available are presented as Theorem 2.4. Some of the principal difficulties that are encountered in applying these results to nonlinear systems are discussed.

Section 2.3: The general results of Section 2.3 are applied to linear systems with gaussian noise. Theorem 2.5 shows that this system has a gaussian *a posteriori* density and that the Kalman filter equations describe the conditional mean and covariance of $p(\underline{X}_k|\underline{Z}_k)$. It is pointed out that it is more convenient, at least in this case, to work with the characteristic function rather than the density itself to derive the desired results.

Section 3.1: The maximum *a posteriori* estimation procedure is presented essentially as a nonrecursive version of the Bayesian approach described in Section 2. This formulation is seen to be identical with the deterministic least-squares problem discussed in Section 1.1 when the plant and measurement noise sequences are assumed explicitly to be additive and gaussian.

Section 3.2: Some aspects of the well-known maximum likelihood procedure are discussed in this section. It is seen that this approach differs from the maximum *a posteriori* through the neglect of the *a priori* distribution assigned to the parameters to be estimated. The two are identical if $p(\underline{X}_k)$ is uniform. Then the maximum *a posteriori* procedure which calls for the determination of the $\underline{X}_k$ that maximizes $p(\underline{X}_k|\underline{Z}_k)$ is identical with the maximum likelihood procedure in which $p(\underline{Z}_k|\underline{X}_k)$ is maximized through the choice of $\underline{X}_k$.

The Cramer-Rao inequality is presented in Theorem 3.1. This inequality provides a lower bound for the error covariance matrix for any unbiased estimate of an unknown parameter. This result is generalized in Theorem 3.2 to the case where the parameter is a random variable with prescribed distribution. The lower bound is the Fisher information matrix. For a linear system with gaussian measurement noise and no plant noise, the Cramer-Rao inequality is shown to provide a lower bound that is identical with a bound derived in Section 3.3.3 of Chapter 1. Also, the Fisher information matrix is shown to be identical to the observability matrix, introduced in Section 3.1 of Chapter 1.

Section 3.3: The determination of the states $\underline{x}_k$ that maximize $p(\underline{x}_k|\underline{Z}_k)$ is shown to again yield the Kalman filter equations. The procedure that is used to accomplish the minimization makes use of the desire to obtain a recursive solution and promises to provide straightforward application to nonlinear systems.

The preceding discussion provides a brief description of the contents of Sections 2 and 3. The reader is directed to the references for more complete treatments of many aspects that have only been touched superficially.

REFERENCES

1. Fel'dbaum, A.A. — *Optimal Control Systems.* (Translated by A.Kraiman). Academic Press, New York, 1965.

2. Aoki, M. — *Optimization of Stochastic Systems: Topics in Discrete-Time Stochastic and Adaptive Systems Optimization.* Academic Press, New York, 1966.

3. Sorenson, H.W. — *A Nonlinear Perturbation Theory for Estimation and Control of Time-Discrete Stochastic Systems.* PhD Dissertation, Department of Engineering, University of California, Los Angeles (also UCLA Engineering Report 68-2), 1968.

4. van Trees, H.L. — *Detection, Estimation, and Modulation Theory.* Part I. John Wiley, New York, 1968.

5. Ho, Y.C. — *The Method of Least-Squares and Optimal Filtering Theory.* Rand Corporation Memo RM-3329-PR, 1962.

6. Ito, K. — *On Stochastic Processes.* Tata Institute for Fundamental Research, Bombay, India, 1961.

7. Stratonovich, R.L. — *Topics in the Theory of Random Noise.* (Translated by R.A.Silverman). Gordon and Breach, New York, 1963.

8. Papoulis, A. — *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, New York, 1965.

9. Sherman, S. — *Non-Mean-Square Error Criterion.* Institute of Radio Engineers, Transactions on Information Theory, IT-4, 1958, pp.125-126.

10. Sorenson, H.W. — *Nonlinear Filtering by Approximation of the a posteriori Density.* International Journal of Control, Vol.8, No.1, pp.33-51.

11. Stratonovich, R.L. — *Conditional Markov Processes.* Theory of Probability and its Applications, Vol.5, No.2, 1960, pp.156-178.

12. Kushner, H.J. — *On the Dynamical Equations of Conditional Probability Density Functions with Applications to Optimal Stochastic Control Theory.* Journal of Mathematical Analysis and Applications, 1964, pp.332-344.

13. Bucy, R.S. — *Nonlinear Filtering Theory.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-10, No.2, April 1965, p.198.

14. Fisher J.R. — *Conditional Probability Density Functions and Optimal Nonlinear Estimation.* PhD Dissertation, University of California, Los Angeles, 1966.

15. Cox, H. — *On the Estimation of State Variables and Parameters for Noisy Dynamic Systems.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-9, No.1, January 1964, pp.5-12.

16. Cox, H. — *Estimation of State Variables via Dynamic Programming.* Joint Automatic Control Conference, Stanford University, 1964, pp.376-381.

17. Detchmendy, D.M.; Sridhar, R. — *Sequential Estimation of States and Parameters in Noisy Nonlinear Dynamical Systems.* Joint Automatic Control Conference, Troy, New York, 1965, pp.56-63.

18. Mortensen, R.E. — *Maximum-Likelihood Recursive Nonlinear Filtering.* Journal of Optical Theory and its Applications, Vol.2, No.6, 1968, pp.386-394.

19. Gura, I.A.; Henriksen, L.J. — *A Unified Approach to Nonlinear Estimation.* Aerospace Corporation, Report TR-0200 (4306-02)-2, Los Angeles, 1969.

20. Cramer, H. — *Mathematical Methods of Statistics.* Princeton University Press, New Jersey, 1961.

CHAPTER 7 - NONLINEAR FILTERING AND COMPARISON WITH KALMAN FILTERING

by

Lawrence Schwartz

Hughes Aircraft Company,
P.O. Box 90919,
Los Angeles, California 90009, USA

NOTATION

$\mathcal{E}$	expectation operator
$\hat{}$	conditional expectation
$\mathcal{L}$	backward diffusion operator
$\tilde{\mathcal{L}}$	forward diffusion operator
R_n	Euclidian N-space
$\triangleq$	equal by definition
$\approx$	approximately equal
$*$	transpose
$'$	derivative
$p(\cdot)$	probability density function
$p(\cdot\|\cdot)$	conditional probability density function
∂	partial differential operator
d	differential operator
$[]_{ij}$	the element in the i^{th} row and j^{th} column of the matrix in brackets
$()_{ave}$	average

CHAPTER 7 - NONLINEAR FILTERING AND COMPARISON WITH KALMAN FILTERING

Lawrence Schwartz

1. INTRODUCTION

The field of nonlinear filtering is quite broad, and a general discussion of the various approaches is not an appropriate subject for a single chapter in a book dedicated to Kalman Filtering. The specific form of nonlinear filtering treated herein is that which is most closely related to Kalman filtering, indeed, that which is the most natural generalization thereof: minimal-variance filtering. To further limit the scope, the derivation is limited to continuous-time minimal-variance filtering.

Curiously enough, the work that led to the successful derivation of nonlinear continuous-time minimal-variance filters does not stem from the paper by Kalman and Bucy[1] which contains the derivation of the linear continuous-time minimal-variance filter. Rather, the impetus derives from Stratonovich[2], in an even earlier paper, who noted that all the information related to the optimal estimate of the state of a system is contained in the conditional probability distribution function of the state given the measurements. Thus, the more general filter is not an outgrowth of the more restricted one but the result of a parallel line of investigation.

While only one approach to nonlinear filtering is discussed in detail, several nonlinear filters are computationally compared with a Kalman filter. All of the nonlinear filters mechanized reduce to the Kalman filter when the dynamical equation of the system is linear, the measurements are linearly related to the state, and the noise processes are *Gaussian*. An outline of the steps in the derivation is essential to set the stage for the following discussion. The first step in the analysis of any physical situation is the specification of a mathematical model; the choice should be made carefully, since the whole analysis depends upon the characteristics of the model. For minimal variance, which is a probabilistic criterion of optimality, the manipulations leading to the filter equations are made particularly simple by assuming that the random processes are white noises. The formal simplicity is gained at the expense of a certain amount of complication in the physical interpretation of the mathematical results.

Given the probabilistic criterion and the white-noise assumption, a natural mathematical model is the stochastic differential equation. Of course, the problem must be such that the stochastic differential equation satisfies existence and uniqueness conditions, which are different from those pertaining to nonstochastic differential equations. The essential difference stems from the fact that for white-noise models there is no bound on the forcing function and global conditions must be satisfied. From the stochastic differential equations for the system and the measurement, it is possible to derive a stochastic partial differential equation for the conditional density function. From the stochastic partial differential equation, in turn, it is possible to derive a stochastic differential equation for the expected value of any scalar function of the state of the system; a finite set of such equations form the vector equation for the filter.

The exact equation for the filter requires the instantaneous evaluation of the conditional expectation of several functions of the state of the system. To simplify the problem, the original model can be replaced by an approximating stochastic differential equation, from which an approximate filter can be derived more simply. It seems reasonable to require that the approximate model equations also satisfy existence and uniqueness conditions.

Guaranteeing existence and uniqueness for the approximate system does not quite do the same for the filter, but slightly stronger conditions on the equations suffice. At this point, one step remains in the validation: that of relating the stochastic differential equation for the filter to an ordinary differential equation for the actual mechanization. It is shown by Schwartz and Stear[3] that nonvalid filter equations similar to those previously derived by Schwartz and Bass[4] and by Fisher[5] for white-noise processes can be made computationally identical to valid filter equations.

2. MATHEMATICAL MODEL

2.1 White Noise

The usual mathematical formulation for a dynamical problem is a differential equation, nowadays most generally written in state-vector form:

$$\frac{dx}{dt} = f(t, x(t), u(t)) \,, \tag{1}$$

where x and f are n-vectors, u is an m-vector, and t is a scalar. The meaning of (1) is well known for most input functions u, but the ensuing analysis deals with white-noise input functions, and (1) must be reinterpreted. The present section explains the problems associated with white-noise inputs, and outlines the development of the necessary calculus of stochastic processes. White noise is often described as a random process with a power spectral density which is a constant or, equivalently, an autocorrelation function which is a Dirac δ-function. It is further noted that such a process has no physical meaning, since it would require infinite signal power. The foregoing definition is valid for stationary white noise, though the autocorrelation-function definition can be extended to the nonstationary case by allowing a time-varying coefficient for the δ-function.

The nonrealizability of white-noise processes is no reason to discard them; they occupy a place with respect to the family of stochastic processes analogous to the place of the Dirac δ-function with respect to functions. Just as the δ-function can be regarded as a limit of a sequence of unit-area pulses of decreasing width, a white-noise process can be considered as the limit of a sequence of processes which are step functions. Moreover, both are useful only when their integrals are considered; indeed, both can be made mathematically rigorous only in terms of their integrals. To be somewhat imprecise, the δ-function may be considered as the derivative of a unit step; with similar imprecision white-noise is the derivative of Brownian motion.

The practical reason for being concerned with white-noise processes is that, when differential equations are forced by white-noise, the solutions are Markov processes, i.e., the future is independent of the past. In other words, the solutions to differential equations forced by white-noise exhibit the stochastic analogue of the property of solutions to differential equations forced by ordinary functions: given the state of the solution at some time and the forcing function from that time on, the subsequent evolution of the solution is stochastically independent of the previous history.

The mathematical problem associated with white-noise is somewhat similar to that associated with the δ-function: the meaning of the integral. The δ-function is not really a function in the ordinary sense of the word, and no theory of integration can result in a value other than zero for

$$\int_{-\infty}^{\infty} \delta(t)\, dt$$

if the δ-function is assumed to be an ordinary function of t. However, by not ascribing values to $\delta(t)$ and considering only its integral, it is possible to construct a meaningful theory. Similarly, no ordinary theory of integration can make sense of

$$\int w(t)\, dt ,$$

where w is a white-noise process. Here again, if no instantaneous value is given to $w(t)$, a useful theory of stochastic integration is possible; that theory is outlined in the following section.

2.2 Stochastic Integrals

The exposition in this section is merely an outline of the mathematical derivation of the stochastic integral, and it does not include any discussion of stochastic integrals of discontinuous random processes. A complete discussion of stochastic integration can be found in Skorokhod[6]. The underlying idea is the description of a white-noise process as the derivative of a Brownian motion; the major difficulty lies in the fact that a Brownian motion fails to be differentiable somewhere in every interval of nonzero length, with probability one. Thus, if b is a Brownian motion, db/dt has no meaning and

$$\int g(t)(db/dt)\, dt$$

is generally not defined, even for continuous functions g. But, if db is an increment of b, it has a well-defined stochastic description, and the Stieltjes integral

$$\int g(t)\, db(t)$$

is a possibly meaningful alternate form. However, b is not a function of bounded variation and even the Stieltjes integral is not defined. The stochastic integral is a stochastically meaningful generalization of the Stieltjes integral, in which the approximating sums are required to converge in probability to the integral, rather than to converge in the ordinary sense.

2.3 Stochastic Differential Equations

The discussion contained in the sequel is limited to the following special case of Equation (1):

$$\frac{dx}{dt} = f(t,x(t)) + g(t,x(t))u(t) , \qquad (2)$$

where x and f are n-vectors, t is a scalar, u is an m-vector unit white-noise process with independent elements, and g is an $n \times m$ matrix. A vector white-noise process has elements derived from Wiener processes;

a Wiener process is a unit Brownian motion, i.e., its increments have zero mean and variance equal to the interval of time over which the increment is defined. While Equation (2) is less general than Equation (1), it is sufficient for most practical applications. Let g^i denote the i^{th} column of g and w_i the i^{th} row of the vector Wiener process from which u is derived. Then, by formal multiplication of Equation (2) by dt and integration of the resulting expression, Equation (2) can be rewritten as

$$x(t) - x(t_0) = \int_{t_0}^{t} f(s,x(s))\, ds + \sum_{i=1}^{n} \int_{t_0}^{t} g^i(s,x(s))\, dw_i(s) \,. \tag{3}$$

In Equation (3), the first integral is an ordinary integral, and the remaining n integrals are stochastic integrals. For simplicity, if the stochastic integral equation is satisfied by a process x with probability one, then Equation (3) is written in the form

$$dx = f(t,x)\, dt + g(t,x)\, dw(t) \,. \tag{4}$$

The simplified form Equation (4) is referred to as a stochastic differential equation and is understood to be a shorthand notation for Equation (3).

The following formula is necessary for the derivation of stochastic differential equations for functions of solutions of other stochastic differential equations. The scalar version of the formula is proved by Skorokhod in Reference 6 (p. 24ff); the vector version follows quite simply. Let x satisfy Equation (4) for $t_0 \leq t \leq t_f$; if a scalar function $\phi(t,x)$ is defined with continuous second cross partial derivatives with respect to the x_i for $t_0 \leq t \leq t_f$ and for all x, then the process $y(t) = \phi(t,x(t))$ satisfies the relation

$$dy = \left(\frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial x} f + \frac{1}{2} g^* \frac{\partial^2 \phi}{\partial x^2} g \right) dt + \frac{\partial \phi}{\partial x} g\, dw \tag{5}$$

where $\partial()/\partial x$ denotes the gradient (row) vector, $\partial^2()/\partial x^2$ denotes the Hessian (matrix of cross partials), and the "star" denotes matrix transpose.

2.4 Relation to the Physical Problem

There are two interfaces between the physical situation and the mathematical model in the filtering problem: the reduction of the dynamics to a stochastic differential equation and the interpretation and mechanization of the stochastic differential equation for the filter as a computational algorithm. The second interface is considered first. The interpretation of the stochastic integral in the context of the actual estimation environment is not at all a trivial matter. The filtering algorithm will be a finite-difference approximation to Equation (2), with u represented by a sampled measurement, not a white-noise. The problem of the interpretation of Equation (2) and the approach to use for the integration is discussed at length by Gray and Caughey[7]; they specify two approaches and propose a list of four pragmatic rules for choosing between the two approaches based on the interpretation of Equation (2). Another treatment of the difference between the two approaches is given by Wong and Zakai[8].

The real difference between the approaches is in the choice of whether to use the ordinary calculus or the stochastic calculus. In the nomenclature of Gray and Caughey[7] the former choice is the physical approach and the latter choice is the mathematical approach. In contrasting the two approaches, the authors are quick to state that neither approach is inherently correct; the choice should be made according to their pragmatic rules:

(i) If $g(t,x)$ is not actually a function of x, both approaches provide identical results.

(ii) If the problem is a strictly mathematical one, the mathematical approach must be used.

(iii) If Equation (8) is either an approximation to or a limit of the discrete problem

$$\left[x(t_{k+1}) - x(t_k)\right]/(t_{k+1} - t_k) = f(t_k, x(t_k)) + g(t_k, x(t_k))\, u(t_k) \,,$$

then the mathematical approach must be used.

(iv) If Equation (2) is either an approximation to a white-noise problem or the limit of a problem with short correlation time, then the physical approach must be used.

The computational effect of the difference between the two approaches is stated by Wong and Zakai[8] as follows: Let (w^n) be a sequence of piecewise linear approximations to the Wiener process in Equation (10) such that $w^n \to w$; then if (x^n) denotes the sequence of corresponding solutions, $x^n \to x$, where x is the solution to

$$dx_i = f_i(t,x(t))\, dt + \frac{1}{2} \sum_{j,k} g_{kj}(t,x(t)) \frac{\partial g_{ij}}{\partial x_k}(t,x(t))\, dt + \sum_j g_{ij}(t,x(t))\, dw_j \,. \tag{6}$$

They state some reservations about the correctness of Equation (6) in the vector case, but the same form is implied by the results of Gray and Caughey[7].

There is also a problem in relating the statistics of the real data to the statistics of the white-noise used in the model; this problem exists at both interfaces and is really the only one at the first. For simplicity, consider the following special case: Let $u(t)$ denote a sequence of pulses of width Δt and of random height given by a Gaussian distribution of zero mean and variance σ^2. The autocorrelation function for u is a triangular spike of width $2\Delta t$ and height σ^2; the area under the spike is then $\sigma^2\Delta t$. It then seems reasonable that the equivalent white-noise be specified by an impulse of weight $\sigma^2\Delta t$. The general result is as follows: If an n-dimensional white-noise is given by a covariance of the form $S(t)\delta(t-\tau)$, an n-dimensional pulse-sequence approximating the process should be chosen from a population given by a covariance of $S(t_i)/\Delta t$ for $t_i \leq t \leq t_i + \Delta t$. The case of continuous u is not quite so direct, though an equivalent formulation can be obtained by using the concept of a correlation time τ, which is a time interval such that $u(t)$ can be considered uncorrelated with $u(t+\tau)$.

Since the mathematical model is constructed under the assumption that the Wiener process has independent elements, one final step is required to model a noise with correlated elements. Let $S(t)\delta(t-\tau)$ be the desired covariance, which implies $S(t)$ is positive semi-definite for all t. Then there exists a matrix (which may be taken as symmetric) $S^{1/2}$ such that $S^{1/2}(S^{1/2})^* = S$. If $dv \triangleq S^{1/2}\,dw$, the white-noise derived from v has the proper covariance. For notational simplicity it may be assumed that $S^{1/2}(t)$ is incorporated into $g(t,x)$, and the formalism of Equations (2), (3), and (4) is still valid.

3. DERIVATION OF APPROXIMATE FILTER EQUATIONS

3.1 Conditional Density Function

Let the dynamic equation of the system be given by Equation (2), and let the measurement be given by

$$y(t) = h(t,x(t)) + r(t)\,v(t)\,, \tag{7}$$

where h is an l-vector, $l \leq n$, v is an l-dimensional unit white-noise and r is a nonsingular symmetric $l \times l$ matrix relating the unit white-noise to the modeled white-noise (equivalent to the matrix $S^{1/2}$ just described). Since the mathematical model cannot handle white-noise directly, it is assumed that the measurement is derived from a process z given by

$$dz = h(t,x(t))\,dt + r(t)\,db(t)\,, \tag{8}$$

where b is an l-dimensional Wiener process. The mathematical model of the system consists of the two vector equations (4) and (8). The problem is to find the minimal-variance estimate of $x(t)$, given the process $z(s)$ for $t_0 \leq s \leq t$; that is, to find the estimate $\hat{x}(t)$ such that the matrix given by

$$\mathcal{E}(x-X)(x-X)^* - \mathcal{E}(x-\hat{x})(x-\hat{x})^*$$

is positive semi-definite, where X is any estimate of x, the processes are evaluated at t, and the symbol $\mathcal{E}$ denotes expectation.

It is a simple exercise to show that the minimal-variance estimate of a random variable, given a related quantity, is simply the conditional expectation, so that

$$\hat{x}(t) = \mathcal{E}(x(t)|z(s),\ t_0 \leq s \leq t) \tag{9}$$

and the problem is to find an equation for the conditional expectation.

The first step is to derive formally an expression for the stochastic differential of $p(x|z)$, the conditional probability density function of $x(t)$ given $z(s)$, $t_0 \leq s \leq t$. It can be shown that

$$p(x|z) = \frac{\mathcal{E}(e^{\Phi}|x)\,p(x)}{\mathcal{E}(e^{\Phi})}\,, \tag{10}$$

where $p(x)$ is the probability density function for $x(t_0)$ and

$$\Phi \triangleq -\frac{1}{2}\int_{t_0}^{t} h^*r^{-1}h\,ds + \int_{t_0}^{t} h^*r^{-2}\,dz\,. \tag{11}$$

Following Bucy[5], let $p(x|z) \triangleq Q/P$. Also, let $\theta \triangleq \mathcal{E}(e^{\Phi}|x)$. Then Q is explicitly a function of θ and $p(x)$, and θ is explicitly a function of a random process Φ given by

$$d\Phi = \frac{1}{2}\,h^*r^{-2}h\,dt + h^*r^{-1}\,db\,. \tag{12}$$

In addition $p(x(t))$ is a function of t, while $x(t)$ and z are assumed fixed. Then, using Equations (5) and (12),

$$dQ = \frac{\partial Q}{\partial t}\,dt + \frac{\partial Q}{\partial \phi}\left(\frac{1}{2}\,h^* r^{-1} h\,dt + h^* r^{-1}\,db\right) + \frac{1}{2}\,\frac{\partial^2 Q}{\partial \phi^2}\,h^* r^{-1} h\,dt \quad . \tag{13}$$

Now

$$\frac{\partial Q}{\partial t} = \frac{\partial \theta}{\partial t}\,p(x(t)) + \theta\,\frac{\partial p(x(t))}{\partial t} = \theta\,\mathcal{L}\,p(x(t)) = \mathcal{L}\,Q \quad , \tag{14}$$

where $\mathcal{L}$ is the forward diffusion operator

$$-\sum_{i=1}^{n}\frac{\partial(f_i\,\cdot)}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2([\sigma\tau^*]_{ij}\,\cdot)}{\partial x_i \partial x_j} \quad . \tag{15}$$

Also,

$$\frac{\partial Q}{\partial \phi} = \frac{\partial^2 Q}{\partial \phi^2} = Q \quad , \tag{16}$$

Substituting Equations (14) and (16) into Equation (13) provides

$$\begin{aligned} dQ &= \mathcal{L}\,Q\,dt + Q(h^* r^{-1} h\,dt + h^* r^{-1}\,db) \\ &= \mathcal{L}\,Q\,dt + Q h^* r^{-1}\,dz \quad . \end{aligned} \tag{17}$$

Using Equations (16) and (15), and the definition of Q/P,

$$\begin{aligned} dP &= d\left(\int_{R_n} Q\,dx\right) = \int_{R_n} (dQ)\,dx \\ &= P\left\{\mathcal{L}\left[\int_{R_n}\left(\frac{Q}{P}\right)dx\right]dt + \left[\int_{R_n}\frac{Q}{P}\,h^* r^{-1}\,dx\right]dz\right\} \\ &= P\,\hat{h}^* r^{-1}\,dz = P\,\hat{h}^* r^{-1} h\,dt + P\,\hat{h}^* r^{-1}\,db \quad . \end{aligned} \tag{18}$$

Using Equation (5), with $x \triangleq P$, $\phi \triangleq P^{-1}$, $f \triangleq P\,\hat{h}^* r^{-1} h$, $g \triangleq P\,\hat{h}^* r^{-1}$, and $w \triangleq b$,

$$d(P^{-1}) = -P^{-1}\,\hat{h}^* r^{-1}(h - \hat{h})\,dt - P^{-1}\,\hat{h}^* r^{-1}\,db \quad .$$

Finally, using Equation (11) with $x \triangleq (P^{-1}, Q)^*$, $\phi \triangleq QP^{-1}$, $f \triangleq (-P^{-1}\,\hat{h}^* r^{-1}(h - \hat{h})$, $\mathcal{L}\,Q + Q\,h^* r^{-1})^*$, $g \triangleq (-P^{-1}\,\hat{h}^* r^{-1}, Q\,h^* r^{-1})^*$, and $w \triangleq b$,

$$\begin{aligned} d\left(\frac{Q}{P}\right) &= \left(\mathcal{L}\left(\frac{Q}{P}\right) + h^* r^{-1} h\,\frac{Q}{P}\right)dt + h^* r^{-1}\left(\frac{Q}{P}\right)db - \hat{h}^* r^{-1}(h - \hat{h})\left(\frac{Q}{P}\right)dt - \\ &\quad - \hat{h}^* r^{-1}\left(\frac{Q}{P}\right)db - \hat{h}^* r^{-1} h\left(\frac{Q}{P}\right)dt \\ &= \mathcal{L}\left(\frac{Q}{P}\right)dt + (h - \hat{h})^* r^{-1}\left(\frac{Q}{P}\right)(dz - \hat{h}\,dt) \quad . \end{aligned} \tag{19}$$

Let ϕ be any scalar function, twice continuously differentiable in x:

$$\hat{\phi} = \int_{R_n} \phi(x)\,p(x|z)\,dx$$

and

$$d\hat{\phi} = \int_{R_n} \phi(x)(dp(x|z))\,dx \quad . \tag{20}$$

Substituting Equation (19) into (20) provides

$$d\hat{\phi} = \int_{R_n} \phi(x)\left(\mathcal{L}\, p(x|z) + (h-\hat{h})^{*} r^{-1} p(x|z)(dz - \hat{h}\, dt)\right) dx \,. \tag{21}$$

Let $\mathcal{L}$ denote the formal adjoint of $\mathcal{L}$,

$$\sum_{i=1}^{n} f_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j=1}^{n} [gg^{*}]_{ij} \frac{\partial^2}{\partial x_i \partial x_j} \,. \tag{22}$$

Then Equation (21) becomes

$$d\hat{\phi} = \widehat{\mathcal{L}\phi}\, dt + (\widehat{\phi h} - \hat{\phi}\hat{h})^{*} r^{-1}(dz - \hat{h}\, dt) \,. \tag{23}$$

3.2 Approximate Filter Equations

The use of Equation (23) as a differential equation of $\hat{x}$ results in

$$d\hat{x}_i = \hat{f}_i\, dt + (\widehat{x_i h} - \hat{x}_i \hat{h})^{*} r^{-1}(dz - \hat{h}\, dt) \,, \tag{24}$$

which is not very practical because $\hat{f}_i$, $\hat{h}$, and $\widehat{x_i h}$ are needed continuously. As the first step in the approximation, let f and h be approximated by a second-degree expansion about $x = \hat{x}$; also, for notational simplicity, suppress the explicit appearance of t as an argument of f, g, and h, since their dependence on time is incidental to the following manipulations. Thus, adopting the summation convention

$$f_i(x) \approx f_i(\hat{x}) + f_{ij}^{(1)}(\hat{x})(x_j - \hat{x}_j) + \tfrac{1}{2} f_{ijk}^{(2)}(\hat{x})(x_j - \hat{x}_j)(x_k - \hat{x}_k) \,, \tag{25}$$

where

$$f_{ij}^{(1)} \triangleq \frac{\partial f_i}{\partial x_j} \quad \text{and} \quad f_{ijk}^{(2)} \triangleq \frac{\partial^2 f_i}{\partial x_j \partial x_k} \,.$$

A similar expression holds for h. From Equation (25),

$$\hat{f}_i(x) \approx f_i(\hat{x}) + \tfrac{1}{2} f_{ijk}^{(2)}(\hat{x})\widehat{(x_j - \hat{x}_j)(x_k - \hat{x}_k)} \,, \tag{26}$$

where $\widehat{(x_j - \hat{x}_j)(x_k - \hat{x}_k)}$ is the conditional covariance of x and is denoted P_{ij}.

Similarly

$$\widehat{x_i f_j}(x) \approx f_{jk}^{(1)}(\hat{x}) P_{ki} + \hat{x}_i \hat{f}_j(x) \,. \tag{27}$$

Using Equations (26) and (27) for f and h in Equation (24) provides

$$d\hat{x}_i \approx f_i(\hat{x})\, dt + \tfrac{1}{2} f_{ijk}^{(2)}(\hat{x}) P_{jk}\, dt + P_{ij} h_j^{(1)}(\hat{x}) r_{ik}^{-1}\left[dz_i - (h_i(\hat{x}) + \tfrac{1}{2} h_{mn}^{(2)}(\hat{x}) P_{mn})\, dt\right] \,. \tag{28}$$

The next step is to find a differential equation for P. In general, even with a second-degree expansion for the nonlinearities, an infinite sequence of differential equations is required, because all the moments are needed to describe the conditional density. However, by assuming an appropriate form for $p(x|z)$, the sequence stops at P. The first assumption, used by Bucy[9], is that third and fourth central conditional moments be neglected. In Schwartz and Bass[8] it is shown that the assumption is reasonable for a distribution with most of the probability mass sufficiently close to the mean. If it is assumed that $p(x|z)$ is Gaussian, the sequence also stops at P, and there is no restriction on the size of the moments.

Since P_{ij} can be written $(\widehat{x_i x_j} - \hat{x}_i \hat{x}_j)$, dP_{ij} is derived in two parts. Let $\phi = x_i x_j$ and use Equation (23) to find $d\widehat{x_i x_j}$. Since

$$\mathcal{L} x_i x_j = f_i x_j + f_j x_i + g_{ik} g_{jk} \,,$$

it follows that

$$d\widehat{x_i x_j} = \widehat{f_i x_j}\, dt + \widehat{f_j x_i}\, dt + \widehat{g_{ik} g_{jk}}\, dt + (\widehat{x_i x_j h_k} - \widehat{x_i x_j}\hat{h}_k) r_{ik}^{-1}(dz_i - \hat{h}_i\, dt) \,. \tag{29}$$

Next, let $\phi = \hat{x}_i\hat{x}_j$ and use Equations (5) and (24) to obtain

$$d\hat{x}_i\hat{x}_j = \frac{\partial(\hat{x}_i\hat{x}_j)}{\partial\hat{x}_k}\left[\hat{f}_k\,dt + (\widehat{x_k h_l} - \hat{x}_k\hat{h}_l)r_{lm}^{-1}(dz_m - \hat{h}_m\,dt)\right] +$$

$$+ \frac{1}{2}\frac{\partial^2(\hat{x}_i\hat{x}_j)}{\partial\hat{x}_k\partial\hat{x}_l}(\widehat{h_m x_k} - \hat{h}_m\hat{x}_k)r_{mn}^{-1}(\widehat{h_n x_l} - \hat{h}_n\hat{x}_l)\,dt \quad . \tag{30}$$

Combining Equations (29) and (30) provides

$$dP_{ij} = (\widehat{f_i x_j} - \hat{f}_i\hat{x}_j + \widehat{f_j x_i} - \hat{f}_j\hat{x}_i + \widehat{g_{ik}g_{jk}})\,dt - (\widehat{x_i h_k} - \hat{x}_i\hat{h}_k)r_{kl}^{-1}(\widehat{x_j h_l} - \hat{x}_j\hat{h}_l)\,dt +$$

$$+ \left[\widehat{x_i x_j h_k} - \widehat{x_i x_j}\hat{h}_k - \hat{x}_j\widehat{h_k x_i} - \hat{x}_i\widehat{h_k x_j} + 2\hat{x}_i\hat{x}_j\hat{h}_k\right]r_{kl}^{-1}(dz_l - \hat{h}_l\,dt) \quad , \tag{31}$$

which is an exact equation. The derivation of the approximate form of Equation (31) is a tedious algebraic exercise, and is covered in Bass, Norum and Schwartz[10]. For completeness, several intermediate results follow:

For either assumption about the conditional density,

$$\widehat{x_i x_j(x_k - \hat{x}_k)} - \widehat{x_i(x_k - \hat{x}_k)}\hat{x}_j - \widehat{x_j(x_k - \hat{x}_k)}\hat{x}_i = 0 \quad .$$

For the first assumption,

$$\widehat{x_i x_j h_k} - \widehat{x_i x_j}\hat{h}_k - \hat{x}_j\widehat{h_k x_i} - \hat{x}_i\widehat{h_k x_j} + 2\hat{x}_i\hat{x}_j\hat{h}_k = -\tfrac{1}{2}P_{ij}(h_{klm}^{(2)}(\hat{x})P_{lm}) \quad . \tag{32}$$

For the other assumption, the right-hand side of Equation (32) is

$$\tfrac{1}{2}(P_{il}P_{mj} + P_{im}P_{jl})h_{klm}^{(2)}(\hat{x}) \quad . \tag{33}$$

For simplicity, let the term in parentheses in Equation (33) be denoted by T_{ijkl}. Then, if g is also expanded in a first degree approximation, the following equations are obtained:

$$dP_{ij} \approx P_{ik}f_{jk}^{(1)}(\hat{x})\,dt + f_{ik}^{(1)}(\hat{x})P_{kj}\,dt - P_{ik}h_k^{(1)}(\hat{x})r_{lm}^{-1}h_{mn}^{(1)}P_{nj}\,dt +$$

$$+ g_{ik}g_{jk}(\hat{x})\,dt + g_{ikm}^{(1)}g_{jkn}^{(1)}(\hat{x})P_{mn}\,dt -$$

$$-\tfrac{1}{2}P_{kl}h_{mkl}^{(2)}(\hat{x})r_{mn}^{-1}(dz_n - h_n(\hat{x}) - \tfrac{1}{2}h_{npq}^{(2)}(\hat{x})P_{pq})P_{ij} \tag{34}$$

and

$$dP_{ij} \approx P_{ik}f_{jk}^{(1)}(\hat{x})\,dt + f_{ik}^{(1)}(\hat{x})P_{kj}\,dt - P_{ik}h_k^{(1)}(\hat{x})r_{lm}^{-1}h_{mn}^{(1)}(\hat{x})P_{nj}\,dt +$$

$$+ g_{ik}(\hat{x})g_{jk}(\hat{x})\,dt + g_{ikm}^{(1)}g_{jkn}^{(1)}(\hat{x})P_{mn}\,dt +$$

$$+ \tfrac{1}{2}T_{ijkl}h_{mkl}^{(2)}(\hat{x})r_{mn}^{-1}(dz_n - h_n(\hat{x}) - \tfrac{1}{2}h_{npq}^{(2)}(\hat{x})P_{pq}) \quad , \tag{35}$$

where Equation (34) uses (32) and Equation (35) uses (33), and $g_{ij}^{(1)}_k$ denotes $\partial g_{ij}/\partial x_k$.

4. SIMULATIONS

4.1 Description

The numerical investigation was conducted for eight different approaches to filtering, for two sets of dynamics, for two measurement schemes each. The two systems were chosen such that one satisfied existence and uniqueness conditions for stochastic differential equations while the other did not. No difficulties were anticipated in connection with filtering for the second system in the actual computer environment, and none were encountered.

4.2 Filter Equations

Because of the large number of cases involved in the study, only first-order systems are considered. The first system is

$$\left.\begin{aligned} &\text{(D1)} \quad \frac{dx}{dt} = -\frac{x}{1+x^2} + v \\ &\text{(M11)} \quad y = \tan^{-1} x + w \\ &\text{(M12)} \quad y = x + w \end{aligned}\right\} \tag{36}$$

where v and w are white-noises. With either measurement scheme, (M11) or (M12), the overall system satisfies existence and uniqueness. The second system, which does not, is given by

$$\left.\begin{aligned} &\text{(D2)} \quad \frac{dx}{dt} = -x^3 + v \\ &\text{(M21)} \quad y = x + x^3 + w \\ &\text{(M22)} \quad y = x + w . \end{aligned}\right\} \tag{37}$$

The filtering schemes are outlined below. For the outline, f and h are the system nonlinearities as in Equations (2) and (7). The white-noise processes, v and w, are assumed stationary with covariances σ_v^2 and σ_w^2 respectively so that $q = \sigma_v$ and $r = \sigma_w$. The subscript "n" denotes nominal, $\hat{x}$ is the estimate of x, p is the approximate covariance, and a prime denotes differentiation. Note that for linear measurements, (M12) and (M22), certain terms vanish and there are only three different filters.

1. *Linear*

The linear filtering algorithm can be applied to a set of equations linearized about an *a priori* nominal motion. The equations, derived by Kalman and Bucy[23] are

$$\left.\begin{aligned} \frac{d\hat{x}}{dt} &= f_n'\hat{x} + \sigma_w^{-2} p h_n'(y - h_n'\hat{x}) \\ \frac{dp}{dt} &= 2pf_n' - \sigma_w^{-2} p^2 h_n'^2 + \sigma_v^2 . \end{aligned}\right\} \tag{38}$$

2. *Quasi-Moment Minimal-Variance*

This is the filter derived by Schwartz and Bass[4], and independently by Fisher[5].

$$\left.\begin{aligned} \frac{d\hat{x}}{dt} &= f(\hat{x}) + \tfrac{1}{2} f''(\hat{x}) p + \sigma_w^{-2} p h'(\hat{x})[y - h(\hat{x}) - \tfrac{1}{2} p h''(\hat{x})] . \\ \frac{dp}{dt} &= 2pf'(\hat{x}) - \sigma_w^{-2} p^2 h'^2(\hat{x}) + \sigma_w^{-2} p^2 h''(\hat{x})[y - h(x) - \tfrac{1}{2} p h''(\hat{x})] + \sigma_v^2 . \end{aligned}\right\} \tag{39}$$

3. *Truncated Minimal-Variance*

This is the filter derived by Bass, Norum, and Schwartz[10].

$$\left.\begin{aligned} \frac{d\hat{x}}{dt} &= f(\hat{x}) + \tfrac{1}{2} f''(\hat{x}) p + \sigma_w^{-2} p h'(\hat{x})[y - h(\hat{x}) - \tfrac{1}{2} p h''(\hat{x})] \\ \frac{dp}{dt} &= 2pf'(\hat{x}) - \sigma_w^{-2} p^2 h'^2(\hat{x}) - \tfrac{1}{2}\sigma_w^{-2} p^2 h''(\hat{x})[y - h(x) - \tfrac{1}{2} p h''(\hat{x})] + \sigma_v^2 . \end{aligned}\right\} \tag{40}$$

4. *Modified Minimal-Variance*

This filter is a compromise between Equations (39) and (40) which is based on difference in the driving terms in the p equation. By dropping the driving term, the filter is simpler, yet the response falls between the responses for the two preceding filters.

$$\left.\begin{aligned} \frac{d\hat{x}}{dt} &= f(\hat{x}) + \tfrac{1}{2} f''(\hat{x}) p + \sigma_w^{-2} p h'(\hat{x})[y - h(\hat{x}) - \tfrac{1}{2} p h''(\hat{x})] \\ \frac{dp}{dt} &= 2pf'(\hat{x}) - \sigma_w^{-2} p^2 h'^2(\hat{x}) + \sigma_v^2 . \end{aligned}\right\} \tag{41}$$

5. *Maximum-Principle Least Squares*

This filter is derived by Dotchmendy and Sridhar[11] for minimizing an integral-square-estimation-error criterion, using deterministic techniques. By the use of Pontriagin's maximum principle the minimization is "reduced" to a two-point boundary-value problem which is solved by an invariant imbedding technique using an approximation to one boundary condition.

$$\frac{d\hat{x}}{dt} = f(\hat{x}) + \sigma_w^{-2} p h'(\hat{x})[y - (\hat{x})]$$

$$\frac{dp}{dt} = 2pf'(\hat{x}) - \sigma_w^{-2} p^2 h'^2(\hat{x}) + \sigma_w^{-2} p^2 h''(\hat{x})[y - h(\hat{x})] + \sigma_v^2 . \qquad (42)$$

Actually, Equation (42) is a special case of the derivation in Dotchmendy and Sridhar[11]; they used arbitrary weighting functions in the criterion integrand, while Equation (42) corresponds to the particular set of weighting functions that result in the Kalman filter for linear dynamics.

6. *Dynamic Programming Least Squares*

This filter is derived by Cox[12] for a criterion which is similar to that used for the previous filter. The minimization is effected by dynamic programming using a quadratic approximation to the cost function.

$$\frac{d\hat{x}}{dt} = f(\hat{x}) + \sigma_w^{-2} p h'(\hat{x})[y - h(\hat{x})]$$

$$\frac{dp}{dt} = 2pf'(\hat{x}) - \sigma_w^{-2} p^2 h'^2(\hat{x}) + \sigma_v^2 . \qquad (43)$$

A look at Equation (43) shows that this filter is essentially equivalent to using linear filtering about the computed mean, a technique that had been used heuristically previously.

7. *Discrete-Measurement Minimal-Variance*

This filter is derived by Jazwinsky[13] for a minimal-variance criterion under the assumption that the measurements arrive at isolated instants. The form presented here is the limiting form for continuous measurements.

$$\frac{d\hat{x}}{dt} = f(\hat{x}) + \tfrac{1}{2} f''(\hat{x}) p + \sigma_w^{-2} p h'(\hat{x})[y - h(\hat{x})]$$

$$\frac{dp}{dt} = 2pf'(\hat{x}) - \sigma_w^{-2} p^2 h'^2(\hat{x}) - \tfrac{1}{2}\sigma_w^{-2} h''(\hat{x})[y - h(\hat{x})] + \sigma_v^2 . \qquad (44)$$

Since the evolution of the system itself is considered continuous, the portion of the equations related to updating the estimate in the absence of measurements agrees with Equations (39), (40), and (41). The loss of the term in $h''(\hat{x})$ is due to the difference between the physical and mathematical approaches discussed in Section 2. In the derivation of Equation (44), Jazwinsky uses the approximation to the conditional density that is used to derive Equation (40).

8. *Modified Discrete-Measurement Minimal-Variance*

This filter is related to Equation (44) in the same way that Equation (39) is related to Equation (40), i.e., the conditional density is assumed to be Gaussian.

$$\frac{d\hat{x}}{dt} = f(\hat{x}) + \tfrac{1}{2} f''(\hat{x}) p + \sigma_w^{-2} p h'(\hat{x})[y - h(\hat{x})]$$

$$\frac{dp}{dt} = 2pf'(\hat{x}) - \sigma_w^{-2} p^2 h'^2(\hat{x}) + \sigma_w^{-2} p^2 h''(\hat{x})[y - h(\hat{x})] + \sigma_v^2 . \qquad (45)$$

The application of Equations (38)-(45) to Equations (36) and (37) is straightforward and is not carried out here.

The computer program used for the simulation study includes a simple rectangular-rule integration with constant step-size. The use of constant step-size allows each filter to be compared on the basis of the same pseudo-random sequence. The random number generator is a combination of a standard uniform random sequence routine plus an

approximate transformation to a Gaussian random sequence. The program has two modes of operation: in one mode, only a single estimation is made with a pre-specified initial condition for the state; in the other mode, several runs are made for random initial conditions, and the statistics of the estimation errors are computed. The output from the program is a computer-prepared plot showing the time-history of the filter response or the error statistics. The plots show every 20th point with linear interpolation between.

4.3 Results

As noted earlier, the simulation study was not intended to be a complete investigation of the computational characteristics of nonlinear filters. Two standard cases were used, which were computationally docile when used with the step-size chosen. For both dynamical equations (D1) of Equation (36) and (D2) of Equation (37), the standard run consists of 5000 points 0.001 second apart for a total problem time of five seconds, with $\sigma_v^2 = 1$ and $\sigma_w^2 = 10$. Only the initial condition differs. Figures 1 and 2 show the state x and the measurement y for the two standard cases. The figures show the nonlinear measurement, but the noise is so large that there is not too much apparent difference between the linear and nonlinear measurements.

In the response to the standard inputs, there is a large difference between the response of the linear systems and the nonlinear systems, while the various nonlinear systems are remarkably similar. Figures 3 and 4 show the estimation error for the two standard cases. As might be expected, linear measurements help the linear system. What might not be expected is that the relative error performance of the nonlinear filters within the shaded region is different for the two cases. To show that the results in Figures 3 and 4 are not peculiar to the particular initial condition and pseudo-random sequences, three representative filters were chosen for a set of ten runs. The mean and mean-square errors are shown in Figures 5 and 6 for initial conditions with variance one for the first case and variance 1/4 for the second case. It can be seen that the error statistics are still quite close, practically indistinguishable from a mean-square error point of view. Moreover, the relative error performance of the chosen filters is qualitatively unchanged.

While on the topic of statistical runs, it is interesting to note the comparison between the output of the p equation for an initial condition of one standard deviation and the mean-square error for ten runs for one case, as shown in Figure 7 for filter 2, dynamics D1, and measurement M11. Moreover, it was found that the change in the mean-square error for a larger number of cases, up to 50, while noticeable, was relatively small.

The effects of changing the statistics of the random sequences are similar for all the filters, and are intuitively reasonable. Changing σ_v has little effect on the initial response, though the higher σ_v the worse the ultimate following. Contrariwise increasing σ_w slows the response noticeably without affecting the ultimate tracking. The error responses for filter 2 with D1 and M11 are shown in Figures 8 and 9.

The runs demonstrating the correctness of the noise model show the effects of improperly matching the filter parameters to the noise statistics. The effect of the mismatch on $x - \hat{x}$ and p is shown in Figures 10 and 11. For all the curves shown, the actual pseudo-random inputs were taken from identical populations; only the filter parameters varied. Evidently a statistical mismatch can seriously affect the performance of the filter. It also appears that the value of p generated by the filter is a good index of the performance, even when the statistics are poorly matched.

The final aspect studied on the computer was the effect of sampling the data. Two approaches to sampling were considered: sample-and-hold and pulse sampling. For the sample-and-hold runs, the effective noise variance was obtained using a Δt equal to the sampling period, rather than the computation interval. For the pulse-sampling runs, σ_w^{-1} was made zero for those intervals during which no measurements were made. Both types of sampling were applied to filter 2 with dynamics D1 and measurements M11. The response plots are shown in Figure 12. As should be expected, sampling is detrimental to the initial response of the filter, though ultimately the estimate settles in to the proper value. From the one case considered, sample-and-hold appears somewhat better than impulse sampling, which may be due to the smoothing effect of the zero-order hold.

There is one problem in mechanizing the filters, which has not yet been discussed, that requires much more investigation: the effects of the filter parameters and the step-size on computational stability. For some combinations of parameters σ_v and σ_w and initial covariance $p(t_0)$, the step-size must be made small to stabilize the computation during the initial portion of the run. The reason for the instability is the large value of the derivatives and the consequent large truncation errors introduced in the integration. A particularly insidious form of computational instability arose in the course of the simulation study: it is possible to be in a conditionally stable region such that one pseudo-random sequence with a given statistical description results in a stable response, while another sequence drawn from the same population does not.

Apparently, nonlinear filtering is superior to linear filtering, although no one nonlinear filter offers any clear-cut advantage over any other from a performance point of view for the problem considered. It also appears that the filters derived from a continuous model are quite amenable to sampled-data use.

ACKNOWLEDGMENT

The research described in this paper was supported in part by Air Force Grant AFOSR-699-67.

REFERENCES

1. Kalman, R.E. Bucy, R.S. — *New Results in Linear Filtering and Prediction Theory.* Transactions, American Society of Mechanical Engineers, Series D, Journal of Basic Engineering, Vol.83, March 1961, pp.95-108.

2. Stratonovich, R.L. — *Conditional Markov Processes.* Theory of Probability and its Applications, Vol.5, No.2, 1960, pp.156-178.

3. Schwartz, L. Stear, E.B. — *Computational Validity of Approximate Nonlinear Minimal-Variance Filters.* Information Sciences, to appear April 1969.

4. Schwartz, L. Bass, R.W. — *Extensions to Optimal Multichannel Nonlinear Filtering.* Hughes Aircraft Company, Culver City, California, Report SSD 60220R, February 21, 1966.

5. Fisher, J.R. — *Conditional Probability Density Functions and Optimal Nonlinear Estimation.* PhD in Engineering, University of California, Los Angeles, 1966.

6. Skorokhod, A.V. — *Studies in the Theory of Random Processes.* Addison-Wesley, Reading, Mass., 1965.

7. Gray, A.H. Caughey, T.K. — *A Controversy in Problems Involving Random Parametric Excitation.* Journal of Mathematics and Physics, Vol.44, September 1965, pp.288-296.

8. Wong, E. Zakai, M. — *On the Relation Between Ordinary and Stochastic Differential Equations.* Electronics Research Laboratory, University of California, Berkeley, California, Report 64-26, August 11, 1964.

9. Bucy, R.S. — *Nonlinear Filtering Theory.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-10. April 1965, p.198.

10. Bass, R.W. et al. — *Optimal Multichannel Nonlinear Filtering.* Hughes Aircraft Company, Culver City, California, Report SSD 50004R, August 25, 1966.

11. Detchmendy, D.M. Sridhar, R. — *Sequential Estimation of States and Parameters in Noisy Nonlinear Dynamical Systems.* Joint Automatic Control Conference, Preprints of Papers (AIAA, AIChE, ASME, IEEE, ISA) June 22-25, 1965, Rensselaer Polytechnic Institute, Troy, New York, pp.56-63.

12. Cox, H. — *Estimation of State Variables via Dynamic Programming.* Joint Automatic Control Conference, Preprints of Papers (AIAA, AIChE, ASME, IEEE, ISA) June 24-26, 1964, Stanford University, Stanford, California.

13. Jazwinsky, A.H. — *Nonlinear Filtering with Discrete Observations.* AIAA Paper 66-38, presented at the AIAA 3rd Aerospace Sciences Meeting, New York, NY, January 24-26, 1966.

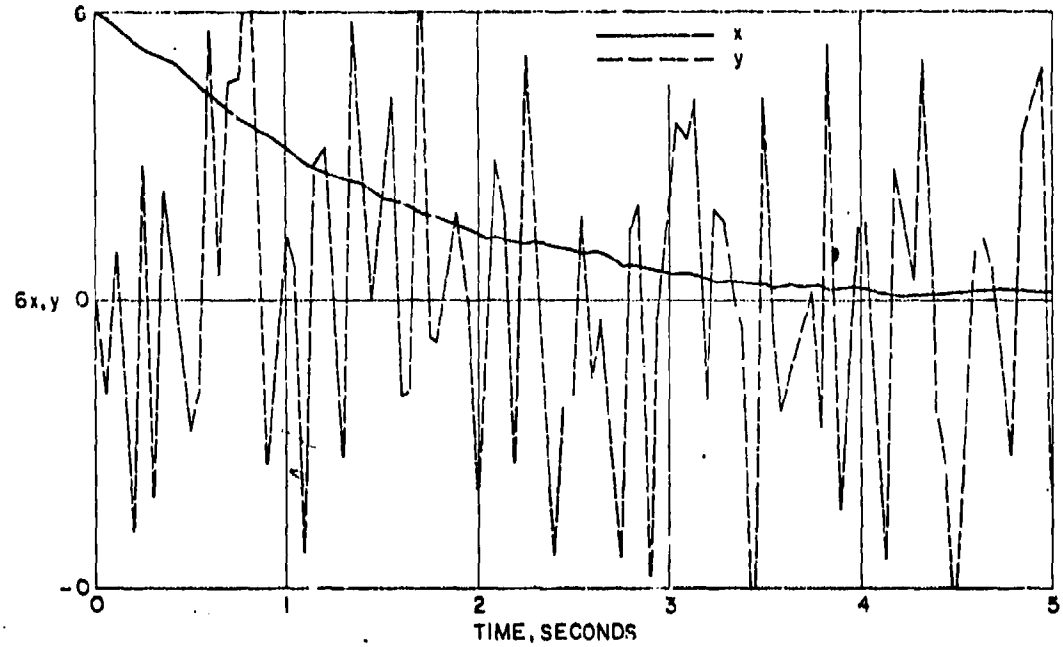


Fig.1 Standard response for dynamics D1 with measurements M11

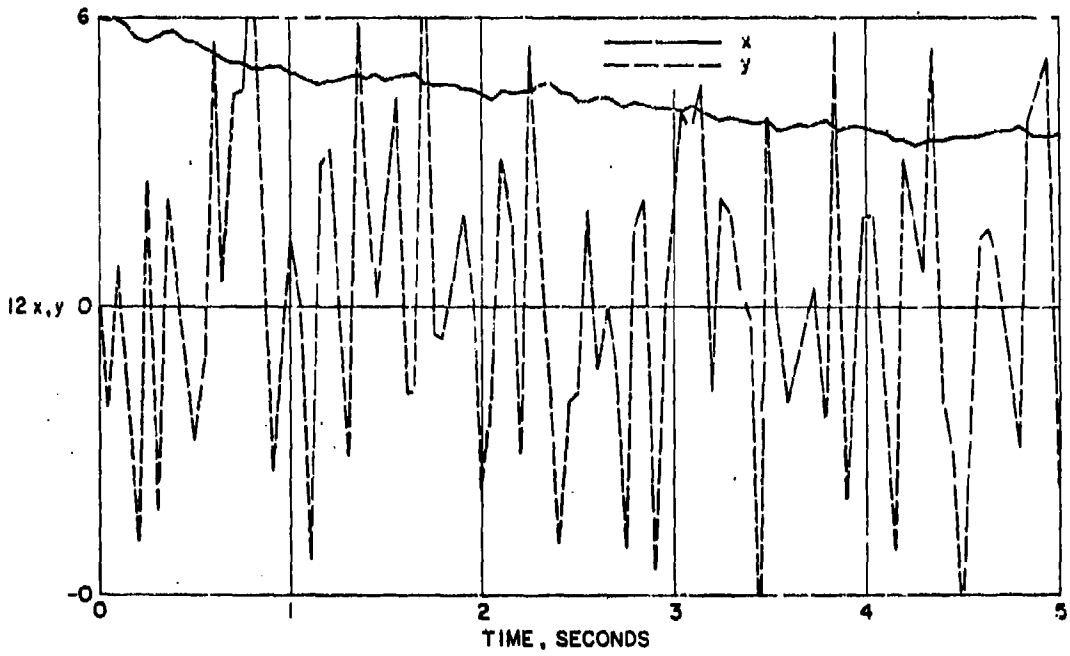


Fig.2 Standard response for dynamics D2 with measurements M21

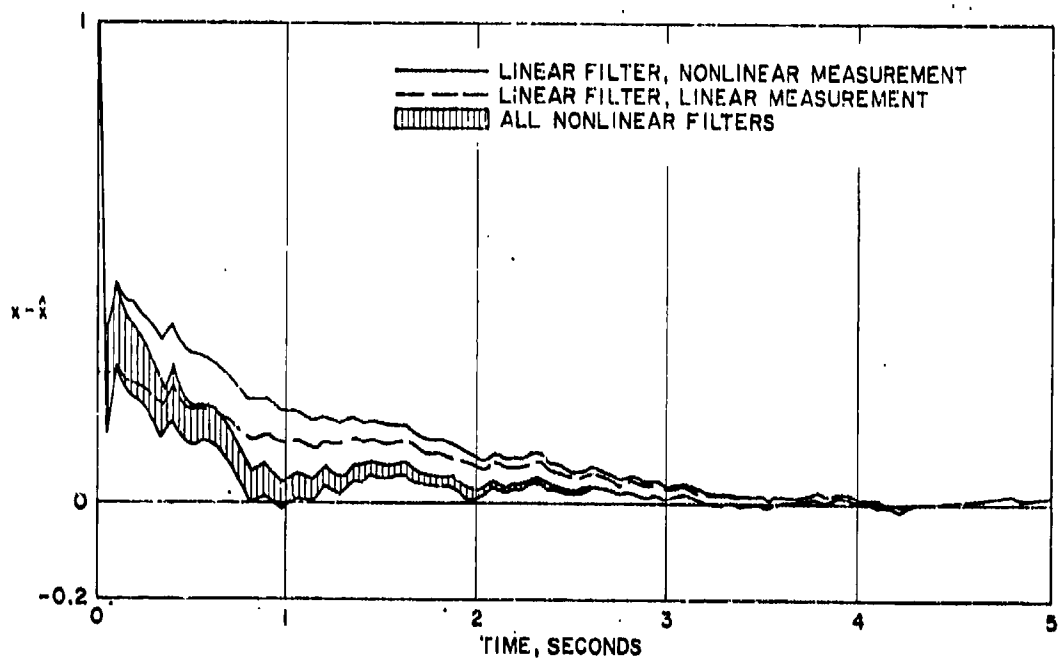


Fig. 3 Estimation error for standard case for dynamics D1

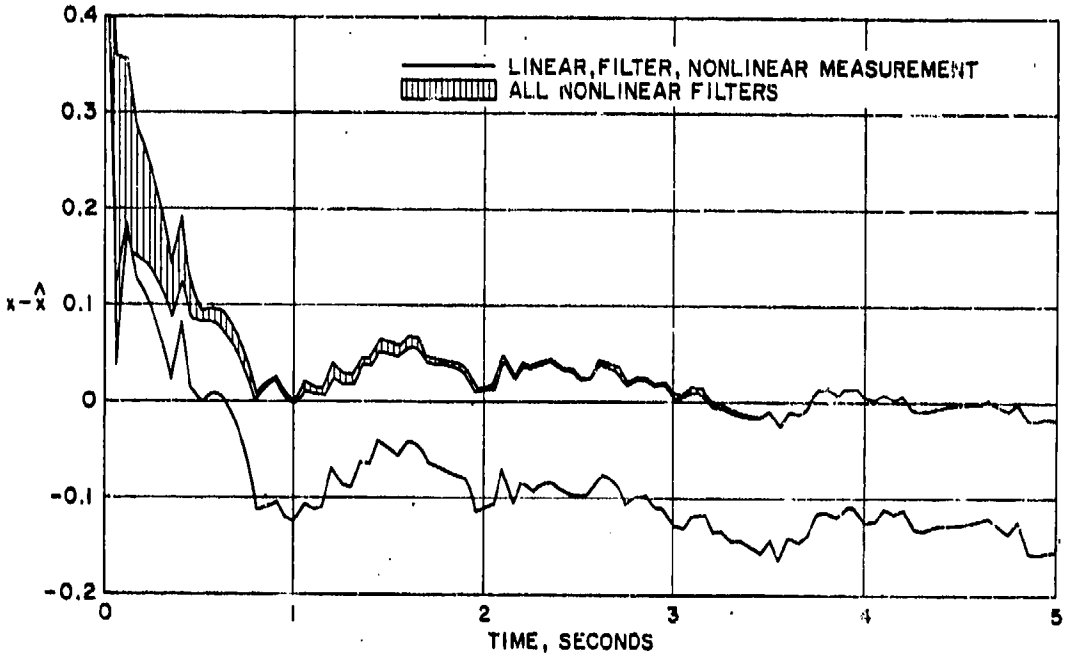


Fig. 4 Estimation error for standard case for dynamics D2

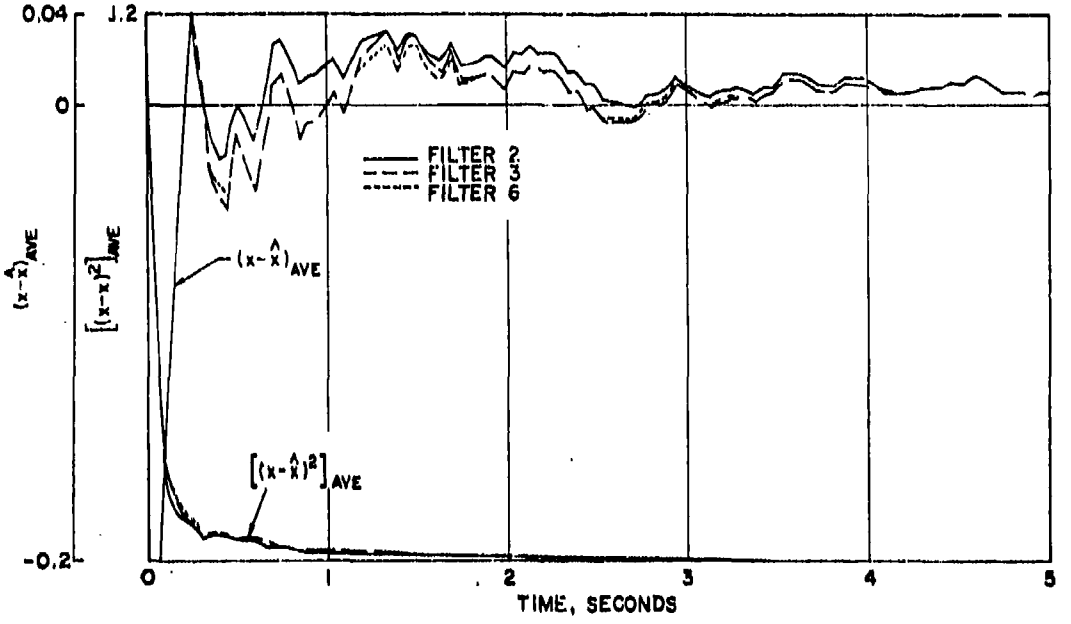


Fig.5 Error statistics for ten runs of dynamics D1 and measurement M11

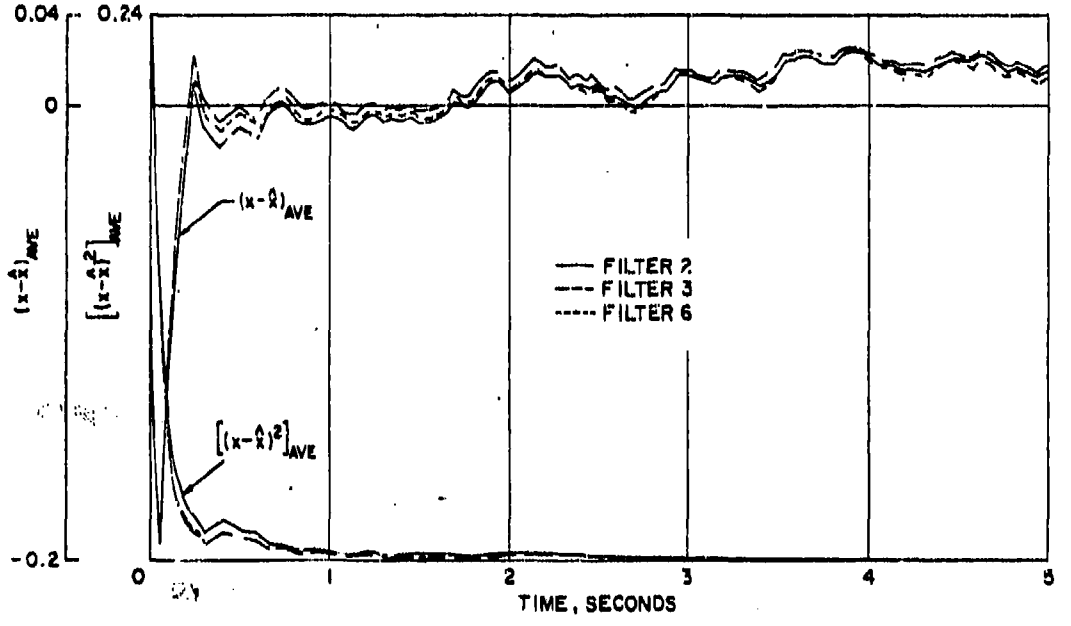


Fig.6 Error statistics for ten runs of dynamics D2 and measurement M21

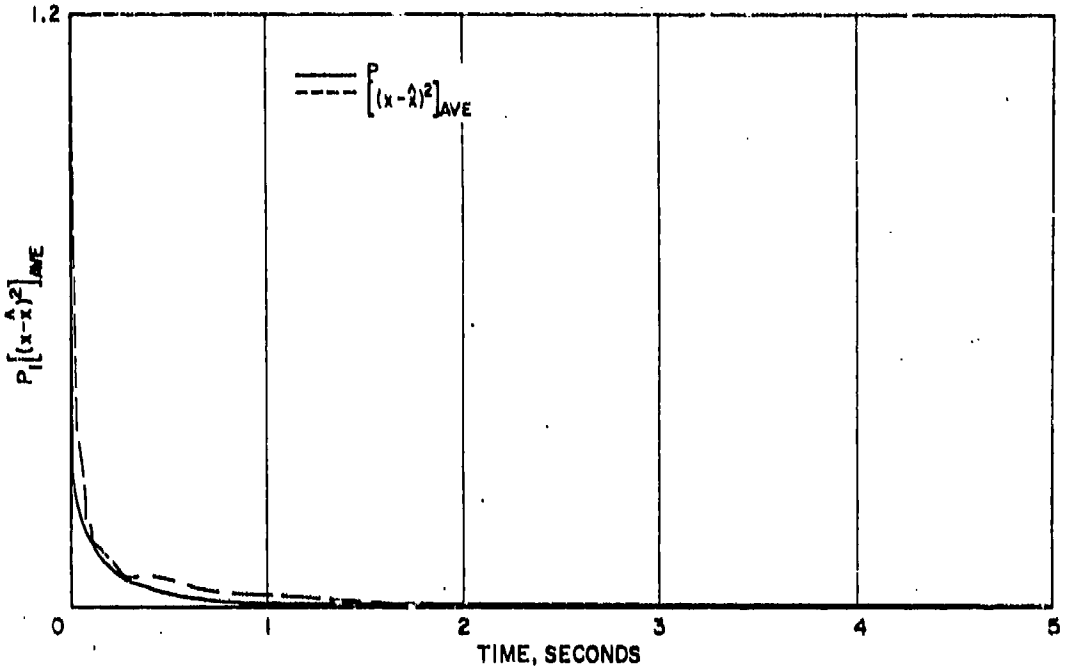


Fig. 7 Comparison of p with measured mean-square error

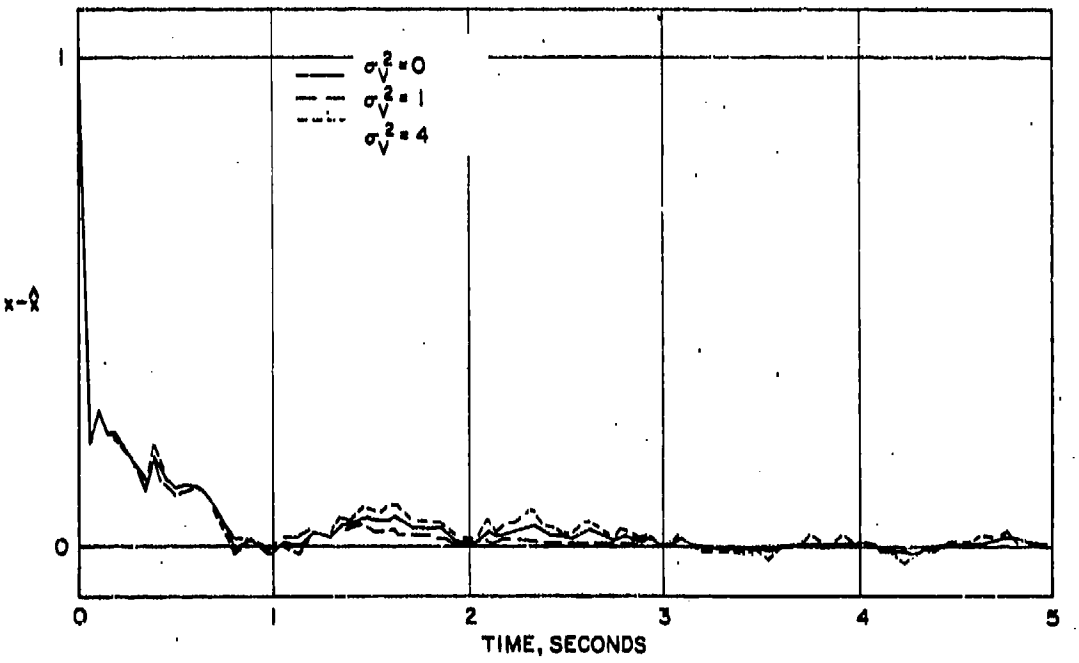


Fig. 8 Effect of varying σ_v

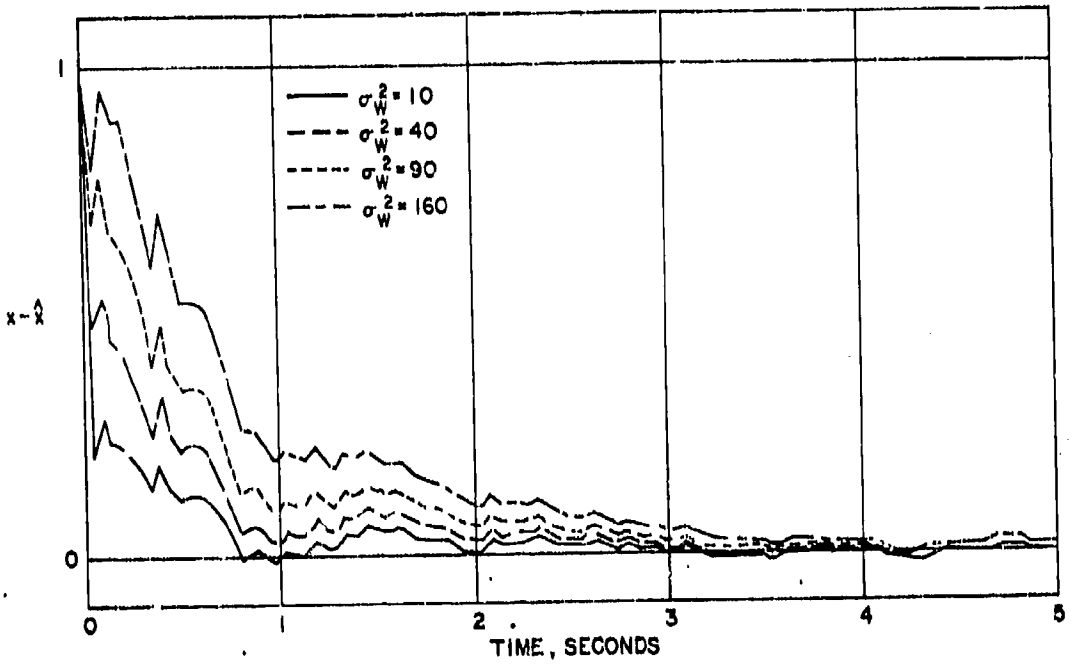


Fig. 9 Effect of varying σ_w

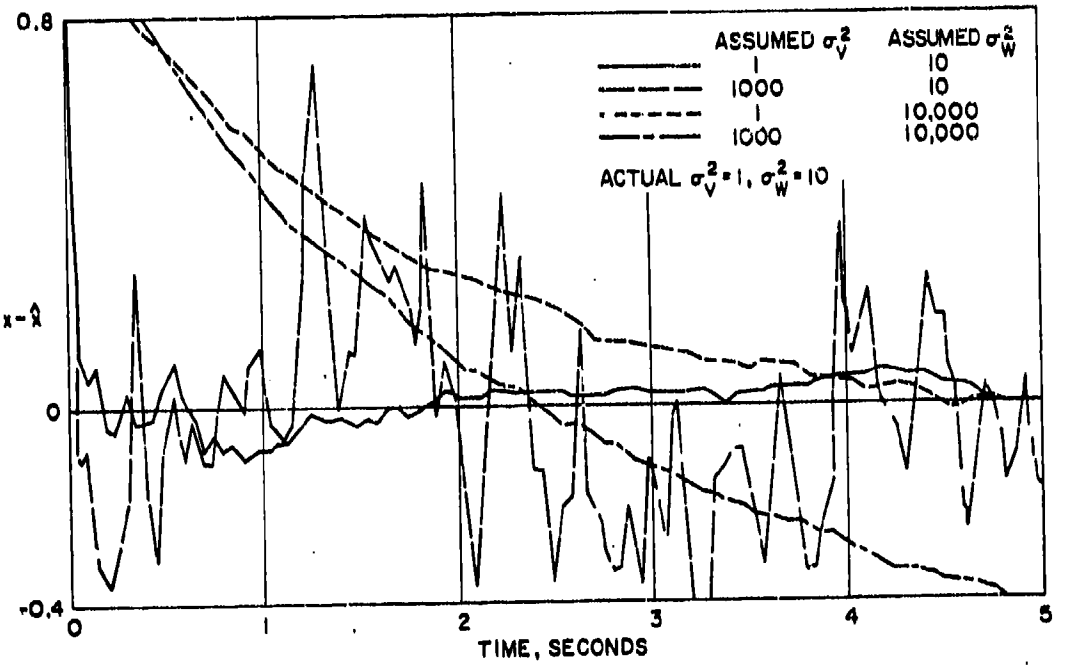


Fig. 10 Effect of statistical mismatch on estimation error

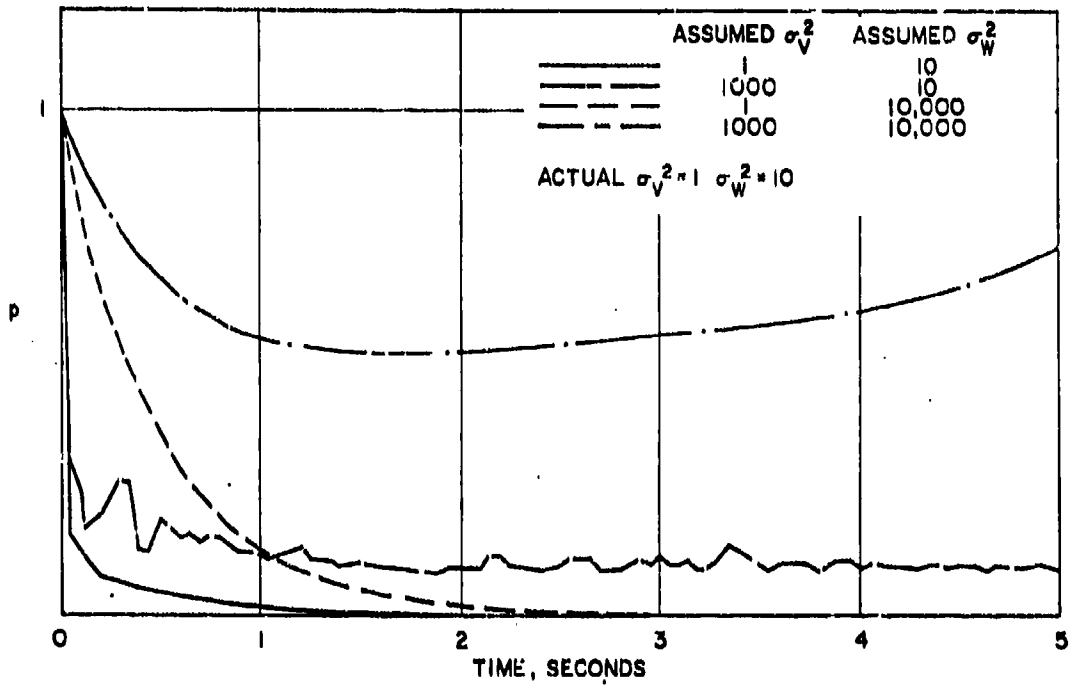


Fig. 11 Effect of statistical mismatch on computed covariance

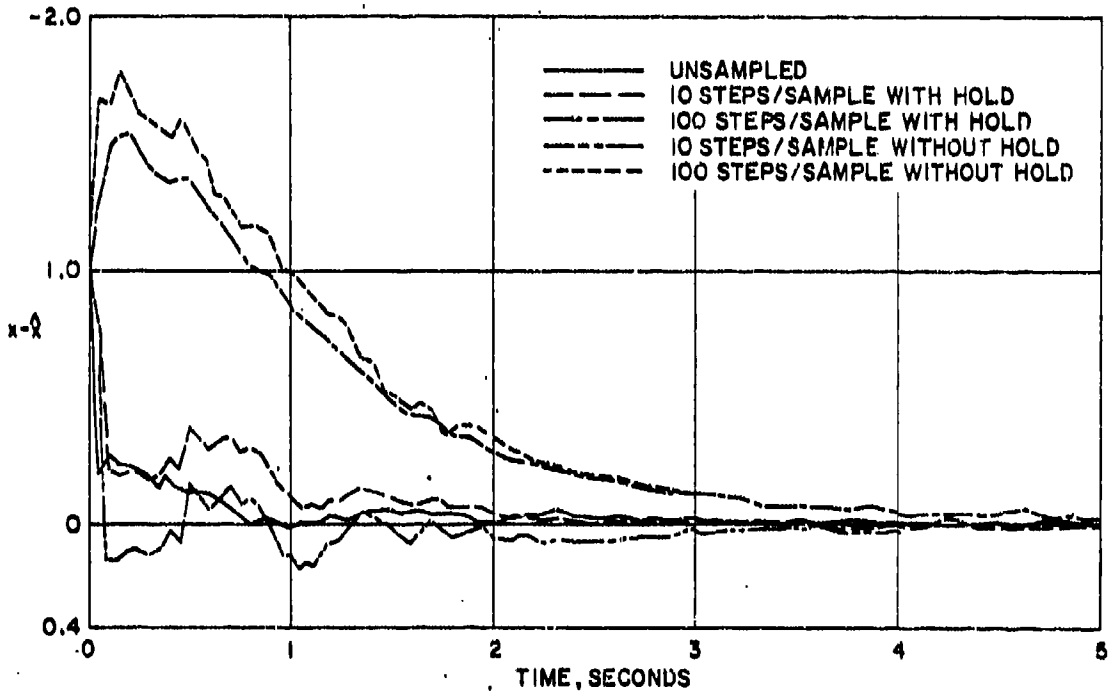


Fig 12 Error response for sampled input

CHAPTER 8 - LINEAR SMOOTHING TECHNIQUES

(POST-FLIGHT DATA ANALYSIS)

by

Herbert E. Rauch

Lockheed Missiles and Space Company
Palo Alto, California, USA

CHAPTER 8 - LINEAR SMOOTHING TECHNIQUES (POST-FLIGHT DATA ANALYSIS)

Herbert E. Rauch

1. INTRODUCTION

1.1 Background

Since the time of Gauss[1] people have been interested in the problem of least-squares estimation. Gauss not only formulated the estimation and devised a technique for solving it (a sequential algorithm called Gaussian elimination, which is used on modern computers), but he applied this to determine the orbit of the asteroid Ceres. The limitation on his solution was that it only applied to the case where the system to be estimated is deterministic and the errors on the individual measurements are uncorrelated. When a system is deterministic, its evolution can be completely specified from the initial conditions and the differential equations of the system. Celestial bodies obey Newton's laws of motion and disturbing forces due to other celestial bodies are known to a reasonable degree of accuracy, so for practical purposes their motion can be considered deterministic. Terrestrial bodies, on the other hand, are subject to disturbing forces whose effects cannot be described exactly beforehand although their statistical properties might be known. When a system is described by differential equations subject to disturbances which are statistically time-varying, the system can be called a stochastic process.

The pioneer work of Wiener[2] discussed and solved the problem of linear least-squares estimation for stochastic processes. However, his results were mainly for stationary, non-time-varying systems, and there was some difficulty in obtaining numerical answers for large order systems. Many papers have appeared since then giving different solutions to the problem under more general conditions. The most widely used solution for filtering and prediction is that of Kalman[3,4] and Bucy[5,6]. The primary advantage of their solutions is that they apply to time-varying systems and the equations which specify the optimum filter are in the form of recursive difference equations which have obvious computational advantages. However, Kalman and Bucy did not consider the important problem of smoothing. Filtering allows one to estimate present values of the variables of interest using present data, while smoothing allows one to estimate past values. A number of papers have presented different derivations of recursive solutions for the smoothing solution from different points of view for both discrete and continuous systems. Bryson and Frazier[7] gave one of the first recursive solutions for continuous systems using the calculus of variations and the method of maximum likelihood. Similar solutions with some alternative forms were given by Cox[8] using Dynamic Programming, Rauch, Tung and Striebel[9] using discrete time manipulations of Bayes's rule and the method of Maximum Likelihood, Lee[10] using the calculus of variations, and Meditch[11] using projection arguments. Two more recent results were by Fraser[12] who expressed the solution as the combination of two filtered estimates from a forward and backward Kalman-Bucy filter, and Kailath and Frost[13,14] who made use of an innovations approach to linear smoothing. In their papers Kailath and Frost also give an interesting summary of work in filtering and smoothing with a list of over forty references.

Obviously, all these solutions should ultimately give the same answer to the same problem, but the different derivations give insight into the reasons for different recursive forms of the smoothing solution. For deterministic systems with linear equations, the solution simplifies considerably because the smoothed estimate of past values is obtained directly from the filtered estimate of the latest value by integrating the differential equations of motion backwards. For linear smoothing of stochastic systems a similar procedure can be used in which differential equations are integrated backwards, starting with the latest filtered estimate. However, there are additional inputs to the differential equations involving earlier filtered estimates or measurements, depending on the form of the solution. The purpose of this presentation is to outline the procedures and solutions which can be used in the linear smoothing problem and to indicate how the results can be applied to a problem in post-flight data analyses which involves tracking a satellite disturbed by stochastic drag forces[15].

1.2 Notation

In general, lower case letters, such as u, v, x, and z, will denote column vectors while upper case letters, such as A, B, F, G, and Φ, will denote matrices. The letter I represents the identity matrix. The superscript prime, as in A' or u', will denote the transpose of the matrix or the transpose of a column vector (which becomes a row vector). The product of a column vector and a row vector, such as xx' is a matrix. The symbol E represents expected value so that $E[A]$ represents the expected value of the quantity A. The superscript -1 as in A^{-1} means the inverse of the matrix A. In all cases it will be assumed that the inverse of the matrix exists, although quite often, the inverse can be replaced by a pseudo-inverse or generalized inverse without changing the results.

1.3 Linear Estimation

For the discrete version of the linear estimation problem the system to be estimated can be described by the following set of matrix difference equations:

$$x_{k+1} = \Phi_{k+1} x_k + u_k \quad \text{for} \quad k = 1, 2, \ldots, N . \tag{1.1}$$

The linear measurements obtained from the system are given by another set of matrix equations,

$$z_k = H_k x_k + v_k \quad \text{for} \quad k = 1, 2, \ldots, N . \tag{1.2}$$

The vector x represents the state of the system which is to be estimated while the vector z represents the known measurements. The vectors u and v are not known exactly, but they are zero mean independent random variables with known covariance. The variable u represents random changes in the state (dynamic noise) while the variable v represents random changes in the measurements (measurement noise). The subscript k represents the value of the quantities at the time of the k^{th} measurement. If the dynamic noise u is identically zero for all time, the system is said to be deterministic. The matrices Φ (transition matrix) and H (output matrix) represent known quantities which can change from one measurement to the next.

Most physical systems will involve nonlinear equations, but it is assumed the above set of linear equations can be obtained by linearizing about some nominal values for the state and the measurements. It may be that the physical system is governed by a set of linear (or linearized) differential equations rather than difference equations, although the measurements will take place at discrete times. In what case the original system differential equations must be integrated to obtain the required difference equation relating the change in state from one measurement to the next. Conversely, under certain conditions, in the limiting case as the time between measurements goes to zero, the discrete system will approach a continuous system.

The optimum estimate will be the linear estimate which minimizes the mean square error. Calculating the estimate requires knowing the mean and covariance of all the random variables of interest, but no higher moments. If all the random variables have a normal probability distribution, the estimate will be the conditional mean of the state given the measurements. Sometimes the estimate is also called the Maximum Likelihood estimate because it maximizes the conditional probability distribution.

Let $\hat{x}_{j/k}$ denote the optimum estimate of the state x_j given all the measurements up to z_k. If j is greater than or equal to k, it is called filtering and prediction. If j is less than k, it is called smoothing. The optimum filtering solution as derived by Kalman, can be written, where B is the matrix gain on the Kalman filter,

$$\left.\begin{aligned} \hat{x}_{k+1/k} &= \Phi_{k+1/k} \hat{x}_{k/k-1} + B\mathfrak{z}_k \\ \mathfrak{z}_k &= z_k - H_k \hat{x}_{k/k-1} . \end{aligned}\right\} \tag{1.3}$$

The initial conditions for the filtering solution are based on the *a priori* information. The quantity $\mathfrak{z}_k$, which contains the new information, is the difference between the actual measurement z_k and the best estimate of the measurement. This quantity will play an important part in the derivation of the smoothing solution.

The above filtering solution is recursive in the sense that the desired estimate is calculated sequentially from previous estimates. The smoothing solution can also be written in a recursive form and, in fact, four different recursive versions of the smoothing solution will be presented here. In theory, these four versions will give identical results but, due to such things as computer programming or running time or the effect of roundoff errors, one way may be superior to another for a particular problem.

The first version, called fixed point smoothing, is useful when it is necessary to calculate the smoothed estimate at only one point. For instance, the smoothed estimate of only the initial conditions of the state might be desired. The form of the solution for fixed point smoothing is similar to that for filtering as shown below, where the matrix gain B is a general weighting coefficient:

$$\hat{x}_{k/N} = \hat{x}_{k/N-1} + B\mathfrak{z}_N . \tag{1.4}$$

The second version, smoothing using filtered estimates, has two parts. First, the Kalman filtering solution is used to process the measurements sequentially, in the order in which they were received, to obtain the filtered estimate (forward sweep). Second, the smoothing solution is used to process the filtered estimate a second time, in reverse order, from the last to the first, to obtain the smoothed estimated (backward sweep). The initial condition for the backward sweep is the filtered estimate of the state at the final time, which is also the smoothed estimate of the state at that time. The form of the solution for smoothing with filtering is shown below, where the matrices A and B are general weighting coefficients:

$$\hat{x}_{k/N} = A\hat{x}_{k+1/N} + B\hat{x}_{k/k} .$$

There is an alternative method in which the backward sweep involves measurements instead of the filtered estimates. This is the third version, which is called smoothing using measurements. The recursive solution can be written schematically as shown below, where the vector w_k is also calculated recursively:

$$\hat{x}_{k/N} = A\hat{x}_{k+1/N} + Bw_k$$

$$w_{k-1} = Cw_k + D(z_k - H_k\hat{x}_{k/N}) .$$

The fourth version, called fixed lag smoothing, can be used when it is desirable to smooth for a fixed number of measurements. The recursive solution involves both filtered estimates and measurements:

$$\hat{x}_{k+1/N+1} = A\hat{x}_{k/N} + B\tilde{x}_{N+1} + C\hat{x}_{k/k} .$$

1.4 Outline of Work

The remainder of this presentation examines the smoothing solution for the discrete system (Section 2) and the continuous system (Section 3), and applies it to a problem in satellite tracking (Section 4). It is shown that the optimum estimate for the discrete system must satisfy the sampled Wiener-Hopf equation and an innovations approach is used to derive a solution. The innovations approach presented here does not have the generality and the elegant mathematics of that of Kailath and Frost[13,14], but it uses some of their ideas. Four different versions of the smoothing solution and the associated covariance are presented.

The relations between the discrete system and the continuous system are examined and the continuous smoothing solutions are presented as the limiting case of the discrete ones. Finally, a problem is outlined involving a stochastic model for satellite drag, and the smoothing solution is applied to actual data.

2. DISCRETE TIME SMOOTHING

2.1 Statement of the Problem

The system to be estimated is described by a set of matrix difference equations,

$$x_{k+1} = \Phi_{k+1}x_k + u_k \quad \text{for} \quad k = 1, 2, \ldots N . \tag{2.1}$$

The measurements obtained from the system are defined by the equations

$$z_k = H_k x_k + v_k \quad \text{for} \quad k = 1, 2, \ldots N , \tag{2.2}$$

where x_k = an $n \times 1$ column vector of state variables .

u_k = an $n \times 1$ column vector of system (input) uncorrelated dynamic noise

v_k = an $r \times 1$ column vector of measurement (output) uncorrelated noise

z_k = an $r \times 1$ column vector of known measurements

Φ_{k+1} = the $n \times n$ transition matrix with known coefficients

H_k = an $r \times n$ matrix with known coefficients.

The dimensions of the measurement z_k can vary from point to point as long as the dimensions of the associated vectors and matrices also vary accordingly. The vectors u_k and v_k represent vector-valued independent random variables with zero mean and a known covariance.

$$\begin{aligned} E(u_k u_j') &= Q_k \quad \text{if} \quad k = j \quad \text{and zero otherwise} \\ E(v_k v_j') &= R_k \quad \text{if} \quad k = j \quad \text{and zero otherwise} \\ E(u_k v_j') &= 0 . \end{aligned} \tag{2.3}$$

In some treatments of this problem, it is assumed the vectors u_k and v_k are correlated if $k = j$. This slight modification will be discussed later. The matrices Q and R could be singular and the implications of this are also discussed later, although for the time being it will be assumed both matrices are non-singular.

In any real estimation problem there is an important distinction between *a priori* information and actual measurements. However, in order to simplify the notation and subsequent explanations, it will be convenient to treat the *a priori* information formally as the initial measurement z_0 , as shown below:

$$z_0 = x_0 + v_0 \quad \text{where} \quad z_0 = E[x_0]$$

and

$$E[v_0 v_0'] = R_0 = E[(z_0 - x_0)(z_0 - x_0)'] . \tag{2.4}$$

Thus, the *a priori* information about the initial value of the state x_0 is that it is a random variable with mean z_0 and covariance R_0 .

The state x_k is the quantity to be estimated. The best estimate of the state x_k , given the measurements z_j for $j = 0, 1, \ldots N$, is denoted by $\hat{x}_{k/N}$, and it is a linear combination of the measurements with coefficients given by the $n \times r$ weighting matrix $A_{k/j}$:

$$\hat{x}_{k/N} = \sum_{j=0}^{N} A_{k/j} z_j \,. \tag{2.5}$$

The estimation procedure is called filtering if k equals N, prediction if k is greater than N, and smoothing if k is less than N. The error in the estimate, denoted $\tilde{x}_{k/N}$, is the difference between the state and the best estimate of the state:

$$\tilde{x}_{k/N} = x_k - \hat{x}_{k/N} = x_k - \sum_{j=0}^{N} A_{k/j} z_j \,. \tag{2.6}$$

The partial derivative of the error with respect to the coefficients $A_{k/j}$ is equal to the measurement z_j,

$$\partial \tilde{x}_{k/N} / \partial A_{k/j} = - z_j \,. \tag{2.7}$$

The best estimate of the state will be defined as the linear estimate which minimizes the mean square error, which is the trace of the covariance of the error $P_{k/N}$:

$$P_{k/N} = E(\tilde{x}_{k/N} \tilde{x}'_{k/N}) \,. \tag{2.8}$$

Since the linear coefficients $A_{k/j}$ are chosen to minimize the mean square error, the condition for this to be a minimum is that the partial derivative with respect to the coefficients is zero,

$$\partial \mathrm{TR}(P_{k/N}) / \partial A_{k/j} = 0 \quad \text{for} \quad j = 0, 1, \ldots N \,. \tag{2.9}$$

Substituting the equation for the error into the covariance gives an expression for the mean square error which is quadratic in the coefficients. Carrying out the differentiation gives an expression which is linear in the coefficients - the sampled Wiener-Hopf equation - which is sufficient to determine the coefficients.

$$\partial \mathrm{TR}(P_{k/N}) / \partial A_{k/j} = 0 = -2E(\tilde{x}_{k/N} z'_j) \quad \text{for} \quad j = 0, 1, \ldots N \,. \tag{2.10}$$

This expression can be interpreted by saying the error in the best estimate, $\tilde{x}_{k/N}$, is uncorrelated with any of the measurements which made up that estimate. Since the best estimate is a linear combination of the measurements, this means the error in the best estimate is uncorrelated with the estimate itself,

$$E(\tilde{x}_{k/N} z'_j) = 0 \quad \text{for} \quad j = 0, 1, \ldots N$$

$$E(\tilde{x}_{k/N} \hat{x}'_{k/N}) = 0 \,. \tag{2.11}$$

2.2 Innovations Approach

The innovation is the new information which comes from each measurement. The innovations approach to linear estimation is based on the assumption that the actual measurements and the innovation process (all the innovations) are equal in the sense that they contain the same information as far as linear operations are concerned. The new information is the difference between the actual measurement z_k and the best estimate of the measurement, given all the previous measurements. The best estimate of the measurement is the output matrix H_k times the best estimate of the state $\hat{x}_{k/k-1}$ while the new information, denoted by $\tilde{z}_k$, is the error in this estimate,

$$\tilde{z}_k = z_k - H_k \hat{x}_{k/k-1} = H_k \tilde{x}_{k/k-1} + v_k$$

$$\text{since} \quad z_k = H_k x_k + v_k \,. \tag{2.12}$$

The best estimate of the measurement has the same properties as the best estimate of the state discussed previously, while the new information has the same properties as the error. The new information of the k^{th} measurement is uncorrelated with any previous measurements and the new information is uncorrelated with the estimate. These properties of the new information can be used to derive a sequential estimation procedure.

Assume the current best estimate $\hat{x}_{k/N}$ can be calculated as the sum of the previous best estimate $\hat{x}_{k/N-1}$ and a coefficient times the new information $\tilde{z}_N$:

$$\hat{x}_{k/N} = \hat{x}_{k/N-1} + B_{k/N} \tilde{z}_N \,. \tag{2.13}$$

If the coefficient $B_{k/N}$ can be chosen so the error in the current best estimate is uncorrelated with any of the measurements, it will indeed be the best estimate because it satisfies the sampled Wiener-Hopf equation. The error in the current best estimate is composed of two terms, the error in the previous best estimate, and the error in the new information:

$$\tilde{x}_{k/N} = x_k - \hat{x}_{k/N} = \tilde{x}_{k/N-1} - B_{k/N} \tilde{z}_k \,. \tag{2.14}$$

Because they are errors in best estimates, these two terms are uncorrelated with the previous measurements,

$$E(\tilde{x}_{k/N-1} z'_j) = 0$$

$$E(\tilde{z}_N z'_j) = 0 \quad \text{for} \quad j = 0, 1, \ldots N-1 \,. \tag{2.15}$$

For the current measurement z_N the condition that the error be uncorrelated determines the coefficient $B_{k/N}$.

$$0 = E(\tilde{x}_{k/N}\tilde{z}_N') = E(\tilde{x}_{k/N-1}\tilde{z}_N') - B_{k/N}E(\tilde{z}_N\tilde{z}_N') . \tag{2.16}$$

Therefore, the sequential estimation procedure in (2.13) gives the best estimate as long as the coefficient $B_{k/N}$ satisfies Equation (2.16). For practical computation a more convenient form for the coefficient is obtained by using the following relations:

$$\begin{aligned} E(\tilde{x}_{k/N-1}\tilde{z}_N') &= E(\tilde{x}_{k/N-1}\tilde{z}_N') = E(x_k\tilde{z}_N') \\ E(\tilde{z}_N z_N') &= E(\tilde{z}_N\tilde{z}_N') . \end{aligned} \tag{2.17}$$

These relations allow the coefficient $B_{k/N}$ to be written in the form

$$\begin{aligned} \hat{x}_{k/N} &= \hat{x}_{k/N-1} + B_{k/N}z_N \\ B_{k/N} &= E(x_k\tilde{z}_N')\,E(\tilde{z}_N\tilde{z}_N')^{-1} . \end{aligned} \tag{2.18}$$

The noise v_k is uncorrelated with the error in the best estimate so the covariance of the new information z_k can be calculated directly using (2.12),

$$\begin{aligned} E(z_k z_k') &= H_k E(\tilde{x}_{k/k-1}\tilde{x}_{k/k-1}')H_k' \\ + E(v_k v_k') &= H_k P_{k/k-1}H_k' + R_k . \end{aligned} \tag{2.19}$$

Multiplying the sequential error (2.14) by the state x_k gives a sequential procedure for calculating the covariance of the error,

$$\begin{aligned} E(\tilde{x}_{k/N}x_k') &= E(\tilde{x}_{k/N-1}x_k') - B_{k/N}E(\tilde{z}_N x_k') \\ P_{k/N} &= P_{k/n-1} - B_{k/N}E(\tilde{z}_N x_k') \\ &= P_{k/N-1} - B_{k/N}E(z_N\tilde{z}_N')B_{k/N}' \end{aligned}$$

since $$E(\tilde{x}_{k/j}x_k) = P_{k/j} . \tag{2.20}$$

To summarize, the sequential form of the optimum estimate and the associated covariance equation are given by (2.18) and (2.20). The form of the solution is quite general and does not depend on the system equations. The weighting coefficient $B_{k/N}$ involves two covariances. One of these is given in (2.19) with $k = N$, but the other covariance is as yet unknown. When $N = k - 1$ the above equations apply to filtering and prediction. When $N > k$ the equations are for fixed point smoothing because they show the change in the smoothed estimate at the k^{th} point.

2.3 Filtering and Fixed Point Smoothing

Up to now it has not been necessary to examine in detail the system equations for the state or the measurement. However, to calculate the coefficient $B_{k/N}$, for the sequential estimation procedure, it will be necessary to find a way of calculating the corresponding covariances. Consider first the single-step estimation problem in (2.18), where k is replaced by $k + 1$ and N is replaced by k . This is the familiar Kalman filter equation, where $B_{k+1/k}$ is the gain matrix for the filter,

$$\begin{aligned} \hat{x}_{k+1/k} &= \hat{x}_{k+1/k-1} + B_{k+1/k}z_k \\ B_{k+1/k} &= E(x_{k+1}\tilde{z}_k')\,E(\tilde{z}_k\tilde{z}_k')^{-1} . \end{aligned} \tag{2.21}$$

For one-step estimation the unknown covariance and the gain matrix can be determined directly.

$$\begin{aligned} E(x_{k+1}\tilde{z}_k') &= E[(\Phi_{k+1}x_k + u_k)(H_k\tilde{x}_{k/k-1} + v_k)'] \\ &= \Phi_{k+1}P_{k/k-1}H_k' \\ B_{k+1/k} &= \Phi_{k+1}P_{k/k-1}H_k'(H_k P_{k/k-1}H_k' + R_k)^{-1} . \end{aligned} \tag{2.22}$$

It should be pointed out here that, for the special case where the dynamic noise u_k and the measurement noise v_k are correlated, the unknown covariance has an additional term:

$$E(x_{k+1}\tilde{z}_k') = \Phi_{k+1}P_{k/k-1}H_k' + E(u_k v_k') .$$

To complete the solution to the one-step estimation problem, it is necessary to calculate a sequential relation for the covariance of the error $P_{k+1/k}$ which is used in the gain matrix. The best prediction must be the best estimate times the transition matrix,

$$\hat{x}_{k+1/k-1} = \Phi_{k+1}\hat{x}_{k/k-1} .$$

The error in the best estimate is

$$\begin{aligned} \tilde{x}_{k+1/k} &= x_{k+1} - \hat{x}_{k+1/k} \\ &= (\Phi_{k+1}x_k + u_k) - [\Phi_{k+1}\hat{x}_{k/k-1} + B_{k+1/k}(H_k\tilde{x}_{k/k-1} + v_k)] \\ &= \psi_{k+1}\tilde{x}_{k/k-1} + u_k - B_{k+1/k}v_k \ . \end{aligned}$$

where $$\psi_{k+1} = \Phi_{k+1} - B_{k+1/k}H_k \ . \qquad (2.23)$$

The covariance of the error is obtained by multiplying the error by the state and taking the expected value:

$$\begin{aligned} P_{k+1/k} &= E(\tilde{x}_{k+1/k}x'_{k+1}) \\ &= E\tilde{x}_{k+1/k}\,(\Phi_{k+1}x_k + u_k)' \\ &= \psi_{k+1}P_{k/k-1}\Phi'_{k+1} + Q_k \ . \end{aligned} \qquad (2.24)$$

The covariance for the current estimate can be written in a similar way,

$$P_{k/k} = \Phi^{-1}_{k+1}\psi_{k+1}P_{k/k-1} = (I - B_{k/k}H_k)P_{k/k-1} \ ,$$

where $$B_{k/k} = \Phi^{-1}_{k+1}B_{k+1/k}$$

and $$P_{k+1/k} = \Phi_{k+1}P_{k/k}\Phi'_{k+1} + Q_k \ . \qquad (2.25)$$

These are Kalman filter equations for the one-step estimation problem. To extend this to the smoothing problem, where N is greater than k, it is necessary to see how the errors in the estimate evolve. The matrix ψ_{k+1} in (2.23) shows how the estimation errors evolve from one point to the next point. Applying this to the points from k to N one can write

$$\begin{aligned} \tilde{x}_{N/N-1} &= \psi_N\psi_{N-1}\cdots\psi_{k+1}\tilde{x}_{k/k-1} + \\ &\quad + \text{terms involving } u_j \text{ and } v_j \text{ for } k \leq j \leq N-1 \ . \end{aligned} \qquad (2.26)$$

Since x_k is uncorrelated with the u_j and v_j, the unknown covariance can be written as

$$\begin{aligned} E(x_k\tilde{x}'_N) &= E[x_k(H_N\tilde{x}_{N/N-1} + v_N)'] \\ &= E(\tilde{x}_{k/k-1}\tilde{x}'_{N/N-1})H'_N \\ &= P_{k/k-1}\psi'_{N/k}H'_N \ . \end{aligned}$$

where $$\psi_{N/k} = \psi_N\psi_{N-1}\cdots\psi_{k+1}$$

and $$\psi_{j+1} = \Phi_{j+1} - B_{j+1/j}H_j \ . \qquad (2.27)$$

2.4 Other Smoothing Solutions

The fixed point smoothing solution given by (2.13), (2.19) and (2.27) will now be used to derive other recursive versions of the smoothing solution. At each point where the new information z_j is received, it is multiplied by the weighting matrix $B_{k/j}$ in (2.18). The smoothed estimates at points k and $k+1$ can be written as the weighted sum of all the new information,

$$\begin{aligned} \hat{x}_{k/N} &= \hat{x}_{k/k} + \sum_{j=k+1}^{N} B_{k/j}\tilde{z}_j \\ \hat{x}_{k+1/N} &= \hat{x}_{k+1/k} + \sum_{j=k+1}^{N} B_{k+1/j}\tilde{z}_j \ . \end{aligned} \qquad (2.28)$$

From (2.27) the previously unknown covariances and the weighting coefficient can be related to a new matrix C_k,

$$\begin{aligned} E(x_k\tilde{x}'_N) &= P_{k/k-1}\psi'_{k+1}\psi'_{N/k+1}H'_N \\ &= C_kP_{k+1/k}\psi'_{N/k+1}H'_N \\ &= C_kE(x_{k+1}\tilde{x}'_N) \ , \end{aligned}$$

where $$C_k = P_{k/k-1}\psi'_{k+1}P^{-1}_{k+1/k} \ . \qquad (2.29)$$

For the weighting coefficients in (2.28),

$$B_{k/j} = C_k B_{k+1/j} .$$

since $$B_{k/j} = E(x_k \tilde{z}_j') E(\tilde{z}_j \tilde{z}_j')^{-1} , \tag{2.30}$$

Subtracting C_k times the second equation in (2.28) from the first equation, and making use of (2.30) to cancel out the terms involving the new information, gives a relation between the smoothed estimate at points k and $k+1$,

$$\hat{x}_{k/N} - \hat{x}_{k/k} = C_k (\hat{x}_{k+1/N} - \hat{x}_{k+1/k}) . \tag{2.31}$$

This second version is called smoothing with filtered estimates. The smoothed estimate at k is calculated from the smoothed estimate at $k+1$ and the filtered estimate. In the same way, the smoothed covariance (2.20) can be written as the sum of terms

$$P_{m/N} = P_{m/k} - \sum_{j=k+1}^{N} B_{m/j} E(\tilde{x}_j \tilde{z}_j') B_{m/j}' \quad \text{for} \quad m = k \quad \text{and} \quad k+1 . \tag{2.32}$$

Premultiplying the equation with $m = k+1$ in (2.31), by C_k and post-multiplying by C_k', subtracting from the equation with $m = k$, and making use of (2.30) gives a relation between the covariance of the smoothed estimates at points k and $k+1$,

$$P_{k/N} - P_{k/k} = C_k (P_{k+1/N} - P_{k+1/k}) C_k' . \tag{2.33}$$

An alternative definition of the matrix C_k which is useful is obtained from (2.25),

$$C_k = P_{k/k} \Phi_{k+1}' P_{k+1/k}^{-1} ,$$

since $$P_{k/k} \Phi_{k+1}' = P_{k/k-1} \psi_{k+1}' . \tag{2.34}$$

A slightly different form of smoothing using the filtered estimates is obtained by grouping terms and making use of some identities, in (2.24). Multiplying (2.31) by Φ_{k+1} and subtracting $\hat{x}_{k/N}$ gives

$$\begin{aligned} \Phi_{k+1} \hat{x}_{k/N} - \hat{x}_{k+1/N} &= (\Phi_{k+1} C_k' - I)(\hat{x}_{k+1/N} - \hat{x}_{k+1/k}) \\ &= - Q_k P_{k+1/k}^{-1} (\hat{x}_{k+1/N} - \hat{x}_{k+1/k}) , \end{aligned}$$

since $$\begin{aligned} \Phi_{k+1} C_k &= \Phi_{k+1} P_{k/k-1} \psi_{k+1}' P_{k+1/k}^{-1} \\ &= (P_{k+1/k} - Q_k) P_{k+1/k}^{-1} . \end{aligned} \tag{2.35}$$

The smoothing solution with measurements, the third version, is written in a similar form, but it requires some detailed algebra, making use of complicated identities, to show the equivalence. The measurement smoothing involves a new vector ω_k which is calculated sequentially and, in the continuous version of the problem, can be interpreted as a Lagrange multiplier. The smoothing solution involving measurements is

$$\begin{aligned} \Phi_{k+1} \hat{x}_{k/N} - \hat{x}_{k+1/N} &= - Q_k \omega_k \\ \omega_{k-1} &= \Phi_{k+1}' \omega_k + H_k' R_k^{-1} (z_k - H_k \hat{x}_{k/N}) , \end{aligned}$$

with $$\omega_N = 0 . \tag{2.36}$$

Comparing (2.35) and (2.36) gives a direct interpretation of the vector ω_k, but no hint how to derive the recursion formula for ω_k,

$$\omega_k = P_{k+1/k}^{-1} (\hat{x}_{k+1/N} - \hat{x}_{k+1/k}) . \tag{2.37}$$

In order to derive the recursion formula it will be necessary to make use of the two relations

$$\begin{aligned} P_{k/k}^{-1} &= P_{k/k-1}^{-1} + H_k R_k^{-1} H_k \\ P_{k/k}^{-1} \hat{x}_{k/k} &= P_{k/k-1}^{-1} \hat{x}_{k/k-1} + H_k R_k^{-1} z_k . \end{aligned} \tag{2.38}$$

These relations arise naturally when using the maximum likelihood derivation, but they seem artificial here. The first relation can be proved by multiplying $P_{k/k}^{-1}$ by $P_{k/k}$ in (2.25) to get the identity. The second relation can be proved by showing that

$$B_{k/k} = P_{k/k} H_k' R_k^{-1} . \tag{2.39}$$

Starting with (2.37) and making use of the two relations as well as (2.31) gives the desired recursive relation for ω_k.

$$\begin{aligned}
\Phi'_{k+1}\omega_k &= \Phi'_{k+1}P^{-1}_{k+1/k}(\hat{x}_{k+1/N} - \hat{x}_{k+1/k}) \\
&= P^{-1}_{k/k}C_k(\hat{x}_{k+1/N} - \hat{x}_{k+1/k}) \\
&= P^{-1}_{k/k}(\hat{x}_{k/N} - \hat{x}_{k/k}) \\
&= (P^{-1}_{k/k-1} + H'_kR^{-1}_kH_k)\hat{x}_{k/N} - P^{-1}_{k/k-1}\hat{x}_{k/k-1} - H'_kR^{-1}_kz_k \\
&= P^{-1}_{k/k-1}(\hat{x}_{k/N} - \hat{x}_{k/k-1}) - H'_kR^{-1}_k(z_k - H_k\hat{x}_{k/N}) \\
&= \omega_{k-1} - H'_kR^{-1}_k(z_k - H_k\hat{x}_{k/N}) \,. \qquad (2.40)
\end{aligned}$$

The fourth and last version is called fixed lag smoothing because the smoothed estimate is always a fixed number of points, s, behind the latest measurement. It can be written as shown below where the vector ω_{k-1} can be defined either by (2.37) or by the recursive relation in (2.36):

$$\hat{x}_{k/N} = \Phi_k\hat{x}_{k-1/N-1} + B_{k/N}\bar{x}_N + Q_{k-1}\omega_{k-1} \,,$$

where $$k = N - s \,. \qquad (2.41)$$

This relation is obtained by replacing k by $k-1$ and N by $N-1$ in (2.36) and substituting the resulting expression for $\hat{x}_{k/N-1}$ into the fixed point solution (2.18). A recursive relation for the fixed lag covariance is obtained by replacing k by $k-1$ and N by $N-1$ in (2.33) and substituting the resulting expression for $P_{k/N-1}$ into the fixed point covariance (2.20).

$$P_{k/N} = P_{k/k-1} - D_{k/N}E(\bar{x}_N\bar{x}'_N)B'_{k/N} + C^{-1}_k(P_{k-1/N-1} - P_{k-1/k-1})(C^{-1}_k)' \,. \qquad (2.42)$$

Thus far, it has been assumed the matrices Q_k and R_k were non-singular. If the matrix Q_k is singular it will not change the form of the solutions, although if Q_k is zero the smoothing solutions simplify to

$$\hat{x}_{k/N} = \Phi^{-1}_{k+1}\hat{x}_{k+1/N} \,.$$

If the matrix R_k is singular the smoothing solution involving measurements cannot be used because it contains R^{-1}_k. For the filtering solution, it is necessary that

$$H_kP_{k/k-1}H'_k + R_k$$

be non-singular, so there may be a problem if both Q_k and R_k are singular.

3. CONTINUOUS TIME SMOOTHING

3.1 Statement of the Problem

The system to be estimated is described by a set of matrix differential equations

$$dx/dt = Fx + u \,, \qquad 0 \leq t \leq T \,. \qquad (3.1)$$

The continuous measurements obtained from the system are defined by the equation

$$z = Hx + v \,, \qquad 0 \leq t \leq T \,, \qquad (3.2)$$

where

x = an $n \times 1$ column vector of state variables

u = an $n \times 1$ column vector of system (input) uncorrelated dynamic noise

v = an $r \times 1$ column vector of measurement (output) uncorrelated noise

z = an $r \times 1$ column vector of continuous known measurements

F = the $n \times n$ dynamics matrix with known coefficients

H = the $r \times n$ output matrix with known coefficients.

The vectors u and v represent vector-valued continuous independent random variables (white noise) with zero mean and known covariance.

$$\begin{aligned}
E(u(t)u(s)') &= Q\delta(t-s) \\
E(v(t)v(s)') &= R\delta(t-s) \\
E(u(t)v(s)') &= 0 \,. \qquad (3.3)
\end{aligned}$$

The property of the delta function $\delta(t)$ is that the function is zero for $t \neq 0$ but the integral of $g(t)$ times the function is $g(0)$ if the integral includes the origin. The *a priori* information is that the initial value of the state, x_0, is a random variable with mean $\bar{x}_0$ and covariance $\bar{P}_0$,

$$E(x_0) = \bar{x}_0 \quad \text{and} \quad E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)'] = \bar{P}_0 . \tag{3.4}$$

The smoothed estimate of the state at time t, given all the measurements up to time T, is denoted by $\hat{x}_{t/T}$. The filtered estimate is denoted by $\hat{x}_{t/t}$ or simply $\hat{x}$. The covariances of the errors in the smoothed and filtered estimates are $P_{t/T}$ and $P_{t/t} = P$ respectively. The linear estimate is a linear combination of the *a priori* information and an integral of the measurements

$$\hat{x}_{t/T} = A_{t/0}\bar{x}_0 + \int_{s=0}^{s=T} A_{t/s} z \, ds . \tag{3.5}$$

The best estimate is the linear estimate which minimizes the integral of the mean square error,

$$\text{Minimize} \int_{t=0}^{t=T} E[(x - \hat{x}_{t/T})(x - \hat{x}_{t/T})'] \, dt . \tag{3.6}$$

3.2 Continuous System as Limiting Form of Discrete System

The discrete system can be considered as the sampled version of the continuous system while the continuous system can be considered as the limiting form of the discrete. These results will be used to derive the continuous best estimate as the limiting case of the discrete best estimate. First, consider the continuous system. The linear differential equation can be integrated from time t to time s to get

$$x(s) = \Phi(s,t)x(t) + u^*(s,t) ,$$

where
$$u^*(s,t) = \int_{\sigma=t}^{\sigma=s} \Phi(s,\sigma)u(\sigma) \, d\sigma , \tag{3.7}$$

and where $\Phi(s,t)$ is the transition matrix which satisfies the differential equations

$$d\Phi(s,t)/ds = F(s)\Phi(s,t)$$

$$\Phi(t,t) = I .$$

Let $s = t + \epsilon$, where ϵ is small, and there is an obvious analogy between the continuous system and the discrete system,

$$x_{k+1} = \Phi_{k+1}x_k + u_k .$$

$$x_{k+1} = x(t+\epsilon)$$

$$\Phi_{k+1} = \Phi(t+\epsilon,t)$$

$$u_k = u^*(t+\epsilon,t) \tag{3.8}$$

Since the covariance of the continuous dynamic noise $u(\sigma)$ is a delta function, the covariance of the random input u_k is calculated as follows:

$$E(u_k u_k') = Q_k = \int_t^{t+\epsilon} \Phi(t+\epsilon,\sigma)Q(\sigma)\Phi(t+\epsilon,\sigma)' \, d\sigma \simeq Q(t)\epsilon . \tag{3.9}$$

Thus, starting with a continuous system, one can take a series of samples and convert it to a discrete system. The linear equation for the measurements can also be integrated from time t to time $t + \epsilon$ to get an "average" value for the measurement. Assume for the moment that $H(t)$ and $x(t)$ are constant over that interval.

$$\bar{z}(t) = (1/\epsilon) \int_t^{t+\epsilon} z(\sigma) \, d\sigma = H(t)x(t) + v^*(t+\epsilon,t)$$

$$v^*(t+\epsilon,t) = (1/\epsilon) \int_t^{t+\epsilon} v(\sigma) \, d\sigma . \tag{3.10}$$

Actually, the integral of the change in $H(t)$ and $x(t)$ will be of order ϵ, but that can be neglected for small ϵ. The analogy between the continuous measurements and the discrete measurements is

$$z_k = \bar{z}(t)$$

$$v_k = v^*(t+\epsilon,t) . \tag{3.11}$$

The covariance of the noise is

$$E(v_k v_k') = R_k = (1/\epsilon^2) \int_t^{t+\epsilon} R(\sigma)\, d\sigma \simeq R(t)/\epsilon \,. \tag{3.12}$$

Next, consider the discrete system. The limiting form of the system difference equation is obtained by making a Taylor series expansion of the transition matrix for small ϵ and neglecting terms of order ϵ^2 denoted by $O(\epsilon^2)$,

$$x(t+\epsilon) = [I + F\epsilon + O(\epsilon^2)]\, x(t) + u(t)$$

$$\lim_{\epsilon \to 0} [x(t+\epsilon) - x(t)]/\epsilon = dx(t)/dt \,. \tag{3.13}$$

To summarize, the limiting form of the discrete system is obtained for small ϵ,

$$\lim_{\epsilon \to 0} [x_{k+1} - x_k]/\epsilon = dx/dt$$

$$\Phi_{k+1} \to I + F\epsilon$$

$$Q_k \to Q\epsilon$$

$$R_k \to R/\epsilon$$

$$H_k \to H$$

$$x_k \to x$$

$$z_k \to z \,. \tag{3.14}$$

3.3 Filtering and Smoothing Solutions

The continuous smoothing solutions are presented here as the limiting cases of the discrete solutions derived earlier. Applying the limiting process to the filtering solution in (2.21) and (2.22) and the covariance in (2.24) gives the continuous filtering solution.

$$dx/dt = F\hat{x} + B\tilde{z}_t$$

$$dP/dt = FP + PF' - BHP + Q$$

$$\text{with} \quad \tilde{z}_t = z - H\hat{x} \quad \text{and} \quad B = PH'R^{-1} \,, \tag{3.15}$$

where $\hat{x} = \hat{x}_{t/t}$ and $P = P_{t/t}$ and the initial conditions are

$$\hat{x}_0 = \bar{x}_0 \quad \text{and} \quad P_0 = \bar{P}_0 \,.$$

The number of terms in the continuous case is fewer than that in the discrete case because

$$[H_k P_{k/k-1} H_k' + R_k]^{-1} \to [HPH' + (R/\epsilon)]^{-1} \to R^{-1}\epsilon \,. \tag{3.16}$$

Sometimes in a discrete problem it might be convenient to integrate the discrete difference equations for the covariance on a digital computer by changing them to the corresponding differential equations. As the limiting process shows, one requirement for this approximation to be valid is that

$$R_k \gg H_k P_{k/k-1} H_k' \,.$$

Applying the limiting process to the fixed point smoothing solution in (2.18), (2.19) and (2.27) and the corresponding covariance in (2.20) gives the continuous fixed point smoothing solution.

$$d\hat{x}_{t/T}/dT = D(T,t)\tilde{z}_T$$

$$dP_{t/T}/dT = -D(T,t)RD(T,t)' \,,$$

where $D(T,t) = P_{t/t}^{-1}\psi(T,t)H'R^{-1}$ and

$\psi(T,t)$ = the matrix which satisfies the differential equation

$$d\Psi(T,t)/dT = (F - BH)\Psi(T,t)$$

$$\Psi(t,t) = I \,. \tag{3.17}$$

Applying the limiting process to the smoothing solution involving the filtered estimate in (2.31) and the corresponding covariances in (2.33) gives the smoothing solution for the continuous case,

$$d\hat{x}_{t/T}/dt = F\hat{x}_{t/T} + QP^{-1}(\hat{x}_{t/T} - \hat{x}_{t/t})$$

$$dP_{t/T}/dt = (F+QP^{-1})P_{t/T} + P_{t/T}(F+QP^{-1})' - Q \,. \tag{3.18}$$

These differential equations must be integrated backwards from T to 0 and the beginning conditions at $t = T$ are the terminal conditions on the filter equations $\hat{x}_{T/T}$ and $P_{T/T}$. The limiting form of the matrix C_k has been written as

$$C_k \to I - \epsilon(F+QP^{-1}) \,. \tag{3.19}$$

Applying the limiting process to the filtering solution involving the measurements in (2.36) gives the smoothing solution for the continuous case

$$dx_{t/T}/dt = Fx_{t/T} + Q\omega$$

$$d\omega/dt = -F'\omega - H'R^{-1}(z - H\hat{x}_{t/T}) \,. \tag{3.20}$$

Once again the differential equations are integrated backwards from T to 0 and the beginning conditions are the terminal conditions on the filter equations $\hat{x}_{T/T}$ and $\omega(T) = 0$.

Finally, applying the limiting process to the fixed lag solution of s seconds gives

$$d\hat{x}_{t/T}/dT = F\hat{x}_{t/T} + D(T,t)\tilde{z}_T + Q\omega$$

$$dP_{t/T}/dT = (F+QP^{-1})P_{t/T} + P_{t/T}(F+QP^{-1})' - Q - D(T,t)RD(T,t)' \,,$$

where $$t = T - s \,. \tag{3.21}$$

4. APPLICATION – STOCHASTIC DRAG MODEL

4.1 Satellite Tracking with Stochastic Drag

This section presents an example of smoothing on satellite tracking where a stochastic model is used to improve accuracy. The practical problems which arise in smoothing are essentially the same as those in filtering. A model for the system equations must be developed as well as a model for the measurements. Values must be chosen for the mean and covariance of the dynamic noise and the measurement noise and for the *a priori* estimate and its covariance. The primary difference between practical problems in filtering and those in smoothing is in the effect of including dynamic noise on the model. If there is no dynamic noise in the model the system is deterministic and the smoothing solution is obtained directly from the final value of the filtered solution. When there is dynamic noise, the smoothing solution cannot be obtained so easily and may require significant additional work. Correlated noise on the measurements can have a similar effect to dynamic noise. To handle correlated noise, the state equations are augmented by including additional state variables for the correlated part of the noise.

The problem of trajectory estimation involves reconstructing the position history of a space vehicle from a series of measurements such as range, angles, or Doppler velocity. When there is only gravity acceleration, due to engine thrust or air drag, the motion is deterministic and can be completely specified by six parameters. These six might be the position and velocity at some initial time, or a set of six orbit elements at the time of the ascending node. From the six parameters, the current and future motion can be determined by numerical integration of the acceleration due to the earth's gravitational field or from a closed-form approximation to the integrated motion.

In general, when drag is included, the assumption is usually made that its effect can also be represented as a deterministic phenomenon (for instance, as a polynomial with undetermined coefficients). This treatment may be adequate as long as drag is not a significant factor in the satellite motion. However, for low altitude satellites, stochastic fluctuations in atmospheric density can cause noticeable changes in the motion.

The particular stochastic model presented here is based on the assumption that the statistical properties of the atmosphere can be adequately represented by a first-order stationary Gauss-Markov process. Therefore, the satellite motion can be determined by eight parameters that are a set of six mean orbit elements, a seventh parameter that is related to the constant average value of the drag, and an eighth parameter that is related to the instantaneous deviation of the stochastic drag from its mean value. A similar model could be developed when there is acceleration due to engine thrust. The stochastic fluctuations in thrust might be due to changes in the direction of thrust or the engine performance.

4.2 Simplified Model for Analysis

The main effect of the drag acceleration is to change the in-track motion of the satellite. The satellite arrives at some position in space a little sooner or later, depending on whether the drag has been a little larger or smaller than expected. The simplified discrete model of the motion will involve four state variables – the in-track angular position y_1, the in-track angular velocity y_2, the mean drag acceleration y_3, and the stochastic drag acceleration y_4.

The continuous stochastic acceleration $b(t)$ is assumed to be a zero-mean first-order Markov process with exponential correlation and a time constant $1/c$. The differential equation describing the stochastic acceleration has stationary zero-mean white noise $u(t)$ as the excitation,

$$\begin{aligned} db(t)/dt &= -cb(t) + u(t) \\ E[u(t)u(t+s)] &= q\delta(s) . \end{aligned} \tag{4.1}$$

The discrete version of the stochastic acceleration is given by a difference equation,

$$\begin{aligned} b(t+T) &= b(t)\exp(-cT) + u^* \\ u^* = u^*(t+T,t) &= \int_{s=t}^{s=t+T} \exp[-c(T-s)]\, u(s)\, ds \\ E[u^*u^*] &= \int_{s=0}^{s=T} \exp[-2c(T-s)]\, q\, ds \\ &= q[1-\exp(-2cT)]/2c . \end{aligned} \tag{4.2}$$

The discrete stochastic acceleration y_4 changes every T seconds when it takes on the value of the continuous variable $b(t)$. The system of differential equations for the angular motion are

$$\begin{aligned} dy_1/dt &= y_2 \\ dy_2/dt &= y_3 + y_4 \\ dy_3/dt &= 0 \\ y_4 &= b(kT) \quad \text{for} \quad kT < t < (k+1)T . \end{aligned} \tag{4.3}$$

Every T seconds noisy measurements of position are made. The difference equations for the discrete system are

$$\begin{aligned} x_{k+1} &= \Phi_{k+1}x_k + u_k \\ z_k &= H_k x_k + v_k \\ E(v_k) &= 0 , \qquad E(v_k v_k') = R_k \\ x_k' &= (y_1, y_2, y_3, y_4) \\ u_k' &= (0,0,0,u^*) \\ \Phi_{k+1} &= \begin{bmatrix} 1 & T & T^2/2 & T^2/2 \\ 0 & 1 & T & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \exp(-cT) \end{bmatrix} \\ H_k' &= (1,0,0,0) . \end{aligned} \tag{4.4}$$

The covariance of the *a priori* information is diagonal:

$$E(x_0x_0') = \begin{bmatrix} p & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & q \end{bmatrix} . \tag{4.5}$$

For the numerical examples the sample time T will be held fixed at one revolution ($T = 1$), the correlation time $1/c$ will be two revolutions ($c = 1/2$) and the accuracy of the measurements will be set equal to unity. Three cases will be examined. The covariances of the stochastic drag q and the *a priori* information p for the three cases are

Case	q	p
1	10^{-2}	1
2	10^{-4}	1
3	10^{-2}	100

In each case 25 measurements are taken, starting with n_1. The diagonal elements of the covariance of the estimates of the state for case 1 are presented in Table I for both filtering and smoothing. Notice how smoothing decreases the error substantially. In Figure 1 the variances of the filtered and smoothed estimates for the angular position y_1 are plotted for both case 1 and case 2. In both cases the variance approaches a steady-state value where the information lost by dynamic noise balances the information gained from additional measurements. Reducing the variance of the stochastic drag reduces the steady-state variance of the estimates. In Figure 2, the variances for the angular position are plotted for both case 1 and case 3. Notice how the initial conditions (the *a priori* information) have a significant effect on the first few estimates, although that rapidly dies out.

4.3 Numerical Results with Real Data

The stochastic drag model has been tested by using data taken during the first three days of satellite 1960 Omicron[18]. The satellite was launched on November 12, 1960, with a period of 90.4 min, and was up for only 47 days. The apogee height, perigee height, and inclination were 614 statute miles, 113 statute miles, and 81.9°, respectively. Over the first three days the average constant drag was 24×10^{-4} fractional decay of the semi-major axis per revolution, which corresponds to a period decay of about 0.2 sec/rev. This satellite was chosen because fluctuations in atmospheric density over the first two days caused a drag change of over 50% from the first day to the second, and because there were a large number of highly accurate Baker-Nunn optical sightings of the satellite. The total data consist of 13 passes over radar stations and six sets of optical sightings, spaced so that there is a gap of about 9 hours between the first and second day, when there were no measurements. (There was also a seventh set of optical sightings, which was discarded because it was completely inconsistent, by scores of miles, with all the rest of the data.)

Results were obtained using both optical and radar data but, in the results presented here, the radar data were used to determine the orbit, whereas the optical data were used as a check. This approach was particularly appropriate because the optical data were highly accurate (perhaps to less than one hundred feet) and also because the radar data were concentrated mainly in the northern hemisphere, whereas the optical data were mainly in the southern hemisphere. With a given set of data, the quantities that can be varied are the initial conditions (the initial estimate and its covariance), the r.m.s. drag of the stochastic model (σ), and the correlation of the stochastic model (θ_0). It was desired to eliminate the effect of the initial conditions, so they were chosen with a rather large covariance and with an estimate which matched the data from the first few radar stations. For all the results presented here the initial conditions were held fixed, while only the r.m.s. drag and correlation of the model were changed.

Since the most important effect of drag is to change the time equation, it is the in-track residuals that should be examined. Because the radar range measurements are so much more accurate than the angle measurements, the in-track residuals in position are determined by comparing the in-track position at the time when the range reaches a minimum (the point of closest approach). The weighted r.m.s. in-track residuals for radar data alone are plotted in Figure 3 as a function of r.m.s. drag, with the correlation held fixed at two revolutions (about 3 hr). For very low r.m.s. drag, the stochastic model reduces to the deterministic model and the weighted r.m.s. residuals approach the value obtained using the deterministic model (6800 ft). Notice that the r.m.s. residuals for the stochastic model reach a minimum at about one-tenth of the value obtained using the deterministic model. For very high r.m.s. drag, the forward estimation procedure tended to be marginally stable. The high value of stochastic drag was used to overcorrect the estimate, first one way then the other. The r.m.s. in-track residuals for radar data alone are plotted again in Figure 4 as a function of the correlation, with the r.m.s. drag held fixed at 4×10^{-7}. The results show that the r.m.s. residuals were not very sensitive to the correlation. Once again, as the correlation gets very large, the stochastic model will approach the deterministic model. The estimate of the stochastic drag is plotted in Figure 5 as a function of revolutions for two values of the correlation with σ held fixed at 4×10^{-7}. Notice how the larger value of the correlation tends to mask the rapid fluctuations in the drag.

A point-by-point comparison of the in-track residuals at each radar pass is presented in Table II for the deterministic model with r.m.s. drag zero and the stochastic model with the r.m.s. drag 4×10^{-7} and the correlation 10 rev. Comparing the two sets of results shows that the stochastic treatment of drag has reduced the residuals from the deterministic treatment using constant drag by an order of magnitude, from thousands of feet to hundreds of feet. In order to check the radar results, the optical data were included with zero weight, i.e., the residuals of the optical sightings were calculated, but no differential corrections were made. The residuals at three of the optical stations were less than 200 ft and at the fourth station, following the gap of 9 hr with no measurements, the residuals were 1500 ft. The prediction of the optical measurements after 3 rev and 6 rev following the end of the radar data gave residuals of 28,000 and 70,000 ft, respectively, indicating that there still might be some stochastic fluctuations on the third day. Also a capsule was separated from the satellite on the 31 orbit, and this may have influenced the drag characteristics.

REFERENCES

1. Gauss, K.F. — *Theory of the Motion of Heavenly Bodies Moving About the Sun in Conic Sections*. Translated by C.H.Davis, Dover Publications, New York, 1963.

2. Wiener, N. — *The Extrapolation, Interpolation, and Smoothing of Stationary Time Series with Engineering Applications*. Wiley, New York, 1949.

3. Kalman, R.E. — *A New Approach to Linear Filtering and Prediction Problems*. Transactions, American Society of Mechanical Engineers, Journal of Basic Engineering, Vol.82, March 1960, pp.34-35.

4. Kalman, R.E. — *New Methods in Wiener Filtering Theory*. Proceedings, First Symposium on Engineering Applications of Random Function Theory and Probability. J.L.Bogdanoff and F.Kozin (eds), Wiley, New York, 1963.

5. Bucy, R.S. — *Optimum Finite-Time Filters for a Special Non-Stationary Class of Inputs*. Johns Hopkins University, Applied Physics Laboratory, Baltimore, Md., Internal Memo BBD-600, 1959.

6. Kalman, R.E., Bucy, R.S. — *New Results in Linear Filtering and Prediction Theory*. Transactions, American Society of Mechanical Engineers, Journal of Basic Engineering, Series D, Vol.83, December 1961, pp.95-107.

7. Bryson, A.E., Frazier, M. — *Smoothing for Linear and Nonlinear Dynamic Systems*. Aeronautical Systems Division, Wright-Patterson AFB, Ohio, Tech. Report ASD-TDR-63-119, February 1963.

8. Cox, H. — *On the Estimation of State Variables and Parameters for Noisy Dynamic Systems*. Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-9, January 1964, pp.5-12.

9. Rauch, H.E., et al. — *Maximum Likelihood Estimates of Linear Dynamic Systems*. AIAA Journal, Vol.3, August 1965, pp.1445-1450.

10. Lee, R.C.K. — *Optimal Estimation, Identification and Control*. MIT Press, Cambridge, Mass., 1964.

11. Meditch, J.S. — *Orthogonal Projection and Discrete Optimal Linear Smoothing*. Journal, Society for Industrial and Applied Mathematics, Vol.5, February 1967, pp.74-89.

12. Fraser, D.C. — *A New Technique for the Optimal Smoothing of Data*. ScD Dissertation, Department of Aeronautical Engineering, Massachusetts Institute of Technology, Cambridge, Mass., January 1967.

13. Kailath, T. — *An Innovations Approach to Least Squares Estimation – Part I: Linear Estimation in Additive White Noise*. Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-13, December 1968, pp.646-655.

14. Kailath, T. Frost, P. — *An Innovations Approach to Least Squares Estimation – Part II: Linear Smoothing in Additive White Noise*. Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.AC-13, December 1968, pp.655-660.

15. Rauch, H.E. — *Optimum Estimation of Satellite Trajectories Including Random Fluctuations in Drag*. AIAA Journal, Vol.3, April 1965, pp.717-722.

TABLE I

Diagonal Elements of the Covariance

Observation Point (k)	*Filtered Estimate* ($P_{k/k}$)			
	$cov(y_1)$	$cov(y_2)$	$cov(y_3)$	$cov(y_4)$
0	1.00	1.00	1.00	0.0100
1	0.69	1.31	0.92	0.0100
2	0.80	1.31	0.54	0.0100
3	0.82	0.96	0.36	0.0100
4	0.79	0.68	0.13	0.0100
5	0.75	0.40	0.07	0.0100
10	0.58	0.15	0.008	0.00995
15	0.50	0.10	0.004	0.0099
20	0.48	0.093	0.0026	0.0099
25	0.47	0.089	0.0020	0.0099
Observation Point (k)	*Smoothed Estimate* ($P_{k/N}$)			
	$cov(y_1)$	$cov(y_2)$	$cov(y_3)$	$cov(y_4)$
25	0.47	0.089	0.0020	0.0099
24	0.26	0.058	0.0020	0.0096
23	0.18	0.036	0.0020	0.0091
22	0.15	0.023	0.0020	0.0085
21	0.15	0.017	0.0020	0.0080
20	0.15	0.015	0.0020	0.0078
15	0.135	0.014	0.0020	0.0076
10	0.135	0.014	0.0020	0.0076
5	0.14	0.015	0.0020	0.0076
1	0.26	0.033	0.0020	0.0089
0	0.45	0.082	0.0020	0.0094

TABLE II

Comparisons of In-Track Residuals, 10^3 ft

	Data[a]	*Revolution*[b]	*Deterministic*[c]	*Stochastic*[d]
1	R	0.4	44.6	2.7
2	O	4.0	...	0.1
3	R	5.1	5.7	0.9
4	R	6.1	0.3	-0.3
5	R	7.1	-1.6	0.5
6	R	8.1	-3.0	0.4
7	R	11.4	-4.4	0.2
8	R	13.4	2.1	0.4
9	O	19.0	...	1.5
10	R	19.1	2.6	0.7
11	O	20.1	...	0.2
12	R	20.1	1.6	0.4
13	O	21.1	...	0.1
14	R	21.1	0.6	-0.1
15	R	22.1	2.6	2.1
16	R	23.1	0.6	0.3
17	R	28.4	4.6	0.1
18	O	31.1	...	28.0
19	O	34.1	...	70.0

[a]Optical = O and radar = R.

[b]Since time of first ascending node.

[c]Radar data with constant drag model.

[d]Radar data with zero weight on optical (for $\sigma = 0.4 \times 10^{-6}$ $\theta_0 = 10$ rev).

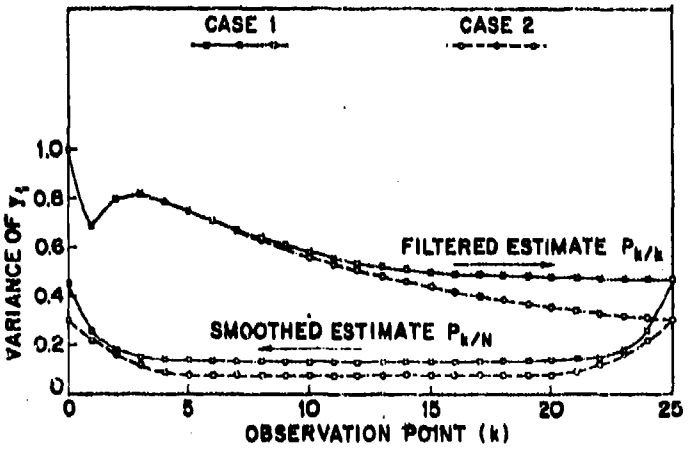


Fig. 1 Variance history for two levels of stochastic disturbances (q)

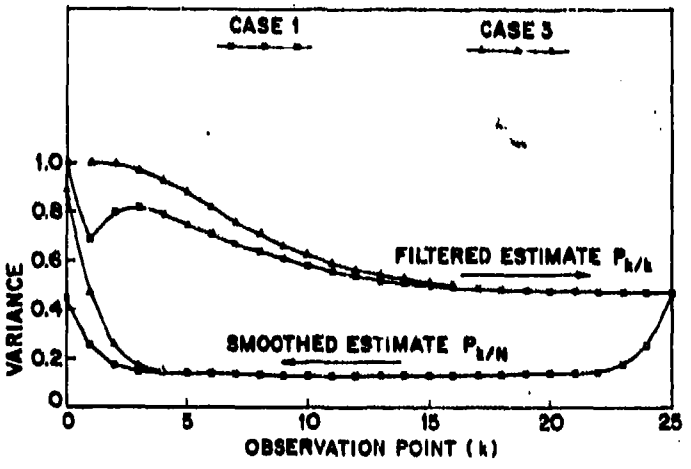


Fig. 2 Variance history for two sets of initial conditions (p)

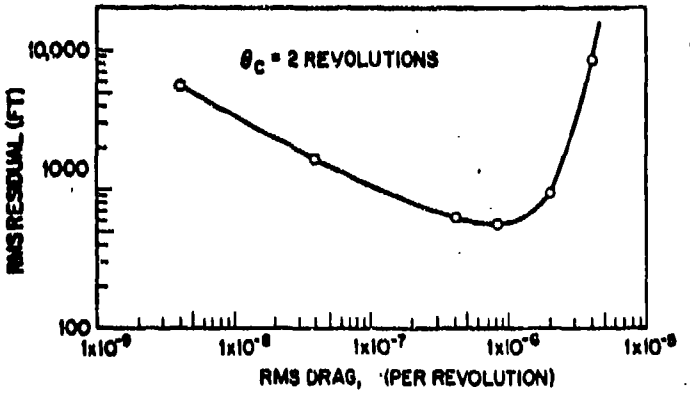


Fig. 3 Root mean square in-track residuals versus root mean square stochastic drag

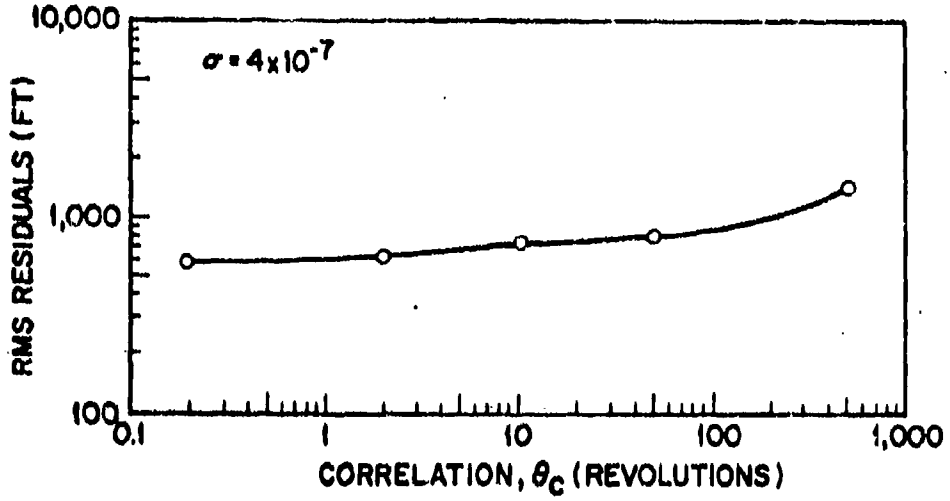


Fig. 4 Root mean square in-track residuals versus correlation

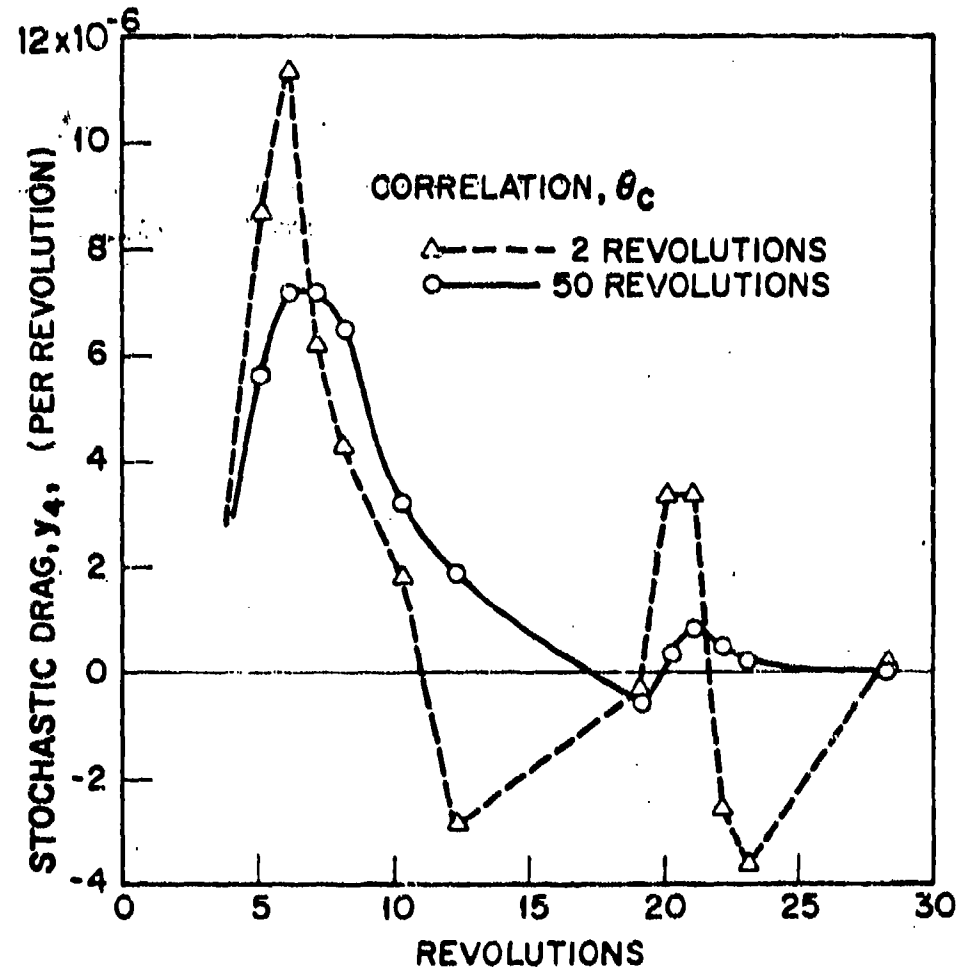


Fig. 5 Estimated stochastic drag versus revolutions

CHAPTER 9 - NONLINEAR SMOOTHING TECHNIQUES

by

John B. Peller

1260 Arbolita Drive
La Habra, California 90631, USA

CHAPTER 9 - NONLINEAR SMOOTHING TECHNIQUES

John B. Peller

1. INTRODUCTION

The so-called "nonlinear smoothing" problem is the problem of estimating the state of some noisy process at some time t, given noisy measurements related to the process over an entire measurement interval which includes the time t. The nonlinear smoothing problem differs from the linear smoothing problem in that the differential equations describing the process and the observations can be nonlinear as well as linear. A typical nonlinear smoothing problem is the post-flight estimation of the flight path of a missile based on tracking system measurements made during the entire duration of the flight. If the estimate of the missile's position and velocity at the midpoint of the flight is desired, the estimate can be based on all of the measurements made, including those made after the midpoint as well as those made before. This is the smoothing problem. If either the differential equations describing the flight of the missile or the measurements made to determine the actual flight path (or both) are nonlinear, the problem of estimating the missile's position and velocity is called a nonlinear smoothing problem. In this chapter, methods of solving the nonlinear smoothing problem will be presented.

Smoothing problems (or any estimation problems for that matter) may be classified according to the criterion by which the quality of the estimate is to be judged. In this chapter, as in the rest of this publication, only probabalistic criteria will be considered. For the smoothing problem, this is a marked departure from practice; normally, smoothing is performed by statistical techniques such as fitting a polynomial to a set of observations in a least squares fashion without considering the known properties of the noise in the measurements or in the plant.

In order to describe the classifications, some definitions are necessary. Let $\underline{X}(t)$ be a vector function of time representing the state of the process or plant. Let $\underline{y}(\underline{X},t)$ be another vector function of both the state $\underline{X}$ and time t; it will be called the observation process. Suppose the process runs from time $t = 0$ to the time $t = T$, and that the observation $\underline{y}(\underline{X},t)$ is recorded for all t in the interval from 0 to T. The smoothing problem is to estimate $\underline{X}(t)$ for some t in the interval by operating on all the observations in the interval. Denote the estimate by $\hat{\underline{X}}(t)$.

If $\hat{\underline{X}}(t)$ is selected so that the probability $\hat{\underline{X}}(t) = \underline{X}(t)$ is maximized, the estimate $\hat{\underline{X}}(t)$ is called the maximum likelihood estimate of $\underline{X}(t)$. If $\hat{\underline{X}}(t)$ is selected so that the covariance of the estimate $E([\underline{X}(t) - \hat{\underline{X}}(t)][\underline{X}(t) - \hat{\underline{X}}(t)]^T)$ (where $E(.)$ is the expectation operator) is minimized, the estimate is called the minimum variance estimate. Minimum variance estimates are particularly useful in missile accuracy analysis. For a linear plant with additive gaussian white noise and linear observations with additive gaussian white noise, the state of the plant $\underline{X}(t)$ is a random process with a gaussian distribution, and any estimate of $\underline{X}(t)$ derived by a linear operation on the observations will also be a gaussian random variable. For this case, both the maximum likelihood estimate and the minimum variance estimate are equal to the mean of a particular gaussian distribution derived by such a linear operation. Consequently, for any kind of linear estimation theory, it is unnecessary to differentiate between the maximum likelihood estimate and the minimum variance estimate.

For a nonlinear plant and nonlinear observations, the probability density function of the estimate conditioned on all the observations is not generally gaussian. In this case, the maximum likelihood estimate (which can be shown to be equal to that value of $\underline{X}(t)$ for which the subject density function has its maximum value) will differ from the minimum variance estimate (which can be shown to be equal to the mean value or first moment of the subject density function). Thus, the procedure used to solve the nonlinear smoothing problem will differ depending upon which type of estimate is desired. Primary emphasis in this chapter will be placed on the problem of finding the minimum variance estimate, because this estimate is more useful in fields such as missile accuracy analysis and other probabalistic applications of the smoothing results. However, a particular approach for use in obtaining maximum likelihood estimates will also be discussed.

Minimum variance smoothing theory will be covered in Sections 2 and 3 of this chapter. In Section 2, "exact" differential equations for the "smoothed expectation" of an arbitary function of the state are derived. In Section 3, these differential equations are approximated by other differential equations which are easily solved on a computer to obtain the minimum variance estimate. Section 4 presents a brief discussion of maximum likelihood nonlinear smoothing.

2. EXACT DIFFERENTIAL EQUATION FOR THE SMOOTHED EXPECTATION

In Section 2.1, the problem is formulated, and initial steps are taken towards deriving the desired exact differential equations. In particular, a differential form is derived for the probability density function of .

the state conditioned on all the observations. In subsequent subsections, adjoint theory and various expansions are employed to convert this to an exact differential equation for the expectation of an arbitrary function of the state, conditioned on all the observations.

2.1 Problem Statement and the Smoothing Density Function

Let the system be represented by the n-dimensional state vector $\underline{x}(t)$, and let the observations $\underline{y}(\underline{x},t)$ be represented by the m-dimensional vector $\underline{z}(t)$. Let these vectors satisfy the following equations:

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x},t) + \dot{\underline{\zeta}}(t) \tag{2.1}$$

and

$$\dot{\underline{z}}(t) = \underline{g}(\underline{x},t) + \dot{\underline{\eta}}(t) , \tag{2.2}$$

where $\dot{\underline{\zeta}}(t)$ and $\dot{\underline{\eta}}(t)$ are formally gaussian white noise processes. Equations (2.1) and (2.2) are considered to be formalisms of the following equations:

$$d\underline{x}(t) = \underline{f}(\underline{x},t)\,dt + d\underline{\zeta}(t) + r , \tag{2.3}$$

$$d\underline{z}(t) = \underline{g}(\underline{x},t)\,dt + d\underline{\eta}(t) + r , \tag{2.4}$$

where r denotes any terms for which the expectations satisfy

$$E(r) = O(dt) , \tag{2.5}$$

$$E(r^2) = O(dt^2) , \tag{2.6}$$

where $O(.)/dt \to 0$ as $dt \to 0$. The function $\underline{f}(\underline{x},t)$ is assumed to be piecewise continuous in t and to possess piecewise continuous first derivatives in $\underline{x}$. No explicit assumptions concerning $\underline{g}(\underline{x},t)$ are made herein. However, assumptions will be made concerning the process $\underline{z}(t)$ and the availability of a filtering density function, the realization of which may constrain $\underline{g}(\underline{x},t)$ to have certain properties, in particular, to be piecewise continuous in both $\underline{x}$ and t. Both functions, $\underline{f}(\underline{x},t)$ and $\underline{g}(\underline{x},t)$, may be nonlinear.

The independent variable t is assumed to be constrained to some set, closed at least on the right, denoted by (t_0,T). It is possible for t_0 to be minus infinity, but T is finite. Let Z_{t_1,t_2} denote the set of random variables $\underline{z}(t)$ for $t_1 \le t \le t_2$:

$$Z_{t_1,t_2} = \{\underline{z}(t) \,;\, t_1 \le t \le t_2\} . \tag{2.7}$$

For notational ease, $Z_{t_0,T}$ will simply be denoted by Z_T. Assume that the conditional distribution of $\underline{x}(t)$, given Z_T, exists and is absolutely continuous. This implies that the conditional density of $\underline{x}(t)$, given Z_T, exists. This density is denoted by $p_{\underline{x}(t)}[\underline{\mu}|Z_T]$, where $\underline{\mu}$ is an n-dimensional argument. Because of the extensive use of this particular density, it is denoted by

$$s(\underline{\mu},t) = p_{\underline{x}(t)}[\underline{\mu}|Z_T] \tag{2.8}$$

and is referred to as the smoothing density function. The numerous occurrences of the density function $p_{\underline{x}(t)}[\underline{\mu}|Z_t]$ make it desirable to give it a special symbol and name. It will therefore be referred to as the filtering density function and denoted as

$$q(\underline{\mu},t) = p_{\underline{x}(t)}[\underline{\mu}|Z_t] . \tag{2.9}$$

As the first step in obtaining a differential form for the smoothing density function $s(\underline{\mu},t)$, the smoothing density function is written in the form

$$p_{\underline{x}(t)}[\underline{\mu}|Z_T] = \int_{-\infty}^{\infty} p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t+\Delta t) = \underline{\nu}\,, Z_T]\, p_{\underline{x}(t+\Delta t)}[\underline{\nu}|Z_T]\, d\nu , \tag{2.10}$$

where the integration is over all n-components of $\underline{\nu}$ (this convention will be used throughout).

At this point, the one and only assumption concerning the observation process is made. The observation process is assumed to be separable[1]. Separability implies that bounds taken over dense sets of points in (t_0,T) are the same as those obtained over the continuous parameter space, and that limits which are valid when approached sequentially are still valid when approached continuously and vice-versa. There is no effective loss of generality to assume that some process is separable, because if some process $\underline{z}(t)$ is not separable, then there exists a process $\bar{\underline{z}}(t)$ which is separable and which equals $\underline{z}(t)$ with probability one:

$$P\{\underline{z}(t) = \bar{\underline{z}}(t)\} = 1 . \tag{2.11}$$

With the separability hypothesis, the process $\underline{z}(t)$ is replaced with a discrete process $\check{\underline{z}}(t)$ which is equal to $\underline{z}(t)$ for t belonging to the set of rationals and equal to any constant for t belonging to the set of irrationals. Thus, only when t is a rational number does $\check{\underline{z}}(t)$ convey any information concerning $\underline{x}(t)$

Because the rationals are countable, $\underline{x}(t)$ can be replaced with $\underline{x}(i)$ and, momentarily replacing t with i, Equation (2.10) becomes

$$p_{\underline{x}(i)}[\underline{\mu}|Z_\tau] = \int_{-\infty}^{\infty} p_{\underline{x}(i)}[\underline{\mu}|\underline{x}(i+1) = \underline{\nu}, Z_\tau] p_{\underline{x}(i+1)}[\underline{\nu}|Z_\tau]\, d\nu . \tag{2.12}$$

The first term in the integral can be written as

$$p_{\underline{x}(i)}[\underline{\mu}|\underline{x}(i+1) = \underline{\nu}, Z_\tau] = p_{\underline{x}(i)}[\underline{\mu}|\underline{x}(i+1) = \underline{\nu}, Z_i, Z_{i+1,\tau}] . \tag{2.13}$$

Because of the gaussian white noise assumption, Equations (2.3) and (2.4) represent Markov processes. Using the Markov property that the past and future are independent when conditioned on the present,

$$p_{\underline{x}(i)}[\underline{\mu}|\underline{x}(i+1) = \underline{\nu}, Z_i, Z_{i+1,\tau}] = p_{\underline{x}(i)}[\underline{\mu}|\underline{x}(i+1) = \underline{\nu}, Z_i] . \tag{2.14}$$

Substituting Equations (2.13) and (2.14) in (2.12),

$$p_{\underline{x}(i)}[\underline{\mu}|Z_\tau] = \int_{-\infty}^{\infty} p_{\underline{x}(i)}[\underline{\mu}|\underline{x}(i+1) = \underline{\nu}, Z_i] p_{\underline{x}(i+1)}[\underline{\nu}|Z_\tau]\, d\nu . \tag{2.15}$$

By Bayes' rule for conditional densities,

$$p_{\underline{x}(i)}[\underline{\mu}|\underline{x}(i+1) = \underline{\nu}, Z_i] = \frac{p_{\underline{x}(i)}[\underline{\nu}|\underline{x}(i+1) = \underline{\mu}, Z_i] p_{\underline{x}(i)}[\underline{\mu}|Z_i]}{p_{\underline{x}(i+1)}[\underline{\nu}|Z_i]} . \tag{2.16}$$

The denominator of Equation (2.16) can be evaluated by

$$p_{\underline{x}(i+1)}[\underline{\nu}|Z_i] = \int_{-\infty}^{\infty} p_{\underline{x}(i+1)}[\underline{\nu}|\underline{x}(i) = \underline{\mu}, Z_i] p_{\underline{x}(i)}[\underline{\mu}|Z_i] d\mu . \tag{2.17}$$

By the Markov property,

$$p_{\underline{x}(i+1)}[\underline{\nu}|\underline{x}(i) = \underline{\mu}, Z_i] = p_{\underline{x}(i+1)}[\underline{\nu}|\underline{x}(i) = \underline{\mu}] . \tag{2.18}$$

Using Equation (2.18), Equations (2.16) and (2.17) can be rewritten, respectively:

$$p_{\underline{x}(i)}[\underline{\mu}|\underline{x}(i+1) = \underline{\nu}, Z_i] = \frac{p_{\underline{x}(i+1)}[\underline{\nu}|\underline{x}(i) = \underline{\mu}] p_{\underline{x}(i)}[\underline{\mu}|Z_i]}{p_{\underline{x}(i+1)}[\underline{\nu}|Z_i]} \tag{2.19}$$

and

$$p_{\underline{x}(i+1)}[\underline{\nu}|Z_i] = \int_{-\infty}^{\infty} p_{\underline{x}(i+1)}[\underline{\nu}|\underline{x}(i) = \underline{\mu}] p_{\underline{x}(i)}[\underline{\mu}|Z_i]\, d\mu . \tag{2.20}$$

Thus, Equation (2.15) becomes

$$p_{\underline{x}(i)}[\underline{\mu}|Z_\tau] = \frac{\int_{-\infty}^{\infty} p_{\underline{x}(i+1)}[\underline{\nu}|\underline{x}(i) = \underline{\mu}] p_{\underline{x}(i)}[\underline{\mu}|Z_i] p_{\underline{x}(i+1)}[\underline{\nu}|Z_\tau]\, d\nu}{p_{\underline{x}(i+1)}[\underline{\nu}|Z_i]} . \tag{2.21}$$

By the separability hypothesis, t and $t + dt$ can be substituted for i and $i + 1$, respectively. Doing this and using Equations (2.8) and (2.9) in Equation (2.21),

$$s(\underline{\mu}, t) = q(\underline{\mu}, t) \int_{-\infty}^{\infty} \frac{p_{\underline{x}(t+dt)}[\underline{\nu}|\underline{x}(t) = \underline{\mu}]}{p_{\underline{x}(t+dt)}[\underline{\nu}|Z_t]} s(\underline{\nu}, t+dt)\, d\nu \tag{2.22}$$

or

$$s(\underline{\mu}, t) = q(\underline{\mu}, t) \frac{\int_{-\infty}^{\infty} p_{\underline{x}(t+dt)}[\underline{\nu}|\underline{x}(t) = \underline{\mu}] s(\underline{\nu}, t+dt)\, d\nu}{\int_{-\infty}^{\infty} p_{\underline{x}(t+dt)}[\underline{\nu}|\underline{x}(t) = \underline{\mu}] q(\underline{\mu}, t)\, d\mu} . \tag{2.23}$$

Note that Equation (2.23) implies that the smoothing density function is not explicitly dependent on the observation process in any manner. Of course, the filter density function $q(\underline{\mu}, t)$ would generally be expressed as a function of the observation vector and its statistics, so that the smoothing density function is implicitly a function of the observation process through the filter density function.

2.2 Differential Equation for the Smoothed Expectation

In this section, preliminary steps are taken to obtain a differential equation for the "smoothed expectation" of an arbitrary function of the state. First, some definitions will be given. The smoothed expectation of an arbitrary function of the state, $h(\underline{x})$, is defined as:

$$E_s h[\underline{x}(t)] = \int_{-\infty}^{\infty} h(\underline{\mu}) s(\underline{\mu},t)\, d\mu . \tag{2.24}$$

For later use, the conditional filtered expectation of an arbitrary function $h(x)$ is defined as

$$E_f h[\underline{x}(t)] = \int_{-\infty}^{\infty} h(\underline{\mu}) q(\underline{\mu},t)\, d\mu . \tag{2.25}$$

Defining

$$dE_s h[\underline{x}(t)] = E_s h[\underline{x}(t)] - E_s h[\underline{x}(t-dt)] \tag{2.26}$$

and formally dividing by dt,

$$\frac{dE_s h[\underline{x}(t)]}{dt} = \frac{1}{dt}\left[E_s h[\underline{x}(t)] - E_s h[\underline{x}(t-dt)]\right] . \tag{2.27}$$

The objective of this section is to obtain an expression of the term on the right side of Equation (2.27).

By definition (Equation (2.24)),

$$E_s h[\underline{x}(t-dt)] = \int_{-\infty}^{\infty} h(\underline{\mu}) p_{\underline{x}(t-dt)}[\underline{\mu}|Z_\tau]\, d\mu \tag{2.28}$$

Substituting Equation (3.22) in Equation (2.28),

$$E_s h[\underline{x}(t-dt)] = \int_{-\infty}^{\infty} d\mu h(\underline{\mu}) \int_{-\infty}^{\infty} p_{\underline{x}(t-dt)}[\underline{\mu}|Z_{t-dt}] \frac{p_{\underline{x}(t)}[\underline{\nu}|\underline{x}(t-dt) = \underline{\mu}]}{p_{\underline{x}(t)}[\underline{\nu}|Z_{t-dt}]} p_{\underline{x}(t)}[\underline{\nu}|Z_\tau]\, d\nu . \tag{2.29}$$

The smoothing density function in Equation (2.29) is evaluated at time t, as is the corresponding density function in $E_s h[\underline{x}(t)]$. This fact is used to advantage by deriving from Equations (2.27) and (2.29) an equation of the form

$$\frac{dE_s h[\underline{x}(t)]}{dt} = E_s\{K_1 h[\underline{x}(t)]\} , \tag{2.30}$$

where K_1 is an operator yet to be determined. The simplest and quickest method of achieving this is first to note that the E_s operator corresponds to a scalar product and then to use the theory of adjoints. Denote the scalar product as

$$E_s h[\underline{x}(t-dt)] = (h(\underline{\mu}),\ s(\underline{\mu}\ ,\ t-dt)) . \tag{2.31}$$

If Equation (2.22) is written in the form

$$s(\underline{\mu}\ ,\ t-dt) = K_2 s(\underline{\mu},t) , \tag{2.32}$$

where K_2 is the operator indicated by Equation (2.22), then Equation (2.31) can be written as

$$E_s h[\underline{x}(t-dt)] = (h(\underline{\mu}),\ K_2 s(\underline{\mu},t)) . \tag{2.33}$$

Denoting the adjoint of K_2 by K_2^*, Equation (2.33) becomes

$$E_s h[\underline{x}(t-dt)] = (K_2^* h(\underline{\mu}),\ s(\underline{\mu},t)) , \tag{2.34}$$

where $K_2^* h(\underline{\mu})$ is given by

$$K_2^* h(\underline{\mu}) = \int_{-\infty}^{\infty} h(\underline{\nu}) \frac{p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu}] p_{\underline{x}(t-dt)}[\underline{\nu}|Z_{t-dt}]}{p_{\underline{x}(t)}[\underline{\mu}|Z_{t-dt}]} d\nu . \tag{2.35}$$

Note that the argument of $K_2^* h$ is the argument after applying K_2^* to h.

Equation (2.27) can now be rewritten as

$$\frac{dE_s h[\underline{x}(t)]}{dt} = \frac{1}{dt}\left[\langle h(\underline{\mu}),\ s(\underline{\mu},t)\rangle - \langle K_2^* h(\underline{\mu}),\ s(\underline{\mu},t)\rangle\right] \tag{2.36}$$

and, by the linearity of scalar products,

$$\frac{dE_s h[\underline{x}(t)]}{dt} = \frac{1}{dt}\langle h(\underline{\mu}) - K_2^* h(\underline{\mu}),\ s(\underline{\mu},t)\rangle\ . \tag{2.37}$$

Equation (2.37) is the basic differential equation for the results to follow. It is necessary to evaluate the operator K_2^* first, and some necessary preliminary steps are presented in Section 2.3.

2.3 Expansion of a Particular Density Function

Before proceeding with the evaluation of the operator K_2^*, it is desirable to find a different expression for the density function $p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu}]$. This is accomplished in this section.

Because the integration in Equation (2.35) is over $\underline{\nu}$, this equation can be rewritten as:

$$K_2^* h(\underline{\mu}) = \frac{1}{p_{\underline{x}(t)}[\underline{\mu}|Z_{t-dt}]}\int_{-\infty}^{\infty} h(\underline{\nu})\, p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu}]\, p_{\underline{x}(t-dt)}[\underline{\nu}|Z_{t-dt}]\, d\nu\ . \tag{2.38}$$

The term $p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu}]$ is evaluated by consideration of the basic plant equation (Eqn. (2.3)) and the properties of the independent increment process. From Equation (2.3),

$$d\underline{\zeta}(t-dt) = \underline{x}(t) - \underline{x}(t-dt) - \underline{f}(\underline{x},\ t-dt)\ dt + r\ . \tag{2.39}$$

The Jacobian of this transformation is unity, so that, by first applying the Markov property and then Equation (2.39),

$$\begin{aligned} p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu}] &= p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu},\ Z_{t-dt}] \\ &= P_{d\underline{\zeta}(t-dt)}[\underline{\mu} - \underline{x}(t-dt) - \underline{f}(\underline{x},\ t-dt)\ dt + r]\ . \end{aligned} \tag{2.40}$$

Let

$$E[d\underline{\zeta}(t)] = \underline{0} \tag{2.41}$$

$$E[d\underline{\zeta}(t)d\underline{\zeta}'(t)] = Q(t)dt + O(dt)\ . \tag{2.42}$$

Then the characteristic function becomes

$$E(e^{i\underline{\omega}'d\underline{\zeta}(t)}) = \exp\left\{-\frac{dt}{2}\,\underline{\omega}'Q(t)\underline{\omega}\right\} + O(dt)\ . \tag{2.43}$$

With the aid of the inverse Fourier transform,

$$p_{d\underline{\zeta}(t)}[\underline{\mu}] = \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}\exp\left\{-i\underline{\omega}'\underline{\mu} - \frac{dt}{2}\,\underline{\omega}'Q(t)\underline{\omega}\right\} d\omega + O(dt)\ . \tag{2.44}$$

Combining Equations (2.40) and (2.44),

$$p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu}] = \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}\exp\left\{-i\underline{\omega}'[\underline{\mu}-\underline{\nu}-\underline{f}(\underline{\nu},\ t-dt)dt + r] - \frac{dt}{2}\,\underline{\omega}'Q(t)\underline{\omega}\right\} d\omega + O(dt)\ . \tag{2.45}$$

It is desirable to expand Equation (2.45) in a power series at this point:

$$\begin{aligned} p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu}] = \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}&\exp[-i\underline{\omega}'(\underline{\mu}-\underline{\nu})]\{1 + i\underline{\omega}'[\underline{f}(\underline{\nu},\ t-dt)\ dt]\} \times \\ &\times\left\{1 - \frac{dt}{2}\,\underline{\omega}'Q(t)\underline{\omega}\right\} d\omega + O(dt)\ . \end{aligned} \tag{2.46}$$

The following identities involving the Dirac delta function are easily verified:

$$\delta(\underline{\mu}-\underline{\nu}) = \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}\exp\{-i\underline{\omega}'(\underline{\mu}-\underline{\nu})\}\,d\omega\,, \tag{2.47}$$

$$\frac{\partial\delta(\underline{\mu}-\underline{\nu})}{\partial\mu_j} = \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}-i\omega_j\exp\{-i\underline{\omega}'(\underline{\mu}-\underline{\nu})\}\,d\omega\,, \tag{2.48}$$

$$\frac{\partial^2\delta(\underline{\mu}-\underline{\nu})}{\partial\mu_j\partial\mu_k} = \frac{1}{(2\pi)^n}\int_{-\infty}^{\infty}-\omega_j\omega_k\exp\{-i\underline{\omega}'(\underline{\mu}-\underline{\nu})\}\,d\omega\,. \tag{2.49}$$

By the properties of the Dirac delta function,

$$\int_{-\infty}^{\infty} f(x,y)\,\delta(x-w)\,dx = f(z,y)\,. \tag{2.50}$$

Because of the ways Equations (2.48) and (2.49) were obtained,

$$\int_{-\infty}^{\infty}\frac{\partial\delta(\underline{\mu}-\underline{\nu})}{\partial\mu_j}f(\underline{\nu})\,d\nu = \frac{\partial}{\partial\mu_j}\int_{-\infty}^{\infty}\delta(\underline{\mu}-\underline{\nu})f(\underline{\nu})\,d\nu \tag{2.51}$$

and

$$\int_{-\infty}^{\infty}\frac{\partial^2\delta(\underline{\mu}-\underline{\nu})}{\partial\mu_j\partial\mu_k}f(\underline{\nu})\,d\nu = \frac{\partial^2}{\partial\mu_j\partial\mu_k}\int_{-\infty}^{\infty}\delta(\underline{\mu}-\underline{\nu})f(\underline{\nu})\,d\nu\,. \tag{2.52}$$

With the aid of Equations (2.47) - (2.49), Equation (2.46) becomes

$$p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu}] = \delta(\underline{\mu}-\underline{\nu}) - \sum_{j=1}^{n}[f_j(\underline{\nu},\,t-dt)\,dt]\frac{\partial\delta(\underline{\mu}-\underline{\nu})}{\partial\mu_j} + \frac{dt}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\frac{\partial^2\delta(\underline{\mu}-\underline{\nu})}{\partial\mu_j\partial\mu_k} + O(dt)\,, \tag{2.53}$$

which is the desired result of this section.

2.4 Completion of Smoothed Expectation Equation

The derivation of the exact differential equation for the smoothed expectation of $h(x)$ can now be completed. By using Equation (2.17), Equation (2.38) becomes

$$K_2^{\dagger}h(\underline{\mu}) = \frac{\int_{-\infty}^{\infty}p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu}]q(\underline{\nu},\,t-dt)h(\underline{\nu})\,d\nu}{\int_{-\infty}^{\infty}p_{\underline{x}(t)}[\underline{\mu}|\underline{x}(t-dt) = \underline{\nu}]q(\underline{\nu},\,t-dt)\,d\nu}\,. \tag{2.54}$$

For convenience, the numerator and denominator of Equation (2.54) will be denoted by Q and P, respectively. With the aid of Equation (2.53), Q can be written as

$$Q = \int_{-\infty}^{\infty}d\nu h(\underline{\nu})q(\underline{\nu},\,t-dt)\left[\delta(\underline{\mu}-\underline{\nu}) - \sum_{j=1}^{n}[f_j(\underline{\nu},\,t-dt)\,dt]\times\right.$$

$$\left.\times\frac{\partial\delta(\underline{\mu}-\underline{\nu})}{\partial\mu_j} + \frac{dt}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\frac{\partial^2\delta(\underline{\mu}-\underline{\nu})}{\partial\mu_j\partial\mu_k}\right] + O(dt)\,. \tag{2.55}$$

Now straightforward application of Equations (2.50) - (2.52) yields

$$Q = h(\underline{\mu})q(\underline{\mu},\,t-dt) - dt\left\{\sum_{j=1}^{n}\frac{\partial}{\partial\mu_j}\left[h(\underline{\mu})q(\underline{\mu},\,t-dt)[f_j(\underline{\mu},\,t-dt)]\right] -\right.$$

$$\left. - \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2}{\partial\mu_j\partial\mu_k}[h(\underline{\mu})q(\underline{\mu},\,t-dt)Q_{jk}(t)]\right\} + O(dt)\,. \tag{2.56}$$

P is simply evaluated by letting $h(\underline{\mu}) = 1$ in Equation (2.56),

$$P = q(\underline{\mu}, t-dt) - dt\left\{\sum_{j=1}^{n}\frac{\partial}{\partial\mu_j}\left[q(\underline{\mu}, t-dt)f_j(\underline{\mu}, t-dt)\right] - \right.$$

$$\left. - \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2}{\partial\mu_j\partial\mu_k}\left[Q_{jk}(t)q(\underline{\mu}, t-dt)\right]\right\} + O(dt) . \qquad (2.57)$$

Let $()_Q$ denote the collection of terms in brackets $()$ in Equation (2.56). Let $()_P$ have a similar meaning for P, as given by Equation (2.57). Then

$$\frac{1}{P} = \frac{1}{q(\underline{\mu}, t-dt)}\,\frac{1}{1 - \dfrac{dt()_P}{q(\underline{\mu}, t-dt)} + O(dt)} . \qquad (2.58)$$

Assuming $()_P/q(\underline{\mu}, t-dt)$ is bounded, $1/P$ can be written as

$$\frac{1}{P} = \frac{1}{q(\underline{\mu}, t-dt)}\left(1 + \frac{dt()_P}{q(\underline{\mu}, t-dt)} + O(dt)\right) . \qquad (2.59)$$

Substituting $()_Q$ and $()_P$ as given in Equations (2.56) and (2.57), respectively, and simplifying, Equation (2.54) becomes

$$\frac{Q}{P} = h(\underline{\mu}) + dt\left\{-\sum_{j=1}^{n}\left[\underline{f}(\underline{\mu}, t-dt)\right]_j\frac{\partial h(\underline{\mu})}{\partial\mu_j} + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\frac{\partial^2 h(\underline{\mu})}{\partial\mu_j\partial\mu_k} + \right.$$

$$\left. + \sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial h(\underline{\mu})}{\partial\mu_j}\frac{\partial Q_{jk}(t)}{\partial\mu_k} + \frac{1}{q(\underline{\mu}, t-dt)}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\frac{\partial h(\underline{\mu})}{\partial\mu_j}\frac{\partial q(\underline{\mu}, t-dt)}{\partial\mu_k}\right\} + O(dt) . \qquad (2.60)$$

In performing the simplifications leading to Equation (2.60), no use was made of the fact that Q is not a function of $\underline{\mu}$. This fact will become important later. Since

$$\frac{1}{q(\underline{\mu}, t-dt)} = \frac{1}{q(\underline{\mu},t) - \dfrac{\partial q(\underline{\mu},t)}{\partial t}dt + O(dt)}$$

$$= \frac{1}{q(\underline{\mu},t)}\left[1 + \frac{\partial q(\underline{\mu},t)}{\partial t}dt + O(dt)\right] \qquad (2.61)$$

and

$$f_j(\underline{\mu}, t-dt) = f_j(\underline{\mu},t) - \frac{\partial f_j(\underline{\mu},t)}{\partial t}dt + O(dt) , \qquad (2.62)$$

it is apparent that Equation (2.60) is still valid when all arguments $t - dt$ are replaced by t. That is, the difference between the two expressions is of $O(dt)$.

By definition,

$$K_0^* h(\underline{\mu}) = \frac{Q}{P} . \qquad (2.63)$$

Comparing Equations (2.30) and (2.37), it is obvious that

$$K_1 f = \frac{1}{dt}(I - K_0^*)f . \qquad (2.64)$$

Equations (2.63) and (2.64) yield

$$K_1 h(\underline{\mu}) = \sum_{j=1}^{n}\left[\underline{f}(\underline{\mu},t)\right]_j\frac{\partial h(\underline{\mu})}{\partial\mu_j} - \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\frac{\partial^2 h(\underline{\mu})}{\partial\mu_j\partial\mu_k} -$$

$$- \sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial h(\underline{\mu})}{\partial\mu_j}\frac{\partial Q_{jk}(t)}{\partial\mu_k} - \frac{1}{q(\underline{\mu},t)}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\frac{\partial h(\underline{\mu})}{\partial\mu_j}\frac{\partial q(\underline{\mu},t)}{\partial\mu_k} + O(dt) . \qquad (2.65)$$

Equation (2.30),

$$\frac{dE_s h[\underline{x}(t)]}{dt} = E_s\{k_j h[\underline{x}(t)]\} , \tag{2.66}$$

combined with Equation (2.65) is the desired differential equation for the smoothed expectation of an arbitrary function of x.

If the substitution

$$h(\underline{\mu}) = \delta(\underline{\mu}-\underline{\nu}) \tag{2.67}$$

is made in both sides of Equation (2.66), the result is

$$\begin{aligned}\frac{ds(\underline{\mu},t)}{dt} = &- s(\underline{\mu},t)\sum_{j=1}^{n}\frac{\partial[\underline{f}(\underline{\mu},t)]_j}{\partial\mu_j} - \sum_{j=1}^{n}[\underline{f}(\underline{\mu},t)]_j + \frac{\partial s(\underline{\mu},t)}{\partial\mu_j} + \\ &+ \frac{1}{2}s(\underline{\mu},t)\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial^2 Q_{jk}(t)}{\partial\mu_j\partial\mu_k} - \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\times \\ &\times\frac{\partial^2 s(\underline{\mu},t)}{\partial\mu_j\partial\mu_k} + \frac{1}{q(\underline{\mu},t)}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\frac{\partial s(\underline{\mu},t)\partial q(\underline{\mu},t)}{\partial\mu_j\partial\mu_k} - \\ &- \frac{s(\underline{\mu},t)}{q^2(\underline{\mu},t)}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\frac{\partial q(\underline{\mu},t)}{\partial\mu_j}\frac{\partial q(\underline{\mu},t)}{\partial\mu_k} + \frac{s(\underline{\mu},t)}{q(\underline{\mu},t)}\sum_{j=1}^{n}\sum_{k=1}^{n}\frac{\partial Q_{jk}(t)}{\partial\mu_j}\frac{\partial q(\underline{\mu},t)}{\partial\mu_k} + \\ &+ \frac{s(\underline{\mu},t)}{q(\underline{\mu},t)}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\frac{\partial^2 q(\underline{\mu},t)}{\partial\mu_j\partial\mu_k} + O(dt) ,\end{aligned} \tag{2.68}$$

which is the differential equation for the smoothing density function.

It is possible to derive a differential equation for the smoothed expectation of an arbitrary function of the state in terms of filtered expectations. Striebel outlined a derivation of such an equation in Reference 2, based on the limit of a discrete formulation. Peller[3] used a similar method based on the formal techniques with differentials used herein. The equation is presented here without derivation (see Peller[3]):

$$\frac{dE_s h[\underline{x}(t)]}{dt} = E_f\left\{L_2\left[\frac{s(\underline{x},t)}{q(\underline{x},t)}\,h(\underline{x})\right]\right\} - E_f\left\{h(\underline{x})L_2\frac{s(\underline{x},t)}{q(\underline{x},t)}\right\} + O(dt) , \tag{2.69}$$

where L_2 is the backward diffusion operator:

$$L_2 h(\underline{\nu}) = \sum_{j=1}^{n} f_j(\underline{\nu}\ ,\ t-dt)\frac{\partial}{\partial\nu_j}h(\underline{\nu}) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}Q_{jk}(t)\frac{\partial^2 h(\underline{\nu})}{\partial\nu_j\partial\nu_k} . \tag{2.70}$$

If the substitution

$$h(\underline{\mu}) = \delta(\underline{\mu}-\underline{\nu}) \tag{2.71}$$

is made in Equation (2.69), the result is the differential equation for the smoothing density function as already given by Equation (2.68), as is to be expected.

Equations (2.65), (2.66), and (2.68) are the principal results; they represent the "exact" differential equations satisfied by the smoothing expectation and smoothing density function. The boundary condition for Equation (2.68) is simply

$$s(\underline{\mu},T) = q(\underline{\mu},T) , \tag{2.72}$$

and the boundary condition for Equations (2.65) and (2.66) is given by the obvious relation

$$E_s\{h[\underline{x}(T)]\} = E_f\{h[\underline{x}(T)]\} . \tag{2.73}$$

As was mentioned earlier, no use was made of the fact that Q is not a function of $\underline{\mu}$. In fact, had use been made of this, it is obvious that one of the terms in Equation (2.65) would have been equal to zero. Because the results remain valid, these "exact" equations can easily be extended to cover the product noise case. First, for conciseness, Equation (2.65) is rewritten to show explicitly the possible dependence of Q on the system state:

$$K_1 h(\underline{\mu}) = \sum_{j=1}^{n} [\underline{f}(\underline{\mu},t)]_j \frac{\partial h(\underline{\mu})}{\partial \mu_j} - \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} Q_{jk}(\underline{\mu},t) \frac{\partial^2 h(\underline{\mu})}{\partial \mu_j \partial \mu_k} -$$

$$- \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial h(\underline{\mu})}{\partial \mu_j} \frac{\partial Q_{jk}(\underline{\mu},t)}{\partial \mu_k} - \frac{1}{q(\underline{\mu},t)} \sum_{j=1}^{n} \sum_{k=1}^{n} Q_{jk}(\underline{\mu},t) \frac{\partial h(\underline{\mu})}{\partial \mu_j} \frac{\partial q(\underline{\mu},t)}{\partial \mu_k} . \tag{2.74}$$

Now suppose

$$d\underline{\zeta}(t) = M(\underline{x},t) d\underline{\omega}(t) , \tag{2.75}$$

where $\underline{\omega}(t)$ is an n-vector increment in an independent increment process and $M(\underline{x},t)$ is an $n \times n$ matrix. Further suppose that

$$E[d\underline{\omega}(t) d\underline{\omega}'(t)] = \Omega(t)dt + O(dt) , \tag{2.76}$$

then, by letting

$$Q(\underline{x},t) = M(\underline{x},t)\Omega(t)M'(\underline{x},t) , \tag{2.77}$$

Equations (2.66) and (2.74) apply directly, and the product noise case is accommodated.

2.5 Comments on the Exact Differential Equations

Equations (2.65), (2.66), and (2.68) together with the boundary conditions given by Equations (2.72) and (2.73) are in theory sufficient to solve any smoothing problem. In fact, Equations (2.68) and (2.72) are sufficient in themselves for this task since once the smoothing density function is found, the smoothed expectation of any function of the system state can be found by integrating that function multiplied by the smoothing density function over all infinity in parameter space.

Although these equations are exact in the mean square sense, they are impossible to solve in practice and they must be replaced by solvable differential equations whose solutions are "approximately" equal to the solutions of these exact equations. The reasons that these equations are impossible to solve in practice are as follows. Equation (2.68) for the smoothing density function is a second-order partial differential equation in all components of the state vector whose domain is all infinity for this problem. The computer requirements that result from this type of problem far exceed what is considered practical by today's standards in computer technology. In addition, it is necessary to have the filtering density function, $q(\mu,t)$, available in analytical form or to have its first and second derivatives with respect to all state vector components available in any form. This is virtually never true in practice.

Now that the difficulties in solving for the smoothing density function are known, a brief note is in order on the difficulties in directly solving the equations for the smoothed expectation. First, note that the filtering density function and its first derivative with respect to all state vector components are required, something we have already seen to be almost never true. Second, note that the smoothing operator on the right side of Equation (2.66) operates on each term in Equation (2.65). This requires the evaluation of the smoothed expectation of new terms. If these are computed by again applying Equations (2.65) and (2.66), still more new terms are introduced. In fact, if the process is continued, a countably infinite number of terms must be evaluated.

Thus, it is clearly impossible to solve directly for the smoothed expectation, or smoothing density function, and approximations which are more easily solved must be sought. In the next section, a particular set of approximations are developed which permit the determination of the minimum variance estimate. In a following section, a method of obtaining a maximum likelihood estimate will be described. This method is based on an entirely different procedure and will be presented without derivation.

3. MINIMUM VARIANCE SMOOTHING APPROXIMATIONS

For linear systems, the smoothing density function is gaussian and is therefore completely described by the smoothed mean and covariance. The case of a nonlinear system is complicated by the fact that the smoothing density function is not necessarily gaussian, and, consequently, the smoothed mean and covariance may not completely describe it. In fact, as mentioned previously, a countably infinite number of parameters are necessary to completely describe smoothed parameters in the general case.

Because of this, approximations involving a finite number of parameters are sought. In the case of nonlinear filtering theory, two general approaches are most evident. One approach is to approximate the conditioned expectation of a function by a Taylor series truncated after the quadratic terms. This is the approach used by Bucy[4], Bass, Norum, and Schwartz[5], and Schwartz[6]. It is equivalent to assuming certain higher order moments to be negligible. The second approach is to use quasi-moments. This approach, used by Fisher[7], allows the conditional density function to be expressed in terms of the conditional mean, conditional covariance, and quasi-moment functions. Both approaches can be used for the nonlinear smoothing problem; the truncated Taylor series approach will be presented in this section. The quasi-moment technique has greater generality, but results

in very complex expressions when applied to the smoothing problem. Section 4.2 very briefly summarizes the idea of quasi-moments. (See Peller[3] for a more extensive discussion of the quasi-moment technique in nonlinear smoothing.)

3.1 Nonlinear Minimum Variance Smoothing

For minimum variance smoothing, the conditional mean of the state is the desired estimate; thus, in Equations (2.64) and (2.65),

$$h_i[\underline{x}(t)] = x_i(t), \qquad i = 1, \ldots, n, \tag{3.1}$$

so that

$$\bar{\bar{\underline{x}}}(t) = \begin{bmatrix} E_s h_1 \\ \vdots \\ E_s h_n \end{bmatrix}, \tag{3.2}$$

where the double overline indicates the minimum variance smoothed estimate. From this point on, it is assumed that Q_{jk} is independent of the state. Using this assumption and substituting Equations (3.1), (2.1) and (2.64) into Equation (2.65),

$$\frac{d\bar{\bar{x}}_i(t)}{dt} = \overline{f_i(\underline{x},t)} - \sum_{k=1}^{n} Q_{ik}(t) \overline{\frac{1}{q(\underline{x},t)} \frac{\partial q(\underline{x},t)}{\partial x_k}}. \tag{3.3}$$

Equation (3.3) is exact in the mean-square sense (and to within $O(dt)$). Approximations are now desired for each term on the right of Equation (3.3).

To approximate $\overline{f_i(\underline{x},t)}$, expand $f_i(\underline{x},t)$ in a Taylor series about $\bar{\bar{\underline{x}}}$ and drop terms of greater than second order:

$$f_i(\underline{x}) \doteq f_i(\bar{\bar{\underline{x}}}) + \sum_{j=1}^{n} \frac{\partial f_i(\bar{\bar{\underline{x}}})}{\partial x_j}(x_j - \bar{\bar{x}}_j) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 f_i(\bar{\bar{\underline{x}}})}{\partial x_j \partial x_k}(x_j - \bar{\bar{x}}_j)(x_k - \bar{\bar{x}}_k), \tag{3.4}$$

where

$$\frac{\partial f_i(\bar{\bar{\underline{x}}})}{\partial x_j} = \left.\frac{\partial f_i(\underline{x})}{\partial x_j}\right|_{\underline{x}=\bar{\bar{\underline{x}}}} \tag{3.5}$$

$$\frac{\partial^2 f_i(\bar{\bar{\underline{x}}})}{\partial x_j \partial x_k} = \left.\frac{\partial^2 f_i(\underline{x})}{\partial x_j \partial x_k}\right|_{\underline{x}=\bar{\bar{\underline{x}}}}. \tag{3.6}$$

Applying the smoothing operator to both sides of Equation (3.4),

$$\overline{f_i(\underline{x})} \doteq f_i(\bar{\bar{\underline{x}}}) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 f_i(\bar{\bar{\underline{x}}})}{\partial x_j \partial x_k} \overline{(x_j - \bar{\bar{x}}_j)(x_k - \bar{\bar{x}}_k)}. \tag{3.7}$$

Noting the jk^{th} element of the smoothing covariance matrix S is given by

$$S_{jk} = \overline{(x_j - \bar{\bar{x}}_j)(x_k - \bar{\bar{x}}_k)}, \tag{3.8}$$

Equation (3.7) becomes

$$\overline{f_i(\underline{x})} \doteq f_i(\bar{\bar{\underline{x}}}) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} S_{jk} \frac{\partial^2 f_i(\bar{\bar{\underline{x}}})}{\partial x_j \partial x_k}. \tag{3.9}$$

An approximation for $\overline{[1/q(\underline{x},t)][\partial q(\underline{x},t)/\partial x_k]}$ is now sought. If $q(\underline{x},t)$ were actually known in the form of some analytic expression, such as Fisher's quasi-moment expansion[7], the desired approximation could easily be attained by techniques like those used to evaluate $\overline{f_i(\underline{x})}$. In practice, $q(\underline{x},t)$ is unlikely to be available in analytic form. For our purposes, it will be assumed that the filtered mean and covariance have been obtained (or approximated) by some technique such as that of Schwartz[6]. To obtain an approximation for $[1/q(\underline{x},t)][\partial q(\underline{x},t)/\partial x_k]$, it will be assumed that $q(\underline{x},t)$ is nearly gaussian, so that it is adequately approximated by a gaussian density function having the computed filtered mean and covariance. Denoting the filtered mean and covariance by $\bar{\bar{\underline{x}}}$ and P, respectively,

$$q(\underline{x},t) \doteq \frac{1}{(2\pi)^{n/2}(|P|)^{1/2}} \exp\left\{-\tfrac{1}{2}[\underline{x} - \bar{\bar{\underline{x}}}]' P^{-1}[\underline{x} - \bar{\bar{\underline{x}}}]\right\}, \tag{3.10}$$

where the time dependence has been suppressed. By straight-forward evaluation (recalling that P^{-1} is symmetric),

$$\frac{\partial q(\underline{x},t)}{\partial x_k} \doteq q(\underline{x},t) \sum_{l=1}^{n} (P^{-1})_{kl}(\bar{x}_l - x_l) . \tag{3.11}$$

Consequently,

$$\frac{1}{q(\underline{x},t)} \frac{\partial q(\underline{x},t)}{\partial x_k} \doteq \sum_{l=1}^{n} [P^{-1}(t)]_{kl}(\bar{x}_l - x_l) . \tag{3.12}$$

Applying the smoothing operator to Equation (3.12),

$$\overline{\frac{1}{q(\underline{x},t)} \frac{\partial q(\underline{x},t)}{\partial x_k}} \doteq \sum_{l=1}^{n} [P^{-1}(t)]_{kl}(\bar{x}_l - \bar{x}_l) . \tag{3.13}$$

With the aid of Equations (3.3), (3.9), and (3.13), the differential equation for $\bar{x}_i$ becomes:

$$\frac{d\bar{x}_i(t)}{dt} = f_i(\bar{\underline{x}},t) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} s_{jk} \frac{\partial^2 f_i(\bar{\underline{x}})}{\partial x_j \partial x_k} + \sum_{k=1}^{n}\sum_{l=1}^{n} q_{ik}(t)[P^{-1}(t)]_{kl}(\bar{x}_l - \bar{x}_l) . \tag{3.14}$$

A differential equation for the covariance is now desired, since it appears in Equation (3.14). By the linearity of the smoothing operator,

$$s_{jk} = \overline{(x_j - \bar{x}_j)(x_k - \bar{x}_k)} = \overline{x_j x_k} - \bar{x}_j \bar{x}_k . \tag{3.15}$$

Since Equation (3.15) implies

$$\dot{s}_{jk} = \frac{d}{dt}\overline{x_j x_k} - \frac{d}{dt}(\bar{x}_j \bar{x}_k) , \tag{3.16}$$

it suffices to evaluate $d(\overline{x_j x_k})/dt$ and $d(\bar{x}_j \bar{x}_k)/dt$. By letting

$$h(x) = x_j x_k \tag{3.17}$$

in Equations (2.64) and (2.65),

$$\frac{d\overline{x_j x_k}}{dt} = \overline{f_j x_k} + \overline{f_k x_j} - q_{jk} - \overline{\sum_{l=1}^{n}(q_{jl}x_k + q_{kl}x_j)\sum_{m=1}^{n}(P^{-1})_{lm}(\bar{x}_m - x_m)} . \tag{3.18}$$

By the linearity of the smoothing operator,

$$\frac{d\overline{x_j x_k}}{dt} = \overline{f_j x_k} + \overline{f_k x_j} - q_{jk} + \sum_{l=1}^{n}\sum_{m=1}^{n} q_{jl}(P^{-1})_{lm}(\overline{x_k x_m} - \bar{x}_k \bar{x}_m) +$$

$$+ \sum_{l=1}^{n}\sum_{m=1}^{n} q_{kl}(P^{-1})_{lm}(\overline{x_j x_m} - \bar{x}_j \bar{x}_m) . \tag{3.19}$$

For mechanization of Equation (3.19), it is sufficient to obtain a mechanizable approximation to $\overline{f_j x_k}$ for all j and k, assuming that Equation (3.19) is mechanized for all j and k. It has already been assumed that the third central moment is negligible. Since

$$E_s[(x_j - \bar{x}_j)(x_k - \bar{x}_k)(x_l - \bar{x}_l)] = E_s[(x_j - \bar{x}_j)(x_k - \bar{x}_k)x_l] - E_s[(x_j - \bar{x}_j)(x_k - \bar{x}_k)\bar{x}_l] , \tag{3.20}$$

this assumption implies that

$$E_s[(x_j - \bar{x}_j)(x_k - \bar{x}_k)x_l] \doteq s_{jk}\bar{x}_l . \tag{3.21}$$

Multiplying Equation (3.4) by x_k,

$$f_j(\underline{x})x_k \doteq f_j(\bar{\underline{x}})x_k + \sum_{l=1}^{n} \frac{\partial f_j(\bar{\underline{x}})}{\partial x_l}(x_l - \bar{x}_l)x_k + \frac{1}{2}\sum_{l=1}^{n}\sum_{m=1}^{n} \frac{\partial^2 f_j(\bar{\underline{x}})}{\partial x_l \partial x_m}(x_l - \bar{x}_l)(x_m - \bar{x}_m)x_k . \tag{3.22}$$

and subtracting $\bar{f}_j\bar{x}_k$ from each side (using the expansion Equation (3.7) for $\bar{f}_j$ on the right side of Equation (3.22)):

$$f_j(\underline{x})x_k - \bar{f}_j\bar{x}_k \doteq f_j(\bar{\underline{x}})x_k + \sum_{l=1}^{n}\frac{\partial f_j(\bar{\underline{x}})}{\partial x_l}(x_l-\bar{x}_l)x_k + \frac{1}{2}\sum_{l=1}^{n}\sum_{m=1}^{n}\frac{\partial^2 f_j(\bar{\underline{x}})}{\partial x_l\partial x_m}(x_l-\bar{x}_l)(x_m-\bar{x}_m)x_k -$$

$$-\bar{x}_k f_j(\bar{\underline{x}}) - \frac{1}{2}\bar{x}_k\sum_{l=1}^{n}\sum_{m=1}^{n}\frac{\partial^2 f_j(\bar{\underline{x}})}{\partial x_l\partial x_m}(x_l-\bar{x}_l)(x_m-\bar{x}_m)\,. \tag{3.23}$$

Applying the smoothing operator to both sides of Equation (3.23) and using Equation (3.21),

$$\overline{f_j(\underline{x})x_k} \doteq \bar{f}_j\bar{x}_k + \sum_{l=1}^{n}\frac{\partial f_j(\bar{\underline{x}})}{\partial x_l}S_{lk}\,. \tag{3.24}$$

With the aid of Equations (3.7) and (3.24), Equation (3.19) becomes

$$\frac{d\overline{x_jx_k}}{dt} \doteq f_j(\bar{\underline{x}})\bar{x}_k + f_k(\bar{\underline{x}})\bar{x}_j + \frac{1}{2}\bar{x}_k\sum_{l=1}^{n}\sum_{m=1}^{n}S_{lm}\frac{\partial^2 f_j(\bar{\underline{x}})}{\partial x_l\partial x_m} +$$

$$+\frac{1}{2}\bar{x}_j\sum_{l=1}^{n}\sum_{m=1}^{n}S_{lm}\frac{\partial^2 f_k(\bar{\underline{x}})}{\partial x_l\partial x_m} + \sum_{l=1}^{n}S_{lk}\frac{\partial f_j(\bar{\underline{x}})}{\partial x_l} +$$

$$+\sum_{l=1}^{n}S_{lj}\frac{\partial f_k(\bar{\underline{x}})}{\partial x_l} - Q_{jk} + \sum_{l=1}^{n}\sum_{m=1}^{n}Q_{jl}(P^{-1})_{lm}(\overline{x_kx_m} - \bar{x}_k\bar{x}_m) +$$

$$+\sum_{l=1}^{n}\sum_{m=1}^{n}Q_{kl}(P^{-1})_{lm}(\overline{x_jx_m} - \bar{x}_j\bar{x}_m)\,. \tag{3.25}$$

It remains to evaluate $d(\bar{x}_j\bar{x}_k)/dt$. It is not really necessary to evaluate this term since, in evaluating S_{jk}, it is possible to generate $\overline{x_jx_k}$ by Equation (3.25) and to generate $\bar{x}_j\bar{x}_k$ as the product of solutions to Equation (3.14). However, it is interesting to study the basic differential equation satisfied by S_{jk}, so an expression for the remaining term of Equation (3.16) will be derived.

For this case, both the ordinary calculus and the stochastic calculus imply (Feller[3]):

$$\frac{d\bar{x}_j\bar{x}_k}{dt} = \bar{x}_j\frac{d\bar{x}_k}{dt} + \bar{x}_k\frac{d\bar{x}_j}{dt}\,. \tag{3.26}$$

Using Equation (3.14) in Equation (3.26),

$$\frac{d\bar{x}_j\bar{x}_k}{dt} \doteq f_k(\bar{\underline{x}})\bar{x}_j + f_j(\bar{\underline{x}})\bar{x}_k + \frac{1}{2}\bar{x}_j\sum_{l=1}^{n}\sum_{m=1}^{n}S_{lm}\frac{\partial^2 f_k(\bar{\underline{x}})}{\partial x_l\partial x_m} +$$

$$+\frac{1}{2}\bar{x}_k\sum_{l=1}^{n}\sum_{m=1}^{n}S_{lm}\frac{\partial^2 f_j(\bar{\underline{x}})}{\partial x_l\partial x_m} + \sum_{l=1}^{n}\sum_{m=1}^{n}Q_{kl}[P^{-1}]_{lm}(\bar{x}_m - x_m)\bar{x}_j +$$

$$+\sum_{l=1}^{n}\sum_{m=1}^{n}Q_{jl}[P^{-1}]_{lm}(\bar{x}_m - x_m)\bar{x}_k\,. \tag{3.27}$$

Using Equations (3.15) and (3.27) and recalling that $S_{jk} = S_{kj}$,

$$\frac{dS_{jk}(t)}{dt} \doteq \sum_{l=1}^{n}\left[S_{lk}\frac{\partial f_j(\bar{\underline{x}})}{\partial x_l} + S_{lj}\frac{\partial f_k(\bar{\underline{x}})}{\partial x_l}\right] + \sum_{l=1}^{n}\sum_{m=1}^{n}\left(Q_{jl}[P^{-1}]_{lm}S_{mk} + Q_{kl}[P^{-1}]_{lm}S_{mj}\right) - Q_{jk}\,. \tag{3.28}$$

Equations (3.14) and (3.28) are the principal results of this section; they are the desired equations for approximate nonlinear minimum variance smoothing.

3.2 Linear Minimum Variance Smoothing

It is interesting to determine the results of applying Equations (3.14) and (3.28) to the case of a linear system. For a linear system, Equation (2.1) becomes

$$\dot{\underline{x}}(t) = F(t)\underline{x}(t) + \underline{\xi}(t) , \tag{3.29}$$

where $F(t)$ is an $n \times n$ matrix. Thus

$$f_i(\bar{\underline{x}},t) = \sum_{j=1}^{n} F_{ij}\bar{x}_j \tag{3.30}$$

and

$$\frac{\partial^2 f_i(\bar{\underline{x}})}{\partial x_j \partial x_k} = 0 . \tag{3.31}$$

Substituting Equations (3.29) and (3.30) into Equation (3.14) yields

$$\frac{d\bar{x}_i}{dt} = \sum_{j=1}^{n} F_{ij}\bar{x}_j + \sum_{k=1}^{n}\sum_{l=1}^{n} Q_{ik}(t)[P^{-1}(t)]_{kl}(\bar{x}_l - x_l) , \tag{3.32}$$

or, equivalently,

$$\frac{d\bar{\underline{x}}_i}{dt} = F\bar{\underline{x}}(t) + Q(t)P^{-1}(t)[\bar{\underline{x}}(t) - \underline{x}(t)] . \tag{3.33}$$

Proceeding similarly for S_{jk} ,

$$\sum_{l=1}^{n} \frac{\partial f_i(\bar{\underline{x}})}{\partial x_l} S_{lk} = \sum_{l=1}^{n} f_{jl}(t)S_{lk} , \tag{3.34}$$

which, when substituted into Equation (3.28), becomes

$$\frac{dS_{jk}(t)}{dt} = \sum_{l=1}^{n}(F_{jl}S_{lk} + F_{kl}S_{lj}) - Q_{jk} + \sum_{l=1}^{n}\sum_{m=1}^{n} Q_{jl}[P^{-1}]_{lm}S_{mk} + \sum_{l=1}^{n}\sum_{m=1}^{n} Q_{kl}[P^{-1}]_{lm}S_{mj} \tag{3.35}$$

or, equivalently,

$$\frac{dS(t)}{dt} = [F(t) + Q(t)P^{-1}(t)]S(t) + S(t)[F(t) + Q(t)P^{-1}(t)] - Q(t) . \tag{3.36}$$

Equations (3.33) and (3.36) are identical to linear smoothing results obtained by Rauch, Tung, and Striebel[6] and repeated elsewhere in this publication. Because the smoothing density function is gaussian, the minimum variance and maximum likelihood estimates are identical and are provided by the solution to Equation (3.33). Boundary conditions for Equations (3.33) and (3.36) are

$$\bar{\underline{x}}(T) = \underline{x}(T) \tag{3.37}$$

and

$$S(T) = P(T) , \tag{3.38}$$

where T is the final time in the smoothing interval.

The fact that Equations (3.14) and (3.28) reduce to forms identical with previously derived results for linear systems is expected. Each approximation introduced in the derivations of the equations in this section is not an approximation at all when applied to linear systems. The filter density, $q(\underline{\mu},t)$, is gaussian in this case, and the third central moments are zero. The expansions about $\bar{\underline{x}}$ are also exact in this case.

3.3 An Iterative Approach

The approximate nonlinear minimum variance smoother has the advantage of relative simplicity, but is based on certain questionable assumptions. Assuming the third central moment to be negligible is equivalent to assuming there is no skewness to the density of $\underline{x}(t)$ conditioned on all the data. However, all sorts of simple nonlinearities produce densities with significant skewness when driven by forcing functions containing gaussian noise. It is, of course, quite possible to use higher order approximations. Whether the improved performance would justify the increased complexity is questionable. Including a moment of order m as the highest moment in the approximation always requires the assumption that the moment of order $m + 1$ is negligible. Because even

order moments (at least for near gaussian cases) are not at all negligible, it is clear that $m + 1$ should not be an even number. Consequently, to include the effects of skewness, both third and fourth order moments should be included. It appears that such a procedure may soon destroy the simplicity of the approach.

If one is truly concerned about the effect of assuming the third central moment to be negligible, the following iterative approach is suggested. In Equations (3.14) and (3.28), the term $\partial f_j(\bar{\underline{x}})/\partial x_l$ was used as a shorthand for $\partial f_j(x)/\partial x_l$ evaluated at $\underline{x} = \bar{\underline{x}}$, and similarly for $\partial^2 f_j(x)/(\partial x_k \partial x_l)$. If these terms are evaluated for $\underline{x} = \hat{\underline{x}}^c$, where $\hat{\underline{x}}^c$ is the c^{th} estimate of $\underline{x}$, then the following iterative equations are suggested to use in place of Equations (3.14) and (3.28):

$$\hat{x}_l^{c+1} = f_l(\hat{\underline{x}}^c, t) + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} s_{jk}\frac{\partial^2 f_j(\hat{\underline{x}}^c)}{\partial x_j \partial x_k} + \sum_{k=1}^{n}\sum_{l=1}^{n} q_{lk}[P^{-1}]_{kl}(\hat{x}_l^c - \bar{x}_l) , \tag{3.39}$$

$$\dot{s}_{jk}^{c+1} = \sum_{l=1}^{n}\left[s_{lk}^c\frac{\partial f_j(\hat{\underline{x}}^c)}{\partial x_l} + s_{lj}^c\frac{\partial f_k(\hat{\underline{x}}^c)}{\partial x_l}\right] + \sum_{l=1}^{n}\sum_{m=1}^{n}\left(q_{jl}[P^{-1}]_{lm}s_{mk}^c + q_{kl}[P^{-1}]_{lm}s_{mj}^c\right) . \tag{3.40}$$

This type of iteration is similar to quasi-linearization, and whereas no specific studies have been performed, it is anticipated that $\hat{\underline{x}}^c \to \bar{\underline{x}}$ and $S^c \to S$ very rapidly, probably to high accuracy with only a couple of iterations. It should be noted that Equations (3.14) and (3.28) should provide better estimates for nonlinear systems than linear theory, and the use of the iterative technique should permit quick convergence to very accurate estimates - even more accurate than Equations (3.14) and (3.28) used in non-iterative fashion.

4. MAXIMUM LIKELIHOOD SMOOTHING

For nonlinear systems, the smoothing density function will not, in general, be gaussian, thus the mean value (minimum variance estimate) will not, in general, be equal to the value for which the smoothing density function has its maximum value. In some cases, one is more interested in the maximum likelihood estimate. There are two approaches to this problem. One approach is to evaluate the maximum likelihood estimate directly, and a technique for this is presented in Section 4.1. The second approach is first to find the smoothing density function and then to find that value of the state for which the smoothing density function has its maximum value (for each t in the smoothing interval). A possible technique of doing this and the resultant difficulties will be briefly discussed in Section 4.2.

4.1 Direct Approach

This direct approach is due to Bryson and Frazier[9]. Let the system be given by the nonlinear vector equation,

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x}, \underline{g}, t) , \tag{4.1}$$

where $\underline{x}(t)$ is the n-dimensional state vector and $\underline{g}(t)$ is a control vector independent of the state. Assume the *a priori* estimates are available:

$$E[\underline{x}(t_0)] = \bar{x}_0 \tag{4.2}$$

$$E\{[\underline{x}(t_0) - \bar{x}_0][\underline{x}(t_0) - \bar{x}_0]'\} = P_0 \tag{4.3}$$

$$E[\underline{g}(t)] = \bar{\underline{g}}(t) \tag{4.4}$$

$$E\{[\underline{g}(t) - \bar{\underline{g}}(t)][g(\tau) - \bar{\underline{g}}(\tau)]'\} = \bar{Q}(t)\delta(t-\tau) . \tag{4.5}$$

Let the relationship between the system state $\underline{x}(t)$ and the measurement process be given by

$$\underline{h}[\underline{z}(t), \underline{x}(t), t] = 0 \tag{4.6}$$

for the case of no measurement noise ($\underline{h}$ is a p-dimension vector). Let

$$E\{[\underline{z}(t) - \bar{\underline{z}}(t)][\underline{z}(\tau) - \bar{\underline{z}}(\tau)]'\} = \bar{R}(t)\delta(t-\tau) \tag{4.7}$$

be an *a priori* estimate of errors in $\underline{z}(t)$. The Dirac delta function appearing in these equations indicates "white noise" in the plant, control, and measurements.

Bryson and Frazier find the maximum likelihood estimate of $\underline{g}(t)$ and the initial condition, $\underline{x}(t_0)$. They also state that the maximum likelihood estimate $\hat{\underline{x}}(t)$ of the state $\underline{x}(t)$ is obtainable from the maximum likelihood estimate $\hat{\underline{g}}(t)$ of the control $\underline{g}(t)$ by the equation

$$\dot{\hat{\underline{x}}} = \underline{f}(\hat{\underline{x}}, \hat{\underline{g}}, t) . \tag{4.8}$$

For linear systems, this is clearly true. However, for nonlinear systems, it is not known to be true in general, and, in fact, the maximum of one probability density does not necessarily map into the maximum of another probability density function simply because of some functional relationship between the two variables. Therefore, the technique will be presented here with the understanding that the result of the computations will be the maximum likelihood estimate of the control, and that $\underline{x}(t)$ is simply that estimate of the state which is consistent with the maximum likelihood estimate of the control, i.e., that which satisfies Equation (4.8).

The maximum likelihood estimate of $\underline{x}(t_0)$ and $\underline{c}(t)$ is found by minimizing the functional,

$$J = \frac{1}{2}\tilde{\underline{x}}'(t_0)\bar{P}_0^{-1}\tilde{\underline{x}}(t_0) + \frac{1}{2}\int_{t_0}^{T} (\tilde{\underline{z}}'\bar{R}^{-1}\tilde{\underline{z}} + \tilde{\underline{c}}'\bar{Q}^{-1}\tilde{\underline{c}})dt , \tag{4.9}$$

where

$$(\tilde{\ }) = (\hat{\ }) - (\bar{\ }) \tag{4.10}$$

with the constraints

$$h(\hat{\underline{z}}, \hat{\underline{x}}, t) = 0 \tag{4.11}$$

and

$$\dot{\hat{\underline{x}}} = f(\hat{\underline{x}}, \underline{c}, t) . \tag{4.12}$$

By adjoining the constraints of Equations (4.11) and (4.12) and applying the calculus of variations, Bryson and Frazier show that the smoothed values of the control and initial condition satisfy

$$\hat{\underline{c}}(t) = \bar{\underline{c}}(t) + \bar{Q}(t)\left(\frac{\partial f}{\partial \underline{u}}\right)'\underline{\lambda}(t) , \tag{4.13}$$

$$\dot{\underline{\lambda}}(t) = -\left(\frac{\partial f}{\partial \underline{x}}\right)'\underline{\lambda}(t) - \left(\frac{\partial h}{\partial \underline{x}}\right)'\left[\left(\frac{\partial h}{\partial \underline{z}}\right)'\right]^{-1}\bar{R}^{-1}(\hat{\underline{z}} - \bar{\underline{z}}) , \tag{4.14}$$

where

$$\hat{\underline{c}}(t) = \bar{\underline{c}}(t) + \bar{Q}\left(\frac{\partial f}{\partial \underline{u}}\right)'\underline{\lambda}(t) , \tag{4.15}$$

$$h(\hat{\underline{z}}, \hat{\underline{x}}, t) = 0 , \tag{4.16}$$

and the boundary conditions,

$$\hat{\underline{x}}(t_0) = \bar{\underline{x}}(t_0) + \bar{P}_0\underline{\lambda}(t_0) \tag{4.17}$$

$$\underline{\lambda}(T) = 0 . \tag{4.18}$$

By this approach, the smoothing problem (for maximum likelihood estimates) becomes a two-point boundary value problem. The numerical solution of two-point boundary value problems on digital computers is often very difficult, although good results are often attained by quasi-linearization techniques. Assuming such techniques succeed in solving the numerical version of this two-point boundary value problem, then this approach has the advantage of not requiring a solution to the filtering problem before the smoothing problem can be worked.

4.3 Quasi-Moment Approach

As mentioned previously, another approach to the maximum likelihood estimation problem is to determine the smoothing density function and then to find that value which maximizes the value of that function; this must be done for each value of t in the smoothing interval. In this section, the technique of expressing a density function in terms of quasi-moments is briefly described, and the reader is then referred to Peller[3] for further information on this approach. The material contained in this section is almost entirely derived from Fisher's work[7], particularly as presented in Reference 10. In fact, this section should simply be considered as a selective abridgment of Fisher's work.

Let $p(\underline{\mu},t)$ be an n-dimensional density function, and let $p_g(\underline{\mu},t)$ be an n-dimensional density function having the same mean and covariance as $p(\underline{\mu},t)$. Let

$$\rho(\underline{\mu},t) = \frac{p(\underline{\mu},t)}{p_g(\underline{\mu},t)} . \tag{4.19}$$

Then, the coefficients of the expansion of $\rho(\underline{\mu},t)$ in a series of multi-dimensional Hermite polynomials are the so-called quasi-moment functions. Since the multi-dimensional Hermite polynomials are a complete set of eigen-functions over n-dimensional Euclidean space, any probability density function can be written as a series of these polynomials, provided $\rho(\underline{\mu},t)$ is square-integrable, i.e., if

$$\int_{-\infty}^{\infty} \rho^2(\underline{\mu},t)\ d\mu < \infty\ . \tag{4.20}$$

It is also possible to approximate $\rho(\underline{\mu},t)$, if it satisfies Equation (4.20), to any specified accuracy in the integrable square error sense by a finite number of terms in the series.

Denoting the mean and covariance by $\underline{m}(t)$ and $C(t)$, respectively, and the characteristic functions of $p(\underline{\mu},t)$ and $p_g(\underline{\mu},t)$ by $\phi(\underline{\alpha},t)$ and $\phi_g(\underline{\alpha},t)$, respectively, Fisher shows that

$$p(\underline{\mu},t) = p_g(\underline{\mu},t) + \sum_{N=3}^{\infty} \frac{(-1)^N}{N!} \underbrace{\sum_{j,k,\ldots,l=1}^{n}}_{N} \underbrace{K_{j,k,\ldots,l}}_{N}(t)\ \frac{\partial^N p_g(\underline{\mu},t)}{\partial\mu_j, \partial\mu_k, \ldots, \partial\mu_l}\ , \tag{4.21}$$

where

$$\underbrace{K_{j,k,\ldots,l}}_{N}(t) = \frac{1}{i^N}\ \frac{\partial^N[\phi(\underline{\alpha},t)/\phi_g(\underline{\alpha},t)]}{\partial\alpha_j, \partial\alpha_k, \ldots, \partial\alpha_l}\bigg|_{\underline{\alpha}=0} \tag{4.22}$$

are the quasi-moment functions of order N . It can easily be shown that all quasi-moments of first and second order are zero.

Because the multi-dimensional Hermite polynomials corresponding to the $n \times n$ matrix $C^{-1}(t)$ are related to a generating function as follows,

$$\underbrace{H_{j,k,\ldots,l}}_{N}(\underline{\mu},t) = (-1)^N \exp\left[\tfrac{1}{2}\underline{\mu}'C(t)\underline{\mu}\right]\ \frac{\partial^N \exp\left[-\tfrac{1}{2}\underline{\mu}'C^{-1}(t)\underline{\mu}\right]}{\partial\mu_j, \partial\mu_k, \ldots, \partial\mu_l}\ , \tag{4.23}$$

$$N\begin{cases} j = 1,2,\ldots,n \\ \vdots \\ l = 1,2,\ldots,n \end{cases}$$

it is possible to rewrite Equation (4.22) in terms of the multi-dimensional Hermite polynomials:

$$p(\underline{\mu},t) = p_g(\underline{\mu},t)\left\{1 + \sum_{N=3}^{\infty} \frac{1}{N!} \underbrace{\sum_{j,k,\ldots,l=1}^{n}}_{N} \underbrace{K_{j,k,\ldots,l}}_{N}(t)\ \underbrace{H_{j,k,\ldots,l}}_{N}[\underline{\mu} - \underline{m}(t)\ ,\ t]\right\}\ . \tag{4.24}$$

Equation (4.24) shows that any probability density function can be expressed directly in terms of the mean, covariance, and quasi-moment functions. Equation (4.24) can be the basis of an approximation for the smoothing density function, $s(\underline{\mu},t)$. Peller[3] gives the basic development for this, although it is so tedious that parts of it are only outlined in terms of the required procedures. While there are certain advantages to the quasi-moment expansion of Equation (4.24), its use for the smoothing problem is not recommended, and it will receive no further discussion here.

5. CONCLUSIONS

This chapter has developed the exact (in the mean-square sense) differential equations for the smoothing density function and the smoothed expectation of an arbitrary function of the state. These results were good for nonlinear systems and observations corrupted by additive gaussian white noise and were extended to the product noise case. In Section 3, approximate differential equations were presented for minimum variance smoothing, although it was assumed that the plant noise was not a function of the state for this case. One technique of maximum likelihood smoothing was presented and another suggested.

The equations developed in this chapter, particularly the minimum variance equations of Section 3, should be useful whenever probabalistic smoothing of data is desired, such as in estimating missile guidance system accuracy from noisy tracking data. Whereas statistical smoothing of data sometimes requires many passes over the data or many iterations to fit some polynomial to the data in a least squares fit, the technique presented here requires just two passes: one pass forward to get the filtered solution, and a second pass backwards to get the smoothed solution. The forward pass should probably use nonlinear filter theory such as that presented by Schwartz[9] to obtain maximum value from the approach.

There are certain unresolved questions regarding the results presented in this chapter. One, all results have been derived formally: little regard to mathematical rigor has been given. The questions of existence and uniqueness of solutions to the various differential equations have not been treated. In any area dealing with

stochastic processes, this could be a serious omission since relatively little is known about the conditions necessary to guarantee existence and uniqueness of solutions. Another question is the sensitivity of the results to uncertainties in the knowledge of the plant and of the various covariance matrices involved. These types of uncertainties are known to be potentially very serious in filter theory; some of these problems are treated elsewhere in this publication.

ACKNOWLEDGMENT

Much of the work contained in this paper was supported by various grants by the United States Air Force and a North American Rockwell fellowship.

REFERENCES

1. Doob, J.L. — *Stochastic Processes.* Wiley, New York, 1953.

2. Striebel, C.T. — *Partial Differential Equations for the Conditional Distribution of a Markov Process Given Noisy Observations.* Journal of Mathematical Analysis and its Applications, Vol.11, No.1-3, July 1965, pp. 151-160.

3. Peller, John B. — *Nonlinear Smoothing Theory and Applications.* PhD Thesis, University of California, Los Angeles, 1966.

4. Bucy, R.S. — *Nonlinear Filtering Theory.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol.10, No.2, April 1965, p.198.

5. Bass, R.W. et al. — *Optimal Multi-Channel Nonlinear Filtering.* Hughes Aircraft Company, Culver City, California, Report SSD 50064R, August 1966.

6. Schwartz, L. — *Approximate Continuous Nonlinear Minimal-Variance Filtering.* Hughes Aircraft Company, Culver City, California, Report SSD 60472R, December 1966.

7. Fisher, James R. — *Conditional Probability Density Functions and Optimal Nonlinear Estimation.* PhD Thesis, University of California, Los Angeles, 1966.

8. Rauch, H.E. et al, — *On the Maximum Likelihood Estimates for Linear Dynamic Systems.* Lockheed Aircraft Corporation, Sunnyvale, California, Technical Report 6-90-63-62, June 1963.

9. Bryson, A.E. Frazier, M. — *Smoothing for Linear and Nonlinear Dynamic Systems.* Published in the Proceedings of the "Optimum System Synthesis Conference", Aeronautical Systems Division, USAF, September 11-13, 1962. (Also Wright-Patterson Air Force Base, Ohio, ASD Technical Documentary Report, ASD-TDR-63-119, February 1963, pp.354-364.)

10. Fisher, J.R. — *Optimal Nonlinear Filtering.* pp.197-300 in "Advances in Control Systems", Vol.5 (edited by C.T.Leondes). Academic Press, New York, 1967.

CHAPTER 10 - GENERAL QUESTIONS ON KALMAN FILTERING IN NAVIGATION SYSTEMS

by

Larry D. Brock* and George T. Schmidt†

* Assistant Design Project Engineer
Hamilton Standard
Farmington, Connecticut, USA

† Principal Engineer
Instrumentation Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts, USA

CHAPTER 10 - GENERAL QUESTIONS ON KALMAN FILTERING IN NAVIGATION SYSTEMS

Larry D. Brock and George T. Schmidt

1. INTRODUCTION

Navigation systems are among the most popular areas for the application of Kalman filtering. Most of the major navigation system manufacturers have developed or proposed systems with Kalman filtering, and it is being used in several vehicles that are in operational use. Two examples are the Apollo space vehicle and the C-5A transport aircraft. Kalman filtering has now become an expected part of almost every proposed new navigation system. The reasons for the popularity of Kalman filtering in navigation systems are not hard to find. There are at least three major complementary factors that have come together at the proper time. These factors are an increased need, the mathematical tools, and the necessary equipment.

The need for effective on-board data processing is great for both space and terrestrial vehicles. The advent of manned travel to the moon, with the requirement for self-contained navigation, required an efficient method for the real-time processing of all navigation measurements. In terrestrial vehicles (both aircraft and marine) the increased need is caused primarily by the introduction of inertial navigation systems that have very good high-frequency response and give a continuous indication of position, velocity, and attitude. They are also self-contained and operate without electromagnetic inputs or outputs. However, they do have errors that grow with time. Thus, for most applications, there is a desire to bound the position error growth with external aids. Accuracy requirements dictate that this external data be processed as efficiently as possible.

A mathematical technique was developed at just the right time to fulfill these needs. The Kalman filter theory, published[1,2] in the early 1960's, was recognized as an ideal solution to the navigation data processing problems for both space[3,4] and terrestrial vehicles[5-10]. These data processing problems can be fitted very nicely into the necessary Kalman filter assumptions, which means that the estimates of the desired navigation outputs are in fact very nearly optimum. This gives the design engineer confidence that the system is the best that can reasonably be expected, without the concern that some other technique might be found that would significantly improve the system. Also, the recursive form of the filter is convenient. A new optimum estimate can be made very shortly after each new measurement is obtained. Nor is it necessary to store a great amount of data or invert a large matrix, as might be necessary in more conventional, least-squares fitting techniques.

However, as excellent a solution as Kalman filtering is to the data processing problem, it would have been of only academic interest if it were not for the fantastic rate of development in electronic circuits, which makes possible very powerful digital flight computers. These three factors have come together at the proper time to create considerable interest in Kalman filtering in the navigation industry. This interest partially motivated this publication.

In this chapter, some of the initial steps necessary for the application of Kalman filtering will be discussed in general. The application of filtering to terrestrial navigation will then be illustrated by simple examples. Two methods are suggested for simplifying the problem in order that it can be more easily handled on a practical computer. Details are then given of the alignment and calibration of the inertial system in a spacecraft on top of a swaying launch vehicle. Practical implementation problems, as well as actual hardware difficulties, are discussed in detail. It is hoped that from these simple examples the reader will be in a better position to appreciate the applications that follow in the text.

2. INITIAL STEPS IN THE APPLICATION OF KALMAN FILTERING

The Kalman filter itself is defined in very precise mathematical terms, but its application to actual physical systems is rarely a precise science. Considerable engineering experience is needed to properly identify the system to which the filter is to be applied, to adequately model that system, and then to develop a practical program that mechanizes the filter in the on-board computer. The optimization of the filter must include many factors which are difficult or impossible to describe mathematically, such as the trade-off between performance and computer size. The statistical parameters are rarely based on the actual statistics of the physical system, because these statistics are either too complicated or are not well known. The parameters are more likely to be chosen by less formal methods which attempt to maximize the performance in spite of the imperfectly known real world. For example, the covariance matrix in the operational Kalman filter will contain an estimate of the r.m.s. position accuracy of the navigation system. Best performance is usually obtained if this number matches fairly closely the actual r.m.s. position accuracy obtained by the system, even if the filter parameter must be "adjusted" to obtain this match. The noises associated with parameters that are modeled might be arbitrarily increased to make up for parameters that were too small or not well enough known to be included in the model.

One of the first steps in applying Kalman filtering is to identify the system on which the filter is to be based. The choice of system is not as obvious as it might seem. The most direct approach is to estimate the desired parameters of the vehicle directly. A functional diagram of this approach is shown in Figure 1. The primary system on which the filter is based is the system of equations that describe the motions of the vehicle itself. This direct method is used in space navigation systems. This approach could also be used in terrestrial navigation systems where the Kalman filter uses all measurements, including inertial measurements, to produce estimates of the position and velocity of the vehicle directly. However, for practical reasons which will be discussed later, an indirect method is used where the filter estimates the error in an inertial navigation system. The corrected inertial output is then used to indicate the position and velocity of the vehicle. In this case, the system on which the filter is based is the one which describes the errors in the inertial system. A functional diagram is shown in Figure 2.

The system on which the filter is to be based is specified by a set of state variables that are defined formally by a set of mathematical relations. In practice, for actual physical systems, there is never enough information to satisfy these mathematical relations perfectly. Further, it is never practical to include more than a few of the major variables in the on-board computer. One of the basic design problems is the choice of state variables for the navigation filter and the associated trade-off of performance versus computer requirements.

As a first step in the choice of state variables, theoretical error analyses and test data can be used to construct a set of state variables which describe the total system much more accurately than the description expected from the on-board filter. For a terrestrial navigation system, this more complete description could require as many as fifty to sixty state variables with many more which could be identified but are not worth considering. The performance of filters based on various limited sets of state variables can then be evaluated by using computer simulations based on the more complete model. One is chosen which is shown by the simulations to best satisfy the system requirements. For terrestrial systems, practical filters have consisted of between seven and twenty state variables. An effective operational test program is the only way to definitely confirm the design of the filter.

The mathematical relations which give the formal definition of the state variables are:

(a) The desired output must be some function of the state variables.

(b) The measurements must be a function of the state variables and uncorrelated errors.

(c) The state variables at one time must be a function of the state variables at a previous time, the controls to the system between the two times, and the uncorrelated (white) noise inputs between the two times.

The first relation can usually be satisfied very easily. The desired output of a navigation system is the position, velocity, and possibly attitude of the vehicle with respect to some navigation coordinate frame. The variables which represent position and velocity can usually be included directly in the state space. The second relation involving the measurements can be more difficult to formulate. The measurements will generally consist of two parts: a desired part which is some function of the states of the vehicle, such as position, velocity, or acceleration; and an undesired error. An error which can be considered uncorrelated between measurements does not require any additional state variables. However, if there are significant components of the error which are correlated, the model for the error must be added to the state variables.

The most effective overall design of the system can usually be obtained by minimizing the amount of preprocessing of the input data before it is used in the filter. One advantage of Kalman filtering is that measurements can be processed in raw form; this can greatly simplify the sensor sub-systems. For example, if an inertial system is to be aided by LORAN, it is not necessary for the LORAN equipment to solve for a position fix, which has then to be inserted in the filter to update the inertial system. The filter can be constructed to accept time difference measurements directly, thus simplifying the LORAN equipment. Another advantage of processing the raw measurements is that the model for measurement error can be much simpler. If the measurements are processed before being used by the filter, the modifications and cross-correlations of the error statistics caused by the preprocessing must be accounted for by the filter. Again using LORAN as an example, the errors in time differences can probably be adequately modeled as simple errors with no cross-correlations. The errors in latitude-longitude coordinate would be complicated functions of LORAN station-vehicle geometry and they would be cross-correlated.

The final step in the forming of the state vector is the addition of any variables necessary to describe the dynamic behavior of the state variables. The total set of state variables is the minimum number of dependent variables in the differential equations that describe the system. The inputs to the system can only be white noises and controls which are deterministic quantities that are known to the filter. Many of the basic sources of error in navigation systems are very highly autocorrelated as, for example, gyro drift rate. To satisfy the requirement for white noise inputs, a shaping filter must be devised which produces the desired correlated error. The input to the shaping filter is white noise, and the dependent variables in the shaping filter are added to the state space. If there are controls into the system which are not known quantities available to the filter, they too would have to be modeled as random noise.

Several general comments will now be made about the formulation of the state variables for space and terrestrial navigation systems. The Kalman filter for a space navigation system is, in many ways, more straightforward than for a terrestrial system. The desired outputs are the position and velocity vectors of the vehicle; this requires six state variables. The measurements usually consist of optical angular measurements between celestial

bodies and are functions of vehicle position. The optical measurements are usually relatively far apart, so that most of the error is uncorrelated with the previous error. In many cases the correlated part of the error will be insignificant, and no additional state variable will be required. In other situations, one additional variable might be required. The inertial measurements made in a spacecraft are obtained only during thrusting maneuvers. During this time there would not normally be any other source of data on the change in velocity. Thus, the inertial measurement is taken as a deterministic control quantity and is not filtered. The dynamic equations for a space vehicle are also clear cut. If the position and velocity of a space vehicle are known at one time, the position and velocity for any future time can be predicted with great accuracy. The uncertainties due to solar and atmospheric pressures, gravitational field uncertainties, and computing inaccuracies are very small unless the vehicle is in a low orbit. Thus, the dynamic behavior can be described with the six basic variables with possibly a small amount of white noise input to account for all the disturbances. Therefore, for an on-board space navigator with negligible correlated measurement errors, the entire filter can be based very nicely on six state variables.

The filter for a terrestrial system is not nearly as clear cut. The desired outputs are again position and velocity. The measurements can come from inertial sensors, Doppler radar, and a great variety of other aids. The errors in any of these sensors are very highly correlated in time. The inertial instruments will add at least six variables – three attitude errors and three gyro drift rates – and possibly more for errors such as scale factor errors and acceleration-sensitive gyro drifts. The Doppler would require at least two variables for scale factor and boresight errors. For radio aids, the sample period may be long enough for correlated errors to be neglected. The dynamic equations are the source of considerable difficulty if the Kalman filter is to be used in the direct approach as shown in Figure 1. This direct formulation requires that a statistical dynamic model for the vehicle be included in the state space. While the future state of a spacecraft can be predicted accurately for weeks, the future position of an aircraft can be predicted accurately for only a few seconds. The model used to describe these random motions is difficult to obtain, contains very high frequency components, and can be highly nonlinear. A direct filter for a terrestrial system is, therefore, difficult to formulate and puts great demand on the computer. Thus, although this direct filter approach is more intellectually satisfying, at the present time it is more of academic interest than of practical importance.

Even though this direct approach may not be practical until more powerful computers are available, the formulation of the navigation problem using this method can be very instructive and aid in the basic understanding of both the filtering process and the navigation process. This filter will give the optimum performance of an inertial navigation system even if no external data are available. This means the Kalman filter automatically solves some of the classical problems in inertial navigation. Optimum self-contained "damping" of the 84-minute and 24-hour error oscillations will be obtained, and the system will give the optimum indication of North by continuously "gyrocompassing". These statements are not as important as they might seem. It turns out that, for the levels of random motion that must be assumed for most vehicles, the optimum solution is only very slightly better than the deterministic solution of the navigation problem that has been used all along. The filter can only improve performance by separating the accelerations due to system errors from the actual accelerations of the vehicle. The system errors do have distinctive dynamic characteristics which make detection possible; however, their magnitudes are so much smaller than the accelerations of the vehicle that they are very difficult to separate. If any improvement could be expected at all, it would be for a system in a very stable vehicle, such as a submarine.

This direct filter will also give optimum performance when other sources of navigation information are added to the inertial data, and even when there are only external aids with no inertial data. For example, this filter could be used to combine Doppler, LORAN, and magnetic heading information without using an inertial system. It therefore has the advantage of continuing to operate after a failure in the inertial system.

If the navigation system does include an inertial system and some other source of navigation data, there is a completely different way of formulating the navigation problem which avoids most of the practical problems of the previous method. Instead of estimating the state of the vehicle directly, the filter is used to estimate the error state of an inertial navigation system. The inertial system follows the high-frequency motions of the vehicle very accurately, but has low-frequency errors which grow with time. The dynamic system on which the filter is based is the set of error equations for the inertial system, which are relatively well known, well behaved, low frequency, and essentially linear. The sample period can range from several seconds up to a minute without greatly influencing the effectiveness of the filter. For these reasons, this method is used for virtually every practical terrestrial Kalman filter mechanization.

The formulation of a Kalman filter for terrestrial navigation can best be illustrated by simple examples. Therefore, both methods are applied to a simplified navigation problem: the specific filter equations are defined, and simulation results are given. First, at the risk of being repetitious with previous chapters, the basic Kalman filter equations are stated. The equations are stated in the form most applicable to the navigation problem and are a mixture of the discrete form and the continuous form. The discrete form is used when a new measurement is introduced, and the continuous form is used between measurements. The equations are also a mixture of linear and nonlinear equations, the full nonlinear equations being used for the simulation of the system and of the measurement process. Linearization of these equations about the estimated state is used for the covariance matrix equations.

3. SUMMARY OF BASIC FILTER EQUATIONS

It is assumed that the state of the entire system, including sensors, vehicle, and environment, can be described by differential equations of the form

$$\dot{\underline{x}}(t) = \underline{f}[\underline{x}(t), \underline{c}(t), \underline{n}(t), t] , \tag{3.1}$$

where $\underline{c}$ are known control inputs and $\underline{n}$ are white noises. For the assumptions to be made here, it can be shown that the control does not affect the form of the optimum filter. Thus, the control variables will not be shown explicitly in the following discussion. It is assumed that measurements are made at discrete times according to the relation

$$\underline{m}(t_n) = \underline{h}[\underline{x}(t_n), \underline{u}(t_n)] , \tag{3.2}$$

where $\underline{u}(t_n)$ are errors in the measurements that are uncorrelated between measurements. (Time-correlated errors in the measurements would have to be included in the state vector $\underline{x}$.) Assuming that the optimal estimates are close enough to the true values for higher-order terms to be neglected, the optimum measurement process is given by Kalman's optimum linear filter. The basic equations are:

At a measurement time

$$\left.\begin{aligned} \hat{\underline{x}} &= \hat{\underline{x}}' + \mathbf{E}'\mathbf{H}^T(\mathbf{H}\mathbf{E}'\mathbf{H}^T + \mathbf{U})^{-1}(\underline{m} - \underline{h}(\hat{\underline{x}}', t)) \\ \mathbf{E} &= \mathbf{E}' - \mathbf{E}'\mathbf{H}^T(\mathbf{H}\mathbf{E}'\mathbf{H}^T + \mathbf{U})^{-1}\mathbf{H}\mathbf{E}' \end{aligned}\right\}$$

Between measurements

$$\left.\begin{aligned} \dot{\hat{\underline{x}}} &= \underline{f}(\hat{\underline{x}}, t) \\ \dot{\mathbf{E}} &= \mathbf{F}\mathbf{E} + \mathbf{E}\mathbf{F}^T + \mathbf{N} , \end{aligned}\right\} \tag{3.3}$$

where the prime (′) indicates conditions that exist just before the measurement. The covariance matrix of estimation errors $\mathbf{E}$ is defined by

$$\mathbf{E} = \langle \delta\underline{x}\, \delta\underline{x}^T \rangle ,$$

where "$\langle\,\rangle$" represents the expected or mean value, $\delta\underline{x}$ is the error in the estimate, and

$$\mathbf{F} = \left.\frac{\partial \underline{f}}{\partial \underline{x}}\right|_{\underline{x}=\hat{\underline{x}}} \qquad \mathbf{H} = \left.\frac{\partial \underline{h}}{\partial \underline{x}}\right|_{\underline{x}=\hat{\underline{x}}'}$$

$$\mathbf{U} = \left.\frac{\partial \underline{h}}{\partial \underline{u}}\,\mathbf{R}\,\frac{\partial \underline{h}^T}{\partial \underline{u}}\right|_{\underline{x}=\hat{\underline{x}}'} \qquad \mathbf{N} = \left.\frac{\partial \underline{f}}{\partial \underline{n}}\,\mathbf{Q}\,\frac{\partial \underline{f}^T}{\partial \underline{n}}\right|_{\underline{x}=\hat{\underline{x}}} .$$

The matrices $\mathbf{R}$ and $\mathbf{Q}$ are defined by

$$\langle \underline{u}(t_n)\underline{u}(t_m)^T \rangle = \mathbf{R}$$

$$\langle \underline{n}(t)\underline{n}(\tau)^T \rangle = \mathbf{Q}\delta(t-\tau) .$$

The optimum filter for a linearized system in which the higher-order terms can be neglected is shown in diagram form in Figure 3, where, at the sample times, the estimate of the state $\hat{\underline{x}}$ is changed by an impulse through the weighting factor, $\mathbf{K} = \mathbf{E}'\mathbf{H}^T(\mathbf{H}\mathbf{E}'\mathbf{H}^T + \mathbf{U})^{-1}$.

4. ILLUSTRATIVE EXAMPLES

Both the direct and indirect forms of the Kalman filter are illustrated by applying them to similar simplified navigation problems. For this example the vehicle is constrained to move only along a meridian on the surface of a spherical non-rotating earth. The vehicle is disturbed by random accelerations with gaussian distribution and an exponential autocorrelation function. Inertial measurements are obtained from an accelerometer that is mounted on a platform that rotates about an axis perpendicular to the meridian plane. The orientation of the platform is controlled by a single-degree-of-freedom gyro. The platform can be commanded to rotate at a given rate by a signal sent to the gyro, and it will also rotate due to gyro drift. The gyro drift rate is assumed to be a random-walk function beginning at some initial bias value. The accelerometer measurements are used by the direct filter to produce statistical estimates of the position and velocity of the vehicle. For the indirect formulation, the accelerometer output is connected directly to the gyro through the proper gain to give a deterministic Schuler-tuned inertial navigator. The measurement in this case is the difference between inertial velocity and velocity from a Doppler radar. The Doppler radar is assumed to have a scale factor error that is also described by a random-walk from an initial bias value. A schematic drawing of the navigation problem with the direct filter is shown in Figure 4. The diagram for the indirect formulation is basically the same except that Doppler data is also an input into the computer.

The steps for applying a statistical filter are: (1) choose the state variables that describe the complete system; (2) determine the dynamic equations that describe these state variables; (3) derive the equations that relate the measurements to the state variables; (4) develop the matrices of partial derivatives that are used in the covariance matrix equations ; and (5) if it can be assumed that the linear approximations of the errors are sufficient and that all signals and initial conditions have gaussian distributions, apply Equations (3.3). These steps are first applied to the direct form of the filter.

The state variables that describe this simplified navigation problem are vehicle acceleration a , velocity v , latitude L , platform misalignment from vertical α , and gyro drift rate d . A diagram of the system is shown in Figure 5. The numbers at the output of the gains K mean discrete changes are made at the corresponding points in the system at the sample times. The equations describing the dynamics of these variables are

$$\left.\begin{aligned} \dot{a} &= -\sigma a + n_a \\ \dot{v} &= a \\ \dot{L} &= v/R \\ \dot{\alpha} &= v/R + d + c \\ \dot{d} &= n_d \end{aligned}\right\} \tag{4.1}$$

where σ is the inverse correlation time of the vehicle acceleration, R is the radius of the earth, and n_a and n_d are the white noise inputs into the models of vehicle acceleration and gyro drift rate, respectively. The $\dot{\alpha}$ equation derives from the fact that the platform is rotated by $\hat{v}/R$ in order to keep it aligned with the vertical while the true vertical rotates by v/R . The platform is also torqued by $\hat{d}$ to compensate for the gyro drift rate d . The known control c is therefore $-\hat{v}/R - \hat{d}$. The equations for the estimated state variables are the same as Equations (4.1) with the actual state variables replaced by the estimated variables and the noise terms deleted. Note that because of c , $\hat{\dot{\alpha}}$ and $\hat{\dot{\alpha}}$ are zero.

The measurement of the accelerometer is given by

$$m = a\cos\alpha + [g - (v^2/R)]\sin\alpha + n_m , \tag{4.2}$$

where g is the acceleration of gravity and n_m is uncorrelated noise in the measurement.

The F and H matrices are obtained by taking the partial derivatives of Equations (4.1) and (4.2) with respect to the state variables. The F matrix is given by

$$F = \begin{bmatrix} -\sigma & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & R^{-1} & 0 & 0 & 0 \\ 0 & R^{-1} & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{4.3}$$

and the H matrix by

$$H = [1 \quad 0 \quad 0 \quad [g - (\hat{v}^2/R)] \quad 0] . \tag{4.4}$$

In the equation for H , the angle α is small such that $\cos\alpha$ can be assumed 1 and terms involving $\sin\alpha$ can be neglected as second order. The noise matrices are given by

$$N = \begin{bmatrix} A & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & D \end{bmatrix} ,$$

where A and D are defined by

$$\langle n_a(t)\, n_a(\tau)\rangle = A\,\delta(t-\tau)$$

$$\langle n_d(t)\, n_d(\tau)\rangle = D\,\delta(t-\tau)$$

and

$$U = \langle n_m^2 \rangle .$$

All terms are now defined so the optimum filter for this problem is completely defined by Equations (3.3).

An example of the response of this meridian system to typical inputs is given in Figures 6-9. The three sources of disturbances in the system are the random-walk gyro drift rate, the disturbance acceleration, and

the random measurement error. The random-walk gyro drift increases the r.m.s. drift by 0.05 meru/hr$^{1/2}$. (A meru is approximately 0.015 deg/hr.) The disturbance acceleration has an r.m.s. value of 0.003 ft/sec^2 and a correlation time of 1 min. The random measurement error has an r.m.s. value of 0.001 ft/sec^2. The sampling period for the measurement is 1 sec. Also shown for the same inputs are the responses of a Schuler-tuned undamped system and a system with lead-lag damping. The accelerations of the vehicle in steady-state motion between maneuvers used in this example are optimistically small even for a very stable vehicle such as a submarine. As the noise increases, the performance of the optimum system approaches very rapidly the performance of a deterministic Schuler-tuned system. However, this example does show that improvement is at least theoretically possible.

A Kalman filter is again applied to essentially the same navigation problem except that the indirect approach is used. The inertial system is mechanized in a closed Schuler loop and the filter uses Doppler data to estimate inertial errors. The state variables that describe the error equations to which the filter is to be applied are: Doppler scale factor error ΔSF, inertial velocity error Δv_i, inertial latitude error ΔL, platform misalignment α, and gyro drift d. The errors are defined as the indicated values minus the true values. A diagram of the system is shown in Figure 10. Written in matrix form, the differential equations for the state variables are

$$\begin{bmatrix} \Delta\dot{SF} \\ \Delta\dot{v}_i \\ \Delta\dot{L} \\ \dot{\alpha} \\ \dot{d} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g & 0 \\ 0 & R^{-1} & 0 & 0 & 0 \\ 0 & -R^{-1} & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta SF \\ \Delta v_i \\ \Delta L \\ \alpha \\ d \end{bmatrix} + \begin{bmatrix} n_s \\ 0 \\ 0 \\ 0 \\ n_d \end{bmatrix}. \tag{4.5}$$

The measurement equation is

$$\Delta m = \begin{bmatrix} -v_i & 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta SF \\ \Delta v_i \\ \Delta L \\ \alpha \\ d \end{bmatrix} + \epsilon_v . \tag{4.6}$$

The F and H matrices are given in Equations (4.5) and (4.6), respectively. The matrices U and N are defined by

$$U = \langle \epsilon_v^2 \rangle$$

$$N = \begin{bmatrix} N_s & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & N_d \end{bmatrix},$$

where

$$\langle n_d(t)\, n_d(\tau) \rangle = N_d\, \delta(t-\tau)$$

$$\langle n_s(t)\, n_s(\tau) \rangle = N_s\, \delta(t-\tau) .$$

The initial estimates for the state variables are assumed to be zero. It is also assumed that there is no initial cross-correlation between state variables. The initial covariance matrix (E) is then a diagonal matrix made up of the expected mean-squared value of each of the variables themselves, since the initial estimates are zero. With numbers specified for all the constants involved, all the information is now available to define the filter using Equations (3.3).

The results for the simulated operation of this system are given for the following typical values:

initial position error	= 0
initial velocity error	= 0
initial tilt	= 0.2 min
initial gyro drift	= 1.0 meru
initial scale factor error	= 0.3%
random-walk gyro drift (N_d)	= (0.1 meru)^2/hr
random scale factor (N_s)	= (0.05%)^2/hr
uncorrelated velocity error ($\sqrt{U}$)	= 0.5 knot.

It is assumed that the vehicle accelerates to 500 knots, holds this velocity for one hour, accelerates to 1200 knots, and then stops after 4 hours. Figures 11-13 show the position, velocity, and tilt errors in the inertial system with and without the filter. Figures 14-15 show the estimates of gyro drift and scale factor error that are made by the filter. Figures 16-20 show the expected uncertainty in each of the state variables. These uncertainties are given by the square root of the diagonal elements of the covariance matrix. It can be seen that there is a drop in most of the uncertainties at one hour when the velocity changed. The change in velocity helps the filter distinguish between inertial errors and Doppler errors, because the Doppler errors depend on velocity while the inertial errors do not.

By comparing Equations (4.3) and (4.5), the dynamic equations for the indirect filter appear very similar to those for the direct filter. In the first element of the state vector, Doppler scale factor replaces vehicle acceleration. There is one additional term which gives the feedback of misalignment error into velocity error that is due to the closing of the Schuler loop. Although the filters appear similar, they have very different characteristics. In the direct filter, the system is dominated by white noise driving the acceleration of the vehicle. The contribution of the inertial errors to the measurement is so small by comparison as to be very nearly unobservable. Also, the sample rate for the filter must be very high to accurately follow the accelerations of the vehicle. On the other hand, the measurement in the indirect filter is made up entirely of system errors - either inertial or Doppler. Thus the filter can very effectively identify and remove navigation errors and is almost independent of vehicle motions. Also, the sample rate only has to be fast enough to keep up with the slowly changing inertial and Doppler errors. Finally, since the system is corrected after each measurement, the predicted measurement difference is zero and there is *no* need to extrapolate the error state variables.

5. APPLICATIONS, CONSEQUENCES, AND PROBLEMS OF FILTERING

Besides the navigation problem, Kalman filtering also greatly aids the initial alignment and calibration of an inertial system. A filter of the same form as that used for navigation can be used, during pre-flight, to process external information in order to estimate the misalignments and miscalibrations of the system. The external information may be just the fact that the aircraft is not moving so that any bias velocity is velocity error. The state variables that are used in the pre-flight normally include platform misalignments, gyro drifts, and any other terms which might be subject to day-to-day variations, such as gyro and accelerometer scale factors. In addition, the calibration program might be run with the platform in several different attitudes to help the filter separate error sources.

The calibration problem and the navigation problem are so similar that in many cases the same computer program can be used. It is necessary to include in the navigation filter most of the major inertial system error sources, such as misalignments and gyro drifts. These parameters are the same ones that must be estimated during the pre-flight calibration program. Thus, with only quantitative changes in some of the statistical values describing the system, the navigation filter will also handle the calibration problem. For example, a filter which is designed to use Doppler radar measurements can be used for calibration by making the Doppler measurement zero and by greatly reducing the assumed error in the measurement.

In most cases, filtering techniques will both improve the accuracy of the calibration process and shorten the time required. A filter can easily handle the problem of motion during the alignment, such as wind buffeting, and can also be mechanized to account for the changing characteristics of components during warm-up. The same type of filter can be used to transfer alignment from a master system to another inertial system, such as in aircraft carrier operations. Alignment and calibration are considered further in Section 7.

Filtering techniques are also useful for post-flight analyses of navigation systems during a test program. In this application, all necessary system data is recorded during flight, along with all available reference data. The reference data is used in the same way as it is used in flight; but reference data is usually much more accurate and the error model for the navigation system can be much more complete. In post-flight analyses even better use of the information can be made by extending the filter to include optimum smoothing techniques, which means that reference data are used from both sides of the point in question.

One of the most important consequences of the use of statistical filtering for navigation is that it changes the basic figure of merit for the component parts of the navigation system. The dynamic and statistical nature of the error becomes very important, in addition to its absolute accuracy. What is ultimately important is the total accuracy of the overall system after the measurements have been processed by the filter. Thus, what is important is the ability of the filter to detect and compensate for the errors. For example, if a component has poor day-to-day stability but the error is nearly constant once warmed-up, the total system accuracy might be better than with a component with better absolute accuracy but with less stability during the flight. The parameters that give a measure of this lack of stability during flight are the noise inputs to the error equations. These are given by the N and U matrices in the filter equations (see Equations (3.3)). Thus, one of the important parameters of a component is the size of this noise term, which is a part of the error model for that component. For example, if gyro drift can truly be described by a random-walk model as used in the example in Section 4, then an important figure of merit for a gyro is the increase in mean-squared gyro drift per unit time. The units for this parameter would be $(\text{meru})^2/\text{hour}$. When a filter is used, this number might be more important than the drift rate itself.

An even more important consideration is the dynamic characteristics of the errors. If a filter is used to combine the information from two systems which have errors with the same dynamic characteristics, then the filter

can do little more than produce a statistical average which might reduce the error around 20 to 30%. On the other hand, if the errors have unique dynamic characteristics, then the filter will be able to distinguish the errors and obtain a possible 70 to 80% improvement. As was mentioned in the examples, this result is advantageous when an inertial system and Doppler radar are combined. Most inertial errors are related to some inertial direction while Doppler errors are related to aircraft axes. Thus, when a turn is made, the filter is, to some degree, able to distinguish and calibrate the errors.

A very valuable by-product of the use of statistical filtering is that an estimate is available at all times for the accuracy of the navigation system. This feature has at least three possible uses. First, the information could be displayed directly. The navigation system would then give the vehicle operator not only the indicated position and velocity, but the estimated accuracy of these numbers. In some situations this information could be invaluable. Secondly, a record could be kept of the actual and estimated error in the system, for example, at terminal points. If there were, on an average, too much discrepancy between how well the system was doing and how well it thought it was doing, this would indicate that the system was not operating properly. The problem could be due either to some component which was exceeding specifications, or to improper modeling of the system. A third use of this statistical data would be the automatic editing of the input data. The filter already has an estimate of the expected error in the measurements. Thus, if a measurement is in error by more than three or four standard deviations, it can be automatically rejected and indication can be given to the operator that something might be wrong.

Two of the more important problems in the use of statistical filtering are the requirements for a large flight computer and the need for adequate statistical models of the component parts of the system. The computational problem is due to the necessity of computing, in real time, the optimum statistical weighting factors. The weighting factors involve the integration of a matrix differential equation for the covariance matrix (see Equations (3.3)). This matrix has dimensions $n \times n$, where n is the number of variables being estimated by the filter. The amount of computation required is roughly proportional to n^3. Thus, the more complete the model, the worse the computational problem. For a particular navigation system there will be a trade-off between accuracy and computer capability. For any system there will be a point beyond which a larger computer gives very little improvement. Considerable knowledge of the system is needed to actually determine the accuracy/computer-size trade-off. This fact leads into the other major problem. Some possible solutions to the computation problem are given in the next section.

The other major problem is the need for statistical models for all the instruments involved in the navigation system. Computer simulations of typical navigation problems using supposedly realistic statistical models produce outstanding performance when Kalman filtering is used. But the results in the real world are not necessarily this good because of unsuspected errors with which the filter is not capable of coping. A detailed statistical error model for an inertial system has not really been essential in the past, but now that the model is actually a part of the system, its determination is much more important. To be confident of having a complete model, it is necessary to test the system in an operational environment as near as possible to the one in which the system is actually going to be used. This flight test program is necessary to assure that the statistical filter can handle any peculiarity in the system. It is very helpful in this flight test program to record all information that is necessary to "re-fly" the flight on a ground based computer. With this recording it is possible to change the filter and determine what the results would have been without having to re-fly the aircraft. It is thus possible to optimize the filter with great savings in flight test expenses.

6. METHODS FOR REDUCING COMPUTATION

The first method for reducing the computational load is to eliminate elements of the covariance matrix that have a negligible contribution to the operation of the system. The linear filter given by Equations (3.3) gives the "mathematically optimum" filter for the model that was assumed to represent the physical system; i.e. terms are included even if they would reduce the mean-squared uncertainty only in the ninth or tenth significant figure. The assumed statistical model, however, represents the actual system to no more than two or three significant figures. The problem is to locate those terms that can be eliminated without significantly degrading the performance.

A method for producing an effective sub-optimum filter, which the authors have used with some success, is to use physical insight into the particular problem to identify the primary flow of information in the optimum filter. The optimum filter is then separated into smaller filters that preserve this primary flow of information while neglecting terms which have negligible effect. The details of the design of the sub-optimum filter are obtained by simulating both the optimum and sub-optimum filters for identical, typical problems. The design of the sub-optimum filter is then adjusted so that its performance matches that of the optimum filter as nearly as possible. This simulation process also indicates the efficiency of the sub-optimum filter. This approach has been relatively successful, particularly for simplifying the computation of the covariance matrix.

The second method for simplifying the optimum filter is to use pre-computed gains. If the filter equations are such that the gains at each measurement are only functions of time and of the *a priori* assumptions of the statistics of the noises and the initial state, then, by specifying the measurement schedule or rate, the gains may be pre-computed and stored in the vehicle's computer. For a small number of measurements this is a practical solution to the problem involved in implementing an optimum filter. For a large number of measurements, the pre-computed gains are usually smoothly varying with respect to time and may be approximated by suitable curves (straight lines, exponentials, etc.) that give almost identical filter responses as the true gains. Some of

the minor gains may even be neglected completely. This technique will be applied to the pre-launch calibration and alignment of an inertial platform in a spacecraft on top of a swaying launch vehicle and is, in fact, similar to the technique used in the Apollo Guidance Navigation and Control System.

7. PRE-LAUNCH CALIBRATION AND ALIGNMENT

7.1 Physical Basis for Calibration and Alignment

The inertial system to be calibrated and aligned includes gyroscopes and accelerometers. The known gravity acceleration magnitude is used to calibrate the accelerometers; the known inertial rate of rotation of the gravity vector at a point on the surface of the earth is used to calibrate the gyros. In this procedure, the system is approximately aligned to the local vertical coordinates (vertical, south, and east), then the gyros are used to instrument an inertial coordinate system. The two approximately horizontal accelerometer outputs (south and east) are used by the optimum filter to generate estimates of the system misalignment with respect to the local vertical coordinates and of the gyro drifts by comparing the measurement of the rotation of the gravity vector with the known rotation rate in the inertial coordinates instrumented by the gyros. The vertical gyro drift is the most difficult quantity to measure since it causes only a third-order effect on the measured acceleration. Gyro failures can be closely associated with changes in drift due to acceleration of gravity along the input axis, so the pre-launch calibration of a gyro in a vertical position is highly desirable.

Since the estimates of the alignment and drift variables will depend on the measurement by the accelerometers of the rotation of the gravity vector in the inertial coordinates instrumented by the gyros, the major disturbances are the accelerometer quantization and the wind-induced sway of the launch vehicle. The model of the system for the optimum filter must include variables due to the sway. The complete filter must be simulated on a digital computer; it will be linear, so that by specifying the measurement schedule the optimum gains may be pre-computed. The gains will be approximated by functions that will be stored in the flight computer. The method for using this simplified filter in other platform positions, an illustration of a system test program, and practical hardware problems will be presented.

7.2 Models

In this section the models of the launch vehicle and the inertial system will be presented. The launch vehicle bending dynamics in the north-south and east-west directions are approximated by identical second-order systems. The wind causing the vehicle sway is assumed to be exponentially correlated with a correlation time of $1/\lambda$ sec. For convenience we assume the wind in each horizontal direction is uncorrelated with the other direction, as is indicated in Figure 21. The correlation function of the white noise required to produce the predicted value of the missile mean-squared sway can be found to be[12]:

$$nw\delta(t-\tau) \;=\; \langle n(t)n(\tau)\rangle \;=\; \frac{\langle p^2\rangle 4\lambda\zeta\omega_n^2(\omega_n^2 + 2\lambda\zeta\omega_n + \lambda^2)}{\lambda + 2\zeta\omega_n}\,\delta(t-\tau)\,, \tag{7.1}$$

where $\langle p^2\rangle$ is the predicted mean-squared missile sway, $n(t)$ is the white noise generating the exponentially correlated wind, and ω_n and ζ are the natural frequency and damping ratio of the second-order approximation to the bending dynamics. The state vector for the sway variables in the south direction is

$$\begin{bmatrix}\dot{p}_s\\ \dot{v}_s\\ \dot{a}_s\end{bmatrix} = \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ -\lambda\omega_n^2 & -\omega_n^2 - 2\zeta\lambda\omega_n & -2\zeta\omega_n - \lambda\end{bmatrix}\begin{bmatrix}p_s\\ v_s\\ a_s\end{bmatrix} + \begin{bmatrix}0\\ 0\\ n(t)\end{bmatrix} \tag{7.2}$$

where p_s, v_s, a_s are the horizontal displacement, velocity and acceleration in the north-south direction. The model for the sway variables in the east direction (p_e, v_e, a_e) has the same form. For the computer simulations to follow, they have the values: $\langle p^2\rangle = 100$ cm^2 east and south, $\lambda = 0.1$ sec^{-1}, $\omega_n = 2.09$ rad/sec, and $\zeta = 0.1$.

The platform orientation with respect to the local vertical coordinate system is described by three small angles α, β, and γ. If the platform axes (x, y, z) were rotated by $-\alpha$, $-\beta$, and $-\gamma$, the axes would coincide with the local vertical coordinates as shown in Figure 22. The state vector equation for this sub-state is given by

$$\begin{bmatrix}\dot{\alpha}\\ \dot{\beta}\\ \dot{\gamma}\\ \dot{d}_x\\ \dot{d}_y\end{bmatrix} = \begin{bmatrix}0 & 0 & \Omega_n & 1 & 0\\ 0 & 0 & \Omega_v & 0 & 1\\ -\Omega_e & -\Omega_v & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0\end{bmatrix}\begin{bmatrix}\alpha\\ \beta\\ \gamma\\ d_x\\ d_y\end{bmatrix} + \begin{bmatrix}t_x - \Omega_v\\ t_y + \Omega_n\\ t_z + d_x\\ 0\\ 0\end{bmatrix}, \tag{7.3}$$

which has been assumed to be, for angle magnitudes of interest, a valid representation of the platform dynamics. Ω_h and Ω_v are the horizontal and vertical components of earth rate at the test site; d_x, d_y, d_z are the constant drifts for the vertical, south, and east gyros; and t_x, t_y, t_z are the torquing rates (if any) applied to the gyros. It is assumed that the torquing rates, the components of earth rate, and the east gyro drift are known perfectly, so that the vector on the right represents known control $g(t)$ and is independent of the state of the system.

Using the small angle approximations, the south and east accelerometer pulse rate outputs due to platform orientation in the gravity field may be written as

$$\begin{bmatrix} \dot{po}_s \\ \dot{po}_e \end{bmatrix} = g \begin{bmatrix} -\gamma \\ \beta \end{bmatrix} . \tag{7.4}$$

where po_s and po_e represent the total pulse counts at some instant of time, and g is the local gravity (cm/sec^2). A 1 cm/sec pulse accelerometer quantization has been assumed.

Accelerometer pulse rates cannot be instantaneously measured, but the total pulses, which make up the output due to sway velocity and orientation in the gravity field, can be counted. The filter must distinguish between the pulse count due to the high frequency sway and the pulse count due to the slow rotation of the gravity vector in the inertial coordinates instrumented by the gyros. Inherent in these measurements are quantization errors. The measurement vector is

$$\underline{m} = \begin{bmatrix} m_s \\ m_e \end{bmatrix} = \begin{bmatrix} v_s \\ v_e \end{bmatrix} + \begin{bmatrix} po_s \\ po_e \end{bmatrix} + \begin{bmatrix} n_s \\ n_e \end{bmatrix} . \tag{7.5}$$

The term "n_m" represents the quantization error at every sampling of the accelerometer pulse count registers. Although the quantization error is uniformly distributed, the measurement noise is assumed to be gaussian with zero mean. The measurements are made once a second with

$$\langle n_m^2 \rangle = ra . \tag{7.6}$$

The state vector is 13-dimensional; the state vector differential equation is

$$d\underline{x}/dt = F\underline{x} + \underline{c} + \underline{n}(t) . \tag{7.7}$$

(See Figure 23 for $\underline{x}$, F, $\underline{c}$, and $\underline{n}$.)

With the derivation of the model for the system, the complete optimum linear filter is defined. The accelerometer pulse count registers will be sampled at constant rates. The estimated state vector is extrapolated between measurements according to

$$\hat{\underline{x}}' = \Phi\hat{\underline{x}} + \underline{c} , \tag{7.8}$$

and the covariance matrix according to

$$E' = \Phi E \Phi^T + S , \tag{7.9}$$

where Φ and S are pre-computed constant matrices for the time step of one second between measurements. They satisfy the following differential equations:

$$dS/dt = FS + SF^T + N \qquad S(0) = 0 \tag{7.10}$$

and

$$d\Phi/dt = F\Phi , \qquad \Phi(0) = I \tag{7.11}$$

which may be integrated on a digital computer for a time step between measurements of 1 second. At the time of a measurement, $\hat{\underline{x}}$ and E are changed according to Equations (3.3) and H, U, and N are given in Figure 23.

7.3 Computer Simulations

A complete nonlinear simulation of the inertial platform in a swaying launch vehicle was made on a digital computer in order to simulate real accelerometer outputs. The initial misalignments were 1 degree on all axes; drifts were 10 meru for the vertical and south gyros and zero for the east gyro. (A meru is approximately 0.015 deg/hr.) The r.m.s. sway in each horizontal direction was 10 cm. The initial conditions for the covariance matrix were 1 deg^2 for the alignment angles and 100 meru^2 for the gyro drifts. The sway variable variances were 100 cm^2, 440 cm^2/sec^2, and 1900 cm^2/sec^4 for position, velocity, and acceleration. All initial cross-correlation terms were assumed zero. The initial estimate of the state was a zero vector.

The response of the filter was excellent for these conditions. A number of runs were first made in which the matrix U was varied so as to cause good agreement between the r.m.s. error as determined by the filter and the actual error. The errors between the estimates and the actual values of azimuth angle, vertical gyro drift and south gyro drift are shown in Figures 24-26. The errors reach small values for the three cases in 15, 40, and 10 minutes respectively. Note that, since the measurements are taken every second, about 2400 samples are

required to estimate the drift of the vertical gyro. Clearly, any significant errors in the model for this system will result in drastic errors in the vertical drift estimation. The errors in the estimates of the two leveling angles (α and β) are negligible after the first few measurements. The estimates of the sway variables are not particularly good, since we are sampling once per second and the sway period is 3 seconds. These simulations were run under assumed perfect knowledge of east gyro drift because of the classical result that east gyro drift cannot be identified from azimuth error. The fact that east gyro drift must be known presents no problem; as seen from Figure 26, the south gyro can be calibrated in about 10 minutes and the error shows little sensitivity to d_z. This gyro can then be placed east by rotating the platform and a complete alignment and calibration made. (The question of other platform positions is discussed later.)

7.4 Design of the Simplified System

The gains for the optimum filter may be pre-computed for all trials, since the measurement times will be the same and the *a priori* assumption for the statistics of the initial state vector and noises will not change. For the problem at hand, the implementation of the gains into the inertial system involved, first, the design of a simplified optimum filter. The gains for each state variable estimate depend on both accelerometer measurements and, in general, one gain is much smaller than the other and can be neglected. In this problem all cross-coupling measurement gains are neglected; e.g., vertical drift estimation depends primarily on the south accelerometer, so the east accelerometer measurement gain for vertical drift estimation is not implemented. Typically the predominant gains vary as in Figure 27. The po_s gain represents the number used to multiply the difference in the measured and predicted output of the south accelerometer and to update po_s. The α gain multiplies the same difference. These gains can be approximated by exponentials and straight-line segments where, at distinct intervals, the time constants and slopes are changed to continually fit the approximate gains to the true gains. The gains for position, velocity, and acceleration quickly reach steady-state values and may be approximated by three constants.

The response of a simplified filter is shown in Figures 24 - 26. The process of design enters, since it required a number of runs using different slopes and time constants for the gains to get a good match with the response of the complete filter. In fact, in the end, the same exponential gains were used for both po_s and po_e, and the same exponential gain magnitude was used for β and γ. The total pre-computed constants were three sway variable gains, two initial conditions for exponentials, and sets of the following five numbers which are changed at ten discrete times: two time constants for exponentials (po_s and β) and three slopes for straight-line segments (α, d_x, d_y). Table I summarizes these comments.

TABLE I

Approximate Gains

Variable	*Accelerometer used*	*Form of approximate gain*
po_s	south	exponential segments
po_e	east	same value as po_s gain
v_s	south	constant
v_e	east	same as v_s
a_s	south	constant
p_s	south	constant
a_e	east	same value as a_s
p_e	east	same value as p_e
α	south	straight-line segments
β	east	exponential segments
γ	south	$-\beta$ gain
d_x	south	straight-line segments
d_y	east	straight-line segments

7.5 Implementation

A slight variation of this simplified filter was implemented in the Apollo Guidance Navigation and Control System. Part of the program was concerned with initialization for platform positions other than the one considered here. The optimum filter, once implemented, does not change for other platform positions; the measurements that the filter gets are made to simulate the standard platform configuration. For example, if the platform axes were vertical, north, and east, and if the sign of the north accelerometer output were changed to simulate a south accelerometer, the filter output for the variables in the south direction need only be interpreted as negative of their true values.

For some applications, it may be desirable to torque the south gyro at negative horizontal earth rate to keep the east axis level. The form of the filter and the filter gains do not change because perfect torquing is assumed; just add negative horizontal earth rate to the extrapolation of the angle β. It has also been found convenient to use a simple first-order extrapolation of the alignment angles ($\alpha = \alpha + \dot{\alpha}\,dt$, etc.). The sway variables are extrapolated according to a sway transition matrix whose elements can be changed to compensate for variations in parameters between different launch vehicles.

Once the optimum filter has been implemented according to the simple method outlined, it can be readily adapted to various problems of alignment and calibration. For example, consider the following system test procedure in which the platform axes are identified as x, y, z:

(1) run a 10-min test with x up, y south, z east to determine y gyro bias drift;

(2) at 10 min read out y drift and use the angle estimates to align the platform to the local vertical coordinates; continue to hold that orientation by torquing the gyros at negative earth rate for 90 sec while counting pulses from the x accelerometer;

(3) orient the platform to x down, y east, z south and run a 10-min test to determine z bias drift;

(4) use the angle estimates at 10 min to align the platform and then torque for 90 sec at negative earth rate while counting x accelerometer pulses;

(5) torque the south gyro with the negative of the horizontal component of earth rate for 45 min while determining vertical drift (x gyro). The y gyro bias, as determined in step 1, is d_z for this step.

This procedure takes about 68 minutes, after which enough information is available to determine y gyro bias drift, z gyro bias, the sum of z gyro bias and acceleration-sensitive drift, and x accelerometer bias and scale factor. One can readily imagine how an automated system test procedure can be set up to completely calibrate the system in the swaying spacecraft. The last step in the program would be an alignment run to ready the system for launch. Through all of this, the basic simplified optimum filter does not change.

7.6 Hardware Problems

The primary source of azimuth error is the uncertainty in calibration of the east gyro drift; the primary source of vertical drift estimation error is due to variations in the east gyro drift during a test. If the east gyro has a large drift due to acceleration along its input axis, then it is desirable to keep the input axis almost horizontal to minimize the variation in east gyro drift by torquing the south gyro at negative horizontal earth rate. Unfortunately, if the bias of the south gyro changes when it is torqued, then the south gyro calibration will yield two answers corresponding to the torqued and untorqued cases, respectively. (Torquing was done in the system test program.) Another problem could occur if the accelerometer biases change during the calibration. Figure 28 shows the effect of an exponentially changing east accelerometer bias on south drift estimation.

Another possible problem area appears when using pulsed integrating accelerometers. If either horizontal accelerometer has a large dead-zone for near zero inputs, large transients in the filter output will appear. The form of the transients will vary depending on the time during the test that the accelerometer goes through the dead-zone. Vertical drift estimation is particularly sensitive to a dead-zone in the south accelerometer. In some cases a transient on the order of 800 meru has been observed; the filter never reached the correct value of drift at the end of 45 minutes because this vertical drift gain is small at the end of the test. As a practical solution to the problem, the platform is deliberately offset (between steps 4 and 5 in the system test program) from the vertical before beginning a vertical drift test so that the accelerometer never goes through null.

These problems are indicative of the strange results that can occur when the model for the system is incorrect. Philosophically, we have designed a total system test and as such it should indicate, in some manner, out-of-specification conditions which would then require lower level testing. The model must therefore include all in-specification conditions and the designer's experience must be used to recognize out-of-specification situations.

7.7 Laboratory Test Results

A significant portion of the filter design was done in the laboratory, since digital simulation can never equal actual hardware testing in pointing out the important problem areas. As part of the filter implementation verification, the Apollo inertial platform was suspended by a cable from the laboratory ceiling. The platform was pushed to simulate spacecraft motion on the launch pad. Comparisons with independent measurement techniques verified the program accuracy to within the limits of the gyro performance.

8. CONCLUSIONS

Navigation systems are one of the most important applications of Kalman filtering. Conversely, Kalman filtering is one of the most important new developments in navigation systems. It is likely that most future navigation systems will involve Kalman filtering in some form. However, considerable engineering experience is required

to develop an effective filter, because the physical systems are never precisely defined, the necessary statistical models are not well known, and the computer is never large enough. It is hoped the discussions in this chapter can be of some help in obtaining the understanding necessary to develop efficient navigation filters.

In summary, we might suggest eight steps for the successful implementation of a filter, such as was illustrated here:

1. Develop models.
2. Digitally simulate model and filter equations.
3. Design simplified filter:
   (a) restrict model,
   (b) decouple variables,
   (c) approximate gains.
4. Digital simulation of system and simplified filter.
5. Implementation into guidance computer.
6. Checkout.
7. Test experience.
8. Recycle to step 1.

The requirement to simulate the complete system with the filter equations has been inserted to demonstrate the filter's ability to estimate the quantities of interest. It has been all-too-common practice to simply run the filter equations and look only at the propagation of the covariance matrix. Finally, as practical experience indicates modeling errors, the sources must be found and practical solutions implemented.

REFERENCES.

1. Kalman, R.E. — *A New Approach to Linear Filtering and Prediction Problems.* Journal of Basic Engineering, Transactions of the American Society of Mechanical Engineers, Vol. 82, March 1960, pp. 35-45.

2. Kalman, R.E.; Bucy, R.S. — *New Results in Linear Filtering and Prediction Theory.* Journal of Basic Engineering, Transactions of the American Society of Mechanical Engineers, Vol. 83, March 1961, pp. 95-108.

3. McLean, J.D. et al. — *Optimal Filtering and Linear Prediction Applied to a Space Navigation System for the Circumlunar Mission.* NASA TN D-1208, March 1962.

4. Battin, R.H. — *A Statistical Optimizing Navigation Procedure for Space Flight.* American Rocket Society Journal, Vol. 32, November 1962, pp. 1681-1696.

5. Åström, K.J. — *Some Problems of Optimal Control in Inertial Guidance.* IBM Research Laboratory, San Jose, California, Research Paper RJ-229, November 1962.

6. Åström, K.J. — *Some Problems of Indicating the Vertical in a Moving Vehicle.* Foa 2 Rapport A 2306-233, Försvarets – Forskningsanstalt, Avdelning 2, Stockholm 80, Sweden, October 1964.

7. Brown, R.G.; Friest, D.T. — *Optimization of a Hybrid Inertial Solar-Tracker Navigation System.* Institute of Electrical and Electronic Engineers International Convention Record, Vol. 12, Part 7, 1964, pp. 121-135.

8. Fagen, S.L. — *A Unified Approach to the Error Analysis of Augmented Dynamically Exact Inertial Navigation Systems.* Institute of Electrical and Electronic Engineers Transactions on Aerospace and Navigation Electronics, Vol. ANE-11, No. 4, December 1964, pp. 234-248.

9. Bona, B.E.; Smay, R.J. — *Optimum Reset of Ship's Inertial Navigation System.* Institute of Electrical and Electronic Engineers Transactions on Aerospace and Electronics Systems, Vol. AES-2, No. 4, July 1966, pp. 409-414.

10. Bona, B.E.; Hutchinson, C.E. — *Optimum Reset of an Inertial Navigator from Satellite Observations.* Proceedings of the National Electronics Conference, Chicago, Illinois, October 1965.

11. Schmidt, G.T. — *The Application of Statistical Estimation Techniques to Inertial Component Calibration.* AGARD-NATO Symposium on Inertial Navigation Components, Braunschweig, Germany, May 1968.

12. James, H.M. et al. — *Theory of Servomechanisms.* Dover Publications, New York, 1965, p.369.

13. Markey, W. Novorka, J. — *The Mechanics of Inertial Position and Heading Indication.* John Wiley, New York, 1961.

14. Schmidt, G.T. Brock, L.D. — *Statistical Estimation in Inertial Navigation Systems.* Journal of Spacecraft and Rockets, Vol.5, No.2, February 1968, pp.146-154.

15. Dunek, H. — *Theory of Error Propagation in Astro-Inertial Guidance Systems for Low-Thrust Earth Orbital Missions.* Proceedings of the American Rocket Society, Preprint No.2682-62, ARS, Los Angeles, California, November 1962.

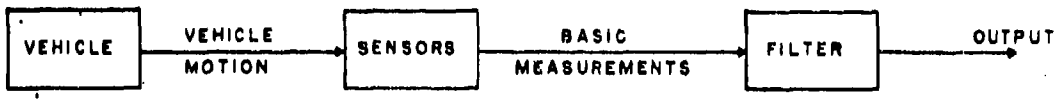


Fig. 1 Navigation problem using direct filter

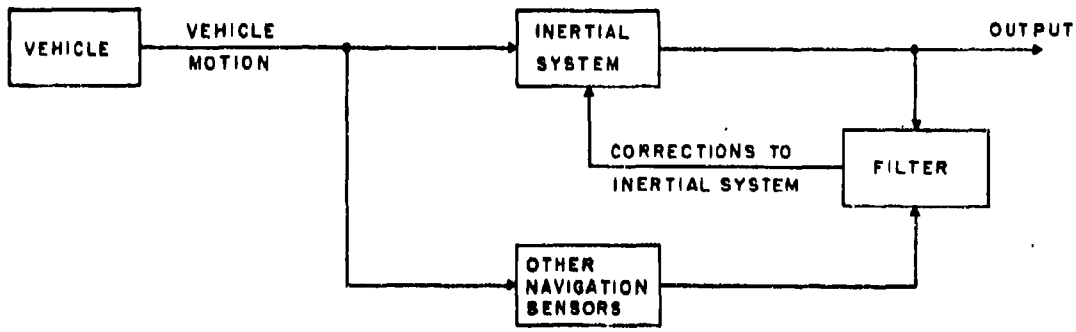


Fig. 2 Navigation problem using indirect filter

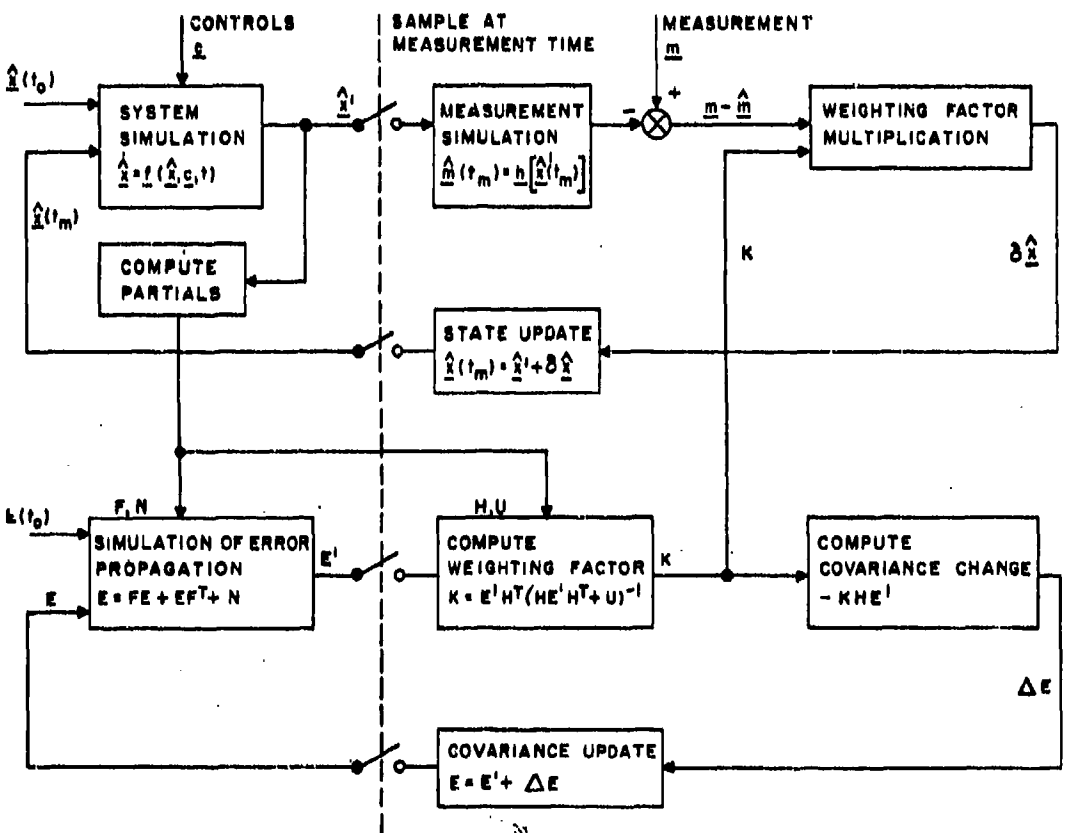


Fig. 3 Kalman filter functional diagram

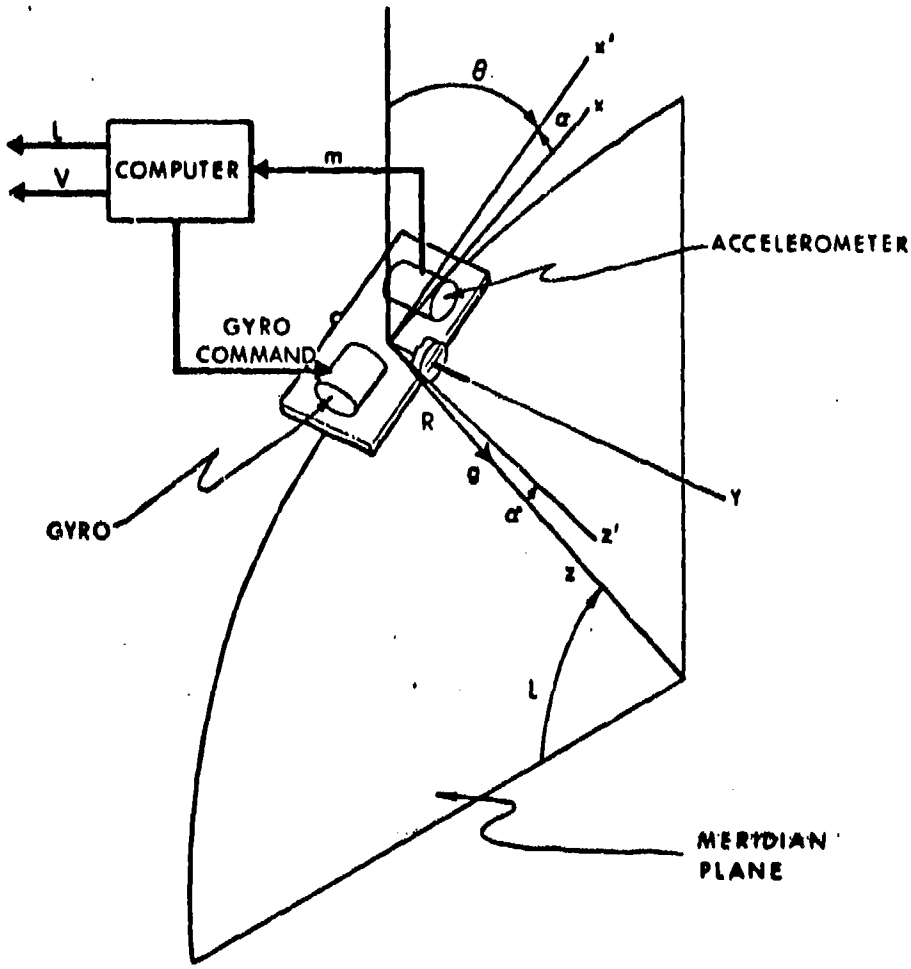


Fig. 4 Navigation system with direct filter

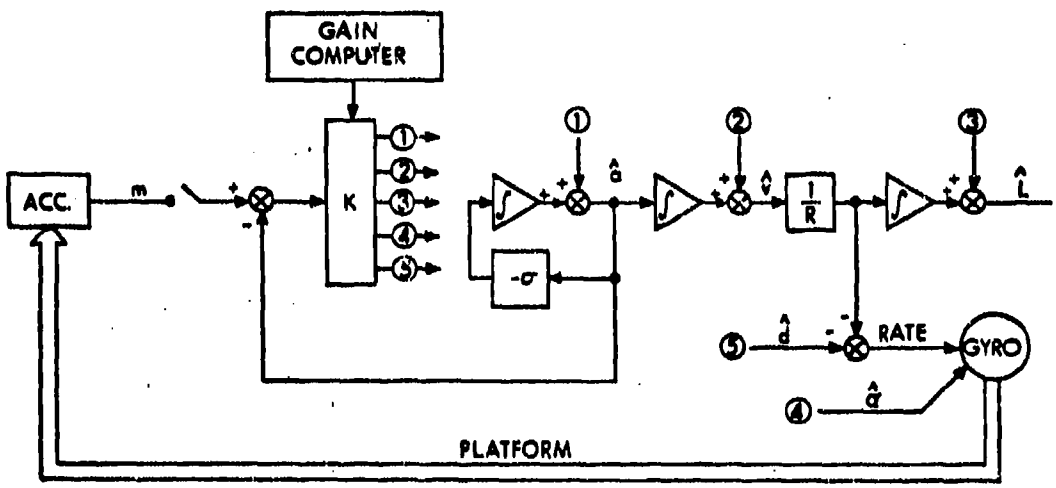


Fig. 5 Diagram of simplified navigation system

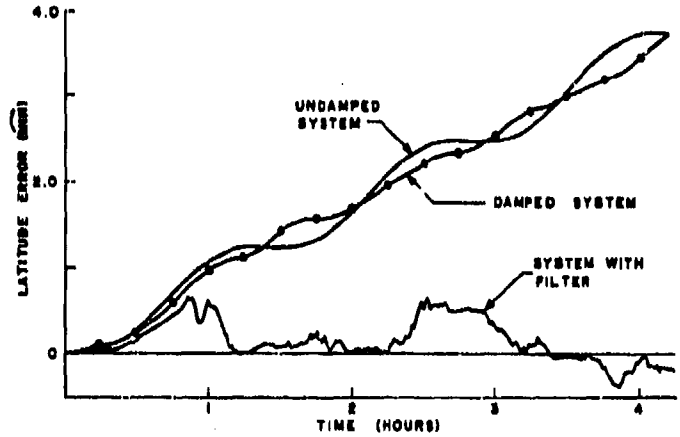


Fig.6 Error in estimated latitude

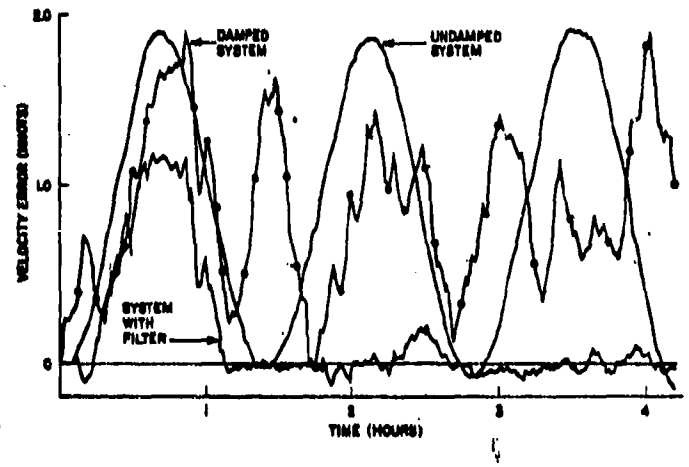


Fig.7 Error in estimated velocity

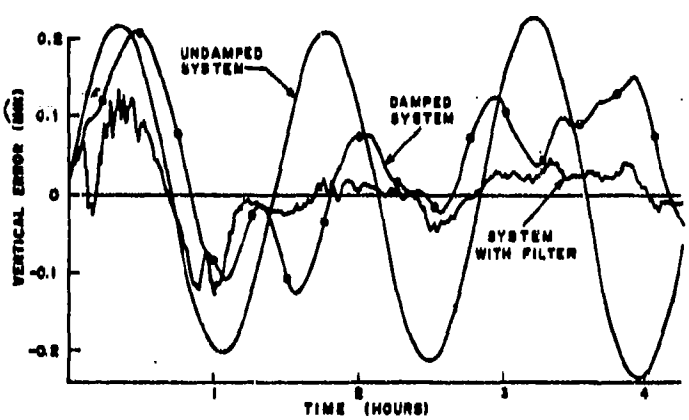


Fig.8 Error in estimated vertical

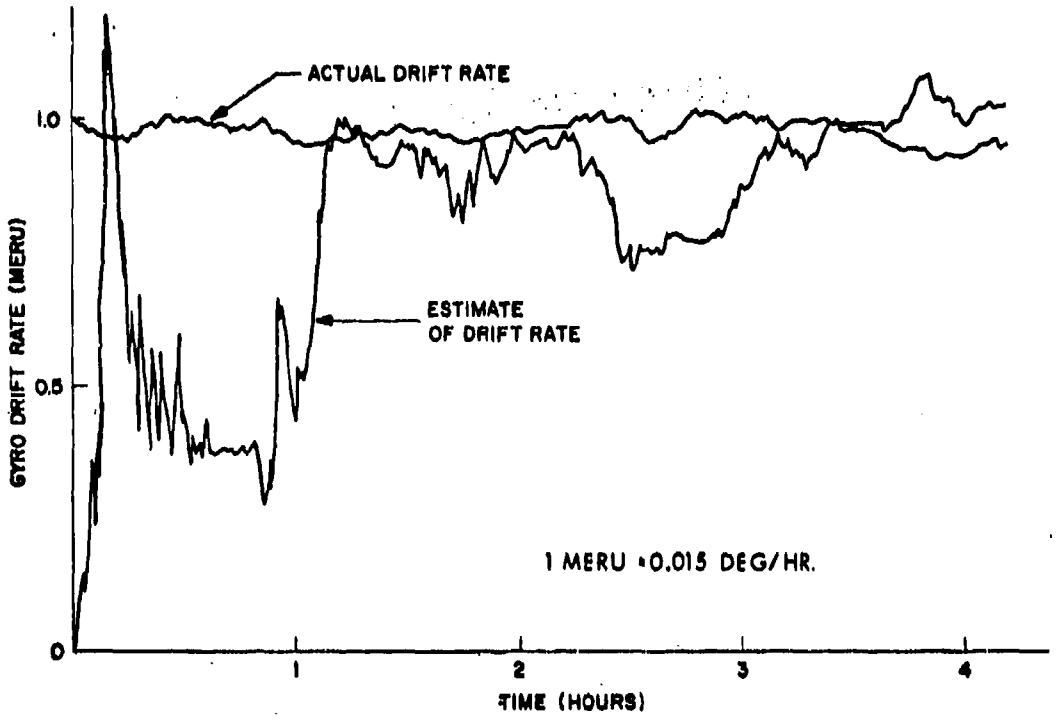


Fig. 9 Estimated and actual gyro drift rate

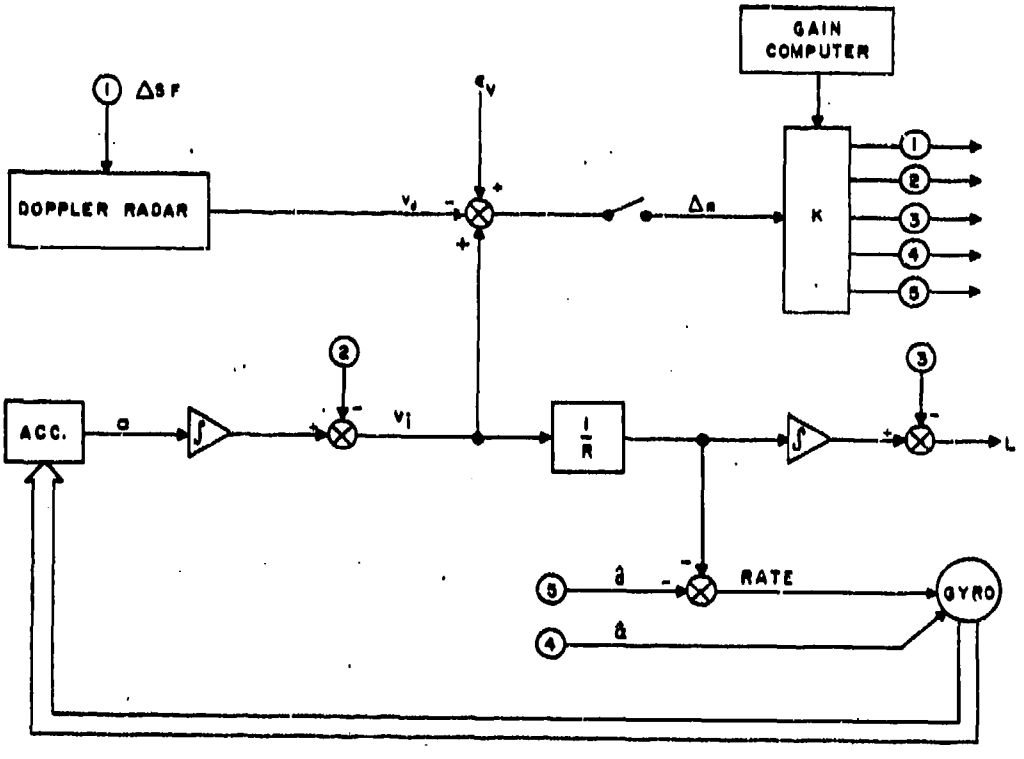


Fig. 10 Navigation system with indirect filter

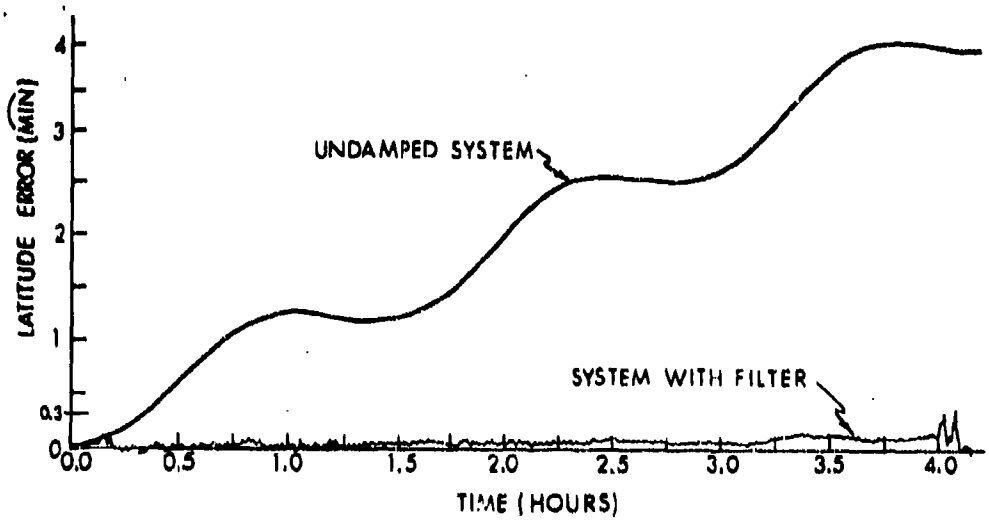


Fig. 11 Error in estimated latitude

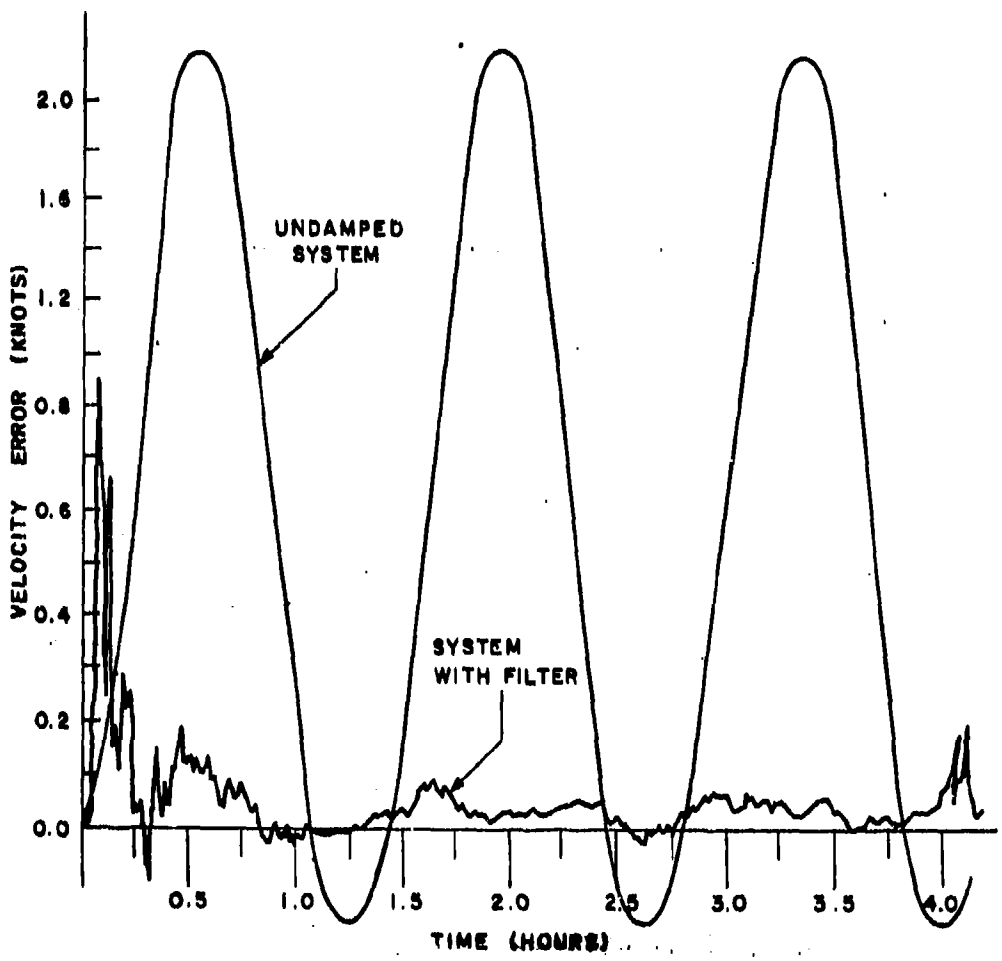


Fig. 12 Error in estimated velocity

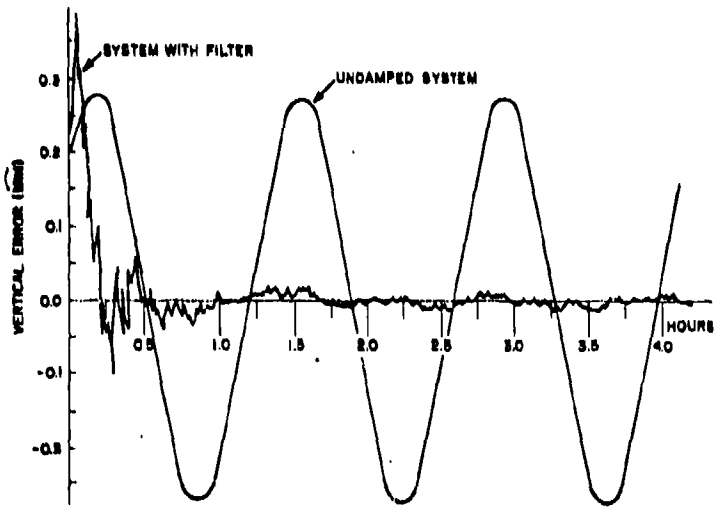


Fig. 13 Error in estimated vertical

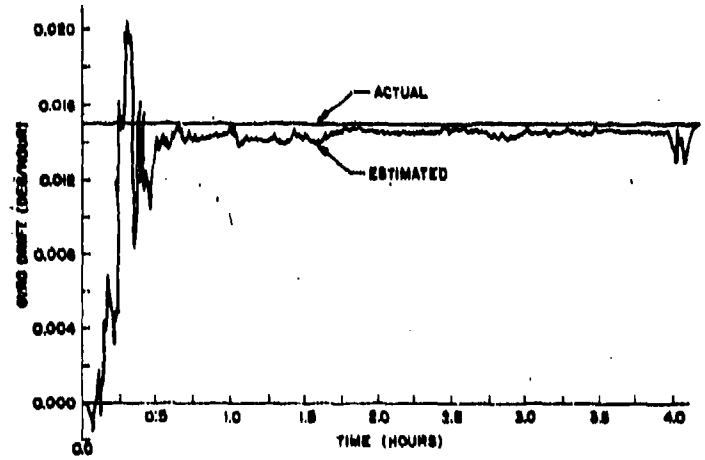


Fig. 14 Actual and estimated gyro drift

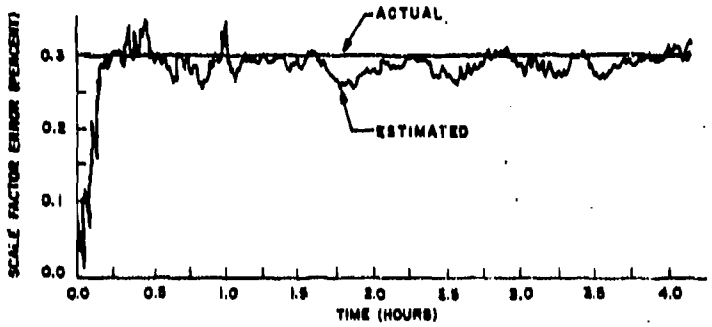


Fig. 15 Actual and estimated scale factor

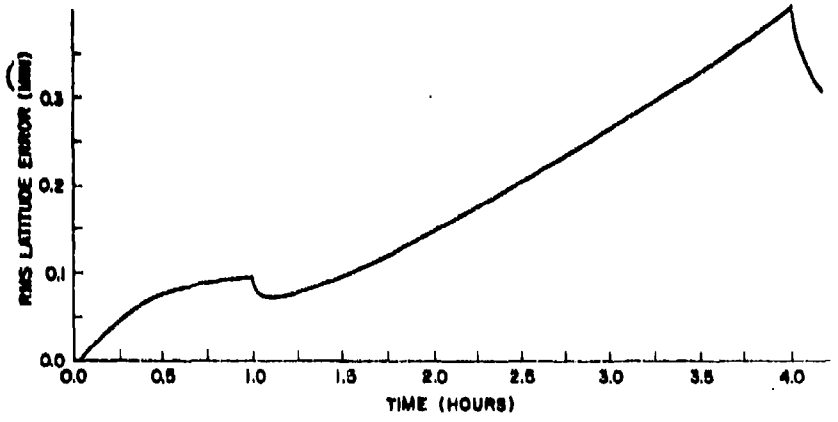


Fig. 16 R.M.S. position uncertainty

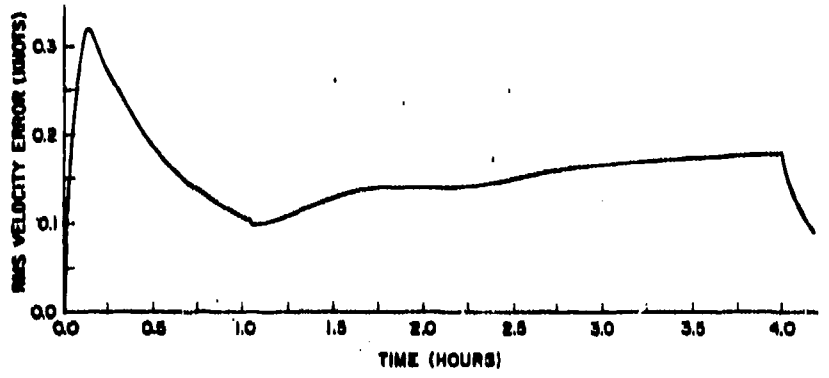


Fig. 17 R.M.S. velocity uncertainty

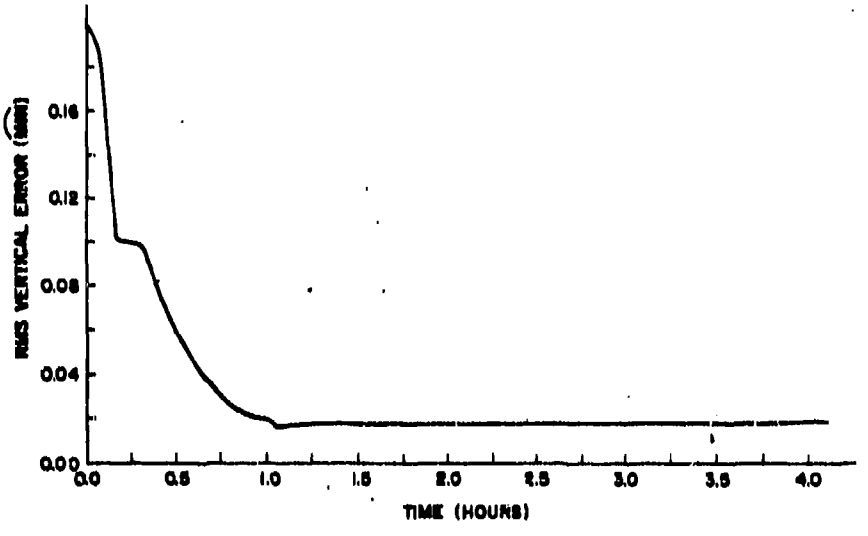


Fig. 18 R.M.S. vertical uncertainty

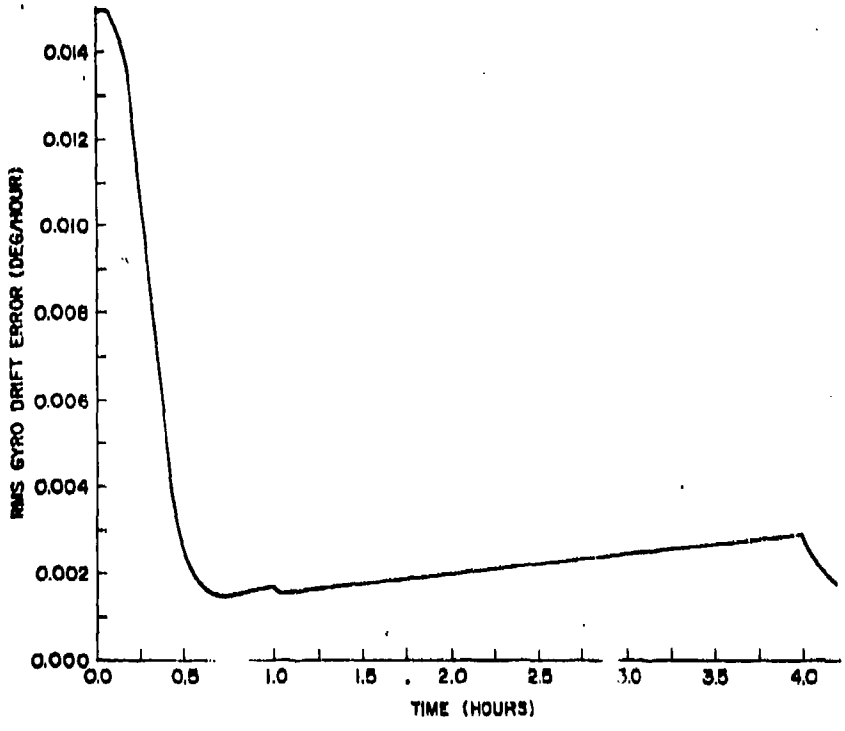


Fig. 19 R.M.S. gyro drift uncertainty

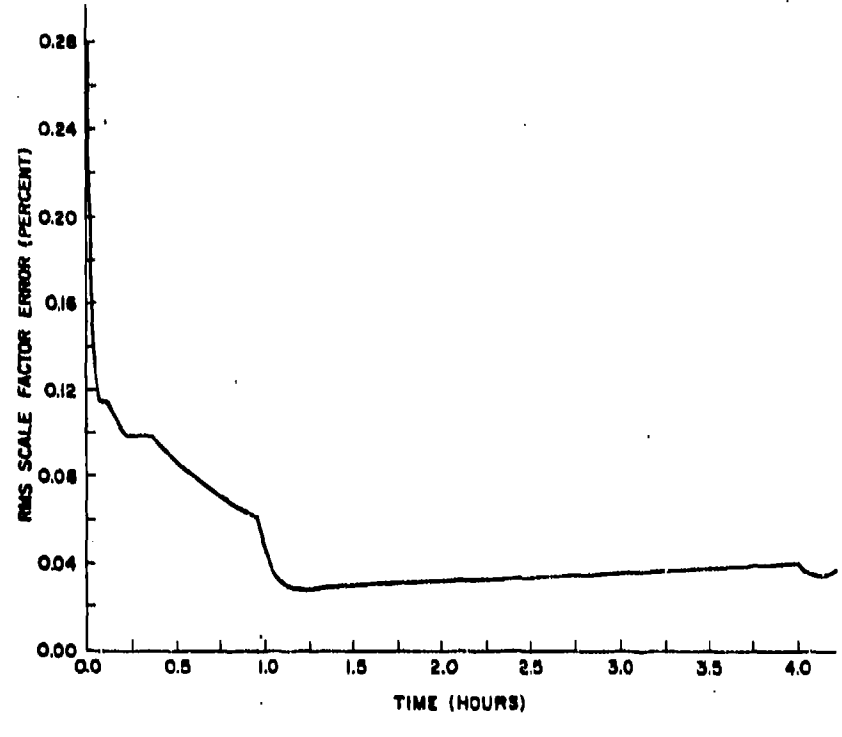


Fig. 20 R.M.S. scale factor uncertainty

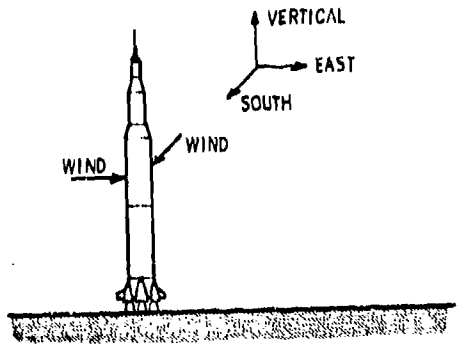


Fig. 21 Launch vehicle subjected to wind

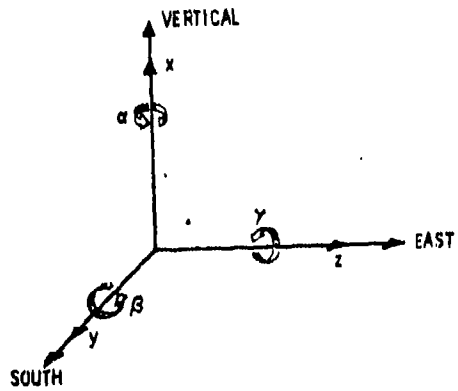


Fig. 22 Small angle definition

$$\underline{x} = \begin{bmatrix} po_s \\ po_e \\ v_s \\ v_e \\ a_s \\ p_s \\ a_e \\ p_e \\ \alpha \\ \beta \\ \gamma \\ d_x \\ d_y \end{bmatrix} \quad \underline{c} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ t_x - \Omega_v \\ t_y + \Omega_h \\ t_z + d_z \\ 0 \\ 0 \end{bmatrix} \quad \underline{n} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ n \\ 0 \\ n \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$F = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -g & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & g & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & f_1 & 0 & f_2 & f_3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & f_1 & 0 & 0 & f_2 & f_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Omega_h & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \Omega_v & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\Omega_h & -\Omega_v & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$f_1 = -\omega_n^2 - 2\zeta\lambda\omega_n$

$f_2 = -\lambda - 2\zeta\omega_n$

$f_3 = -\lambda\omega_n^2$

$$U = \begin{bmatrix} ra & 0 \\ 0 & ra \end{bmatrix} \quad H = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

N = 13x13 ZERO MATRIX EXCEPT FOR ROW 5, COLUMN 5 AND ROW 7, COLUMN 7, WHICH ARE BOTH nw

Fig. 23 Filter matrices and vectors

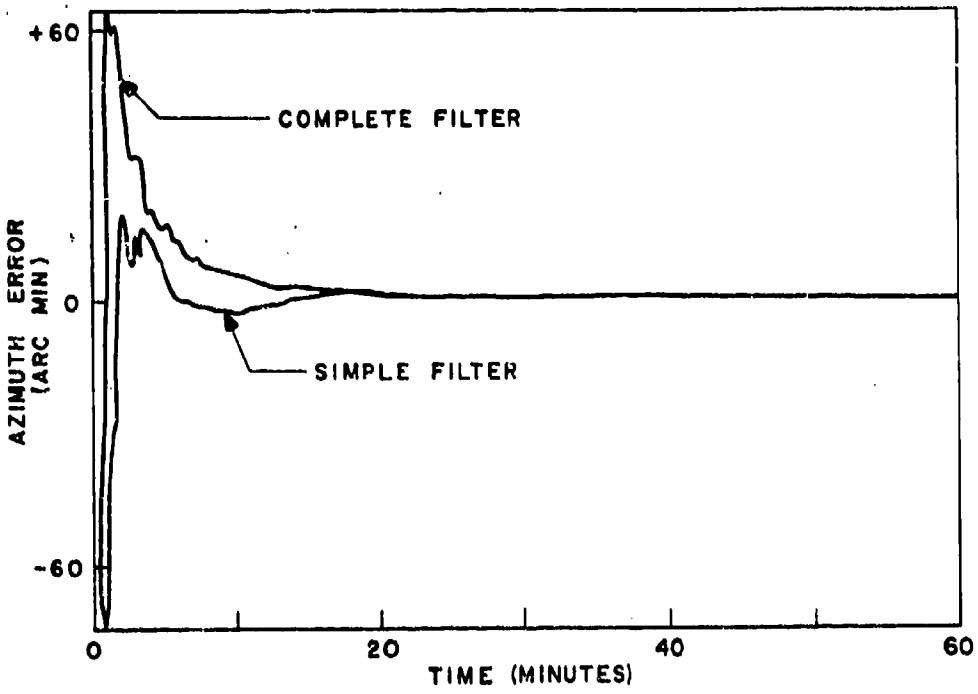


Fig. 24 Azimuth error for complete and simplified filter

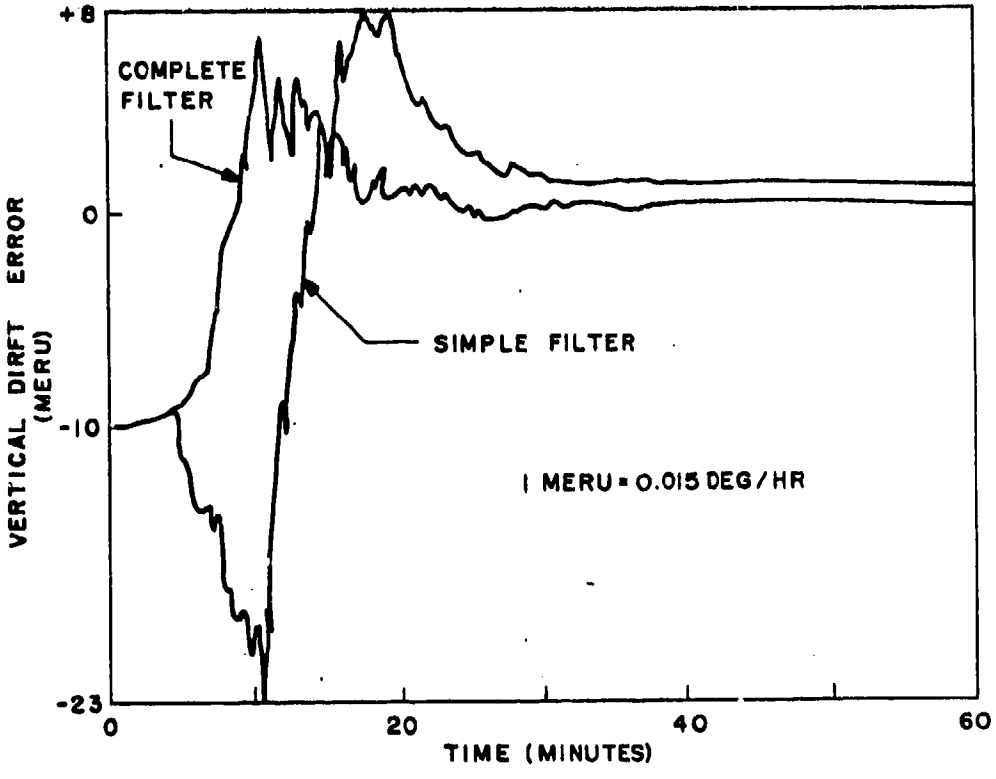


Fig. 25 Vertical drift error for complete and simplified filter

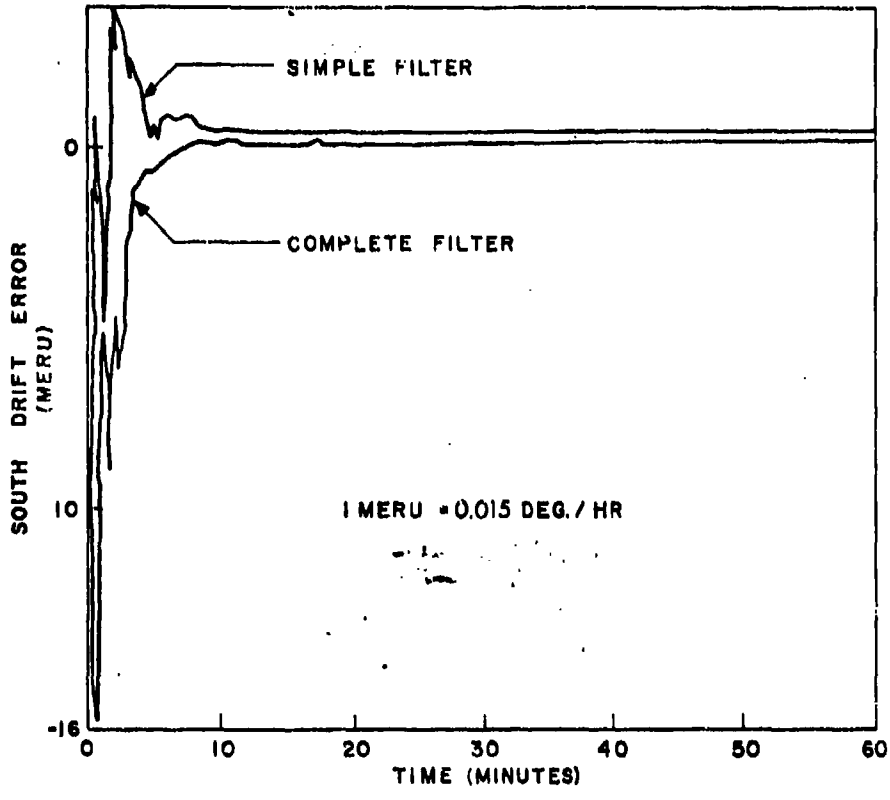


Fig. 26 South drift error for complete and simplified filter

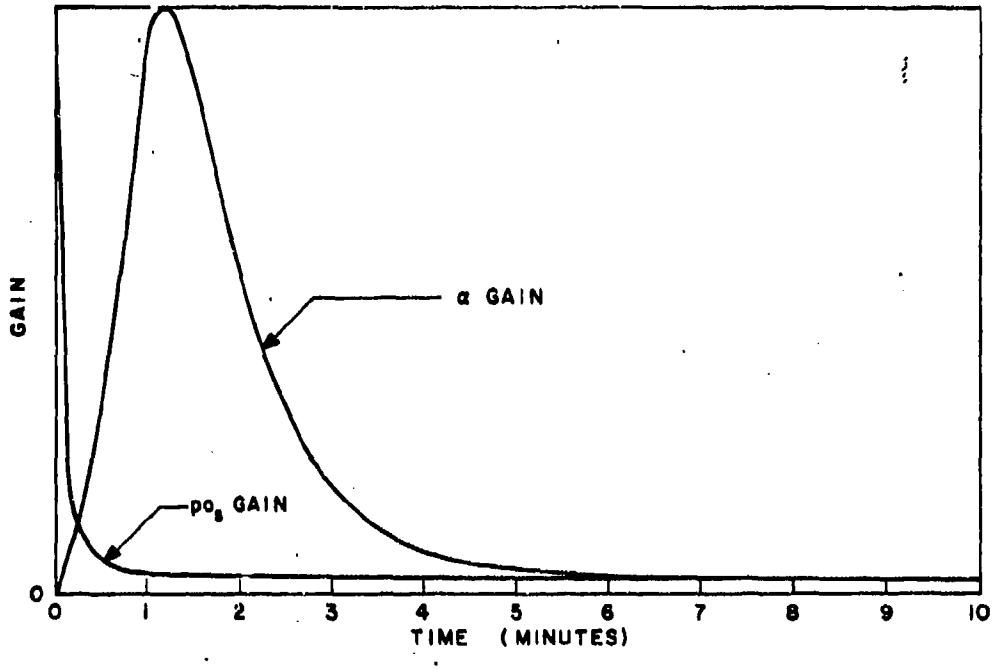


Fig. 27 Form of typical gains

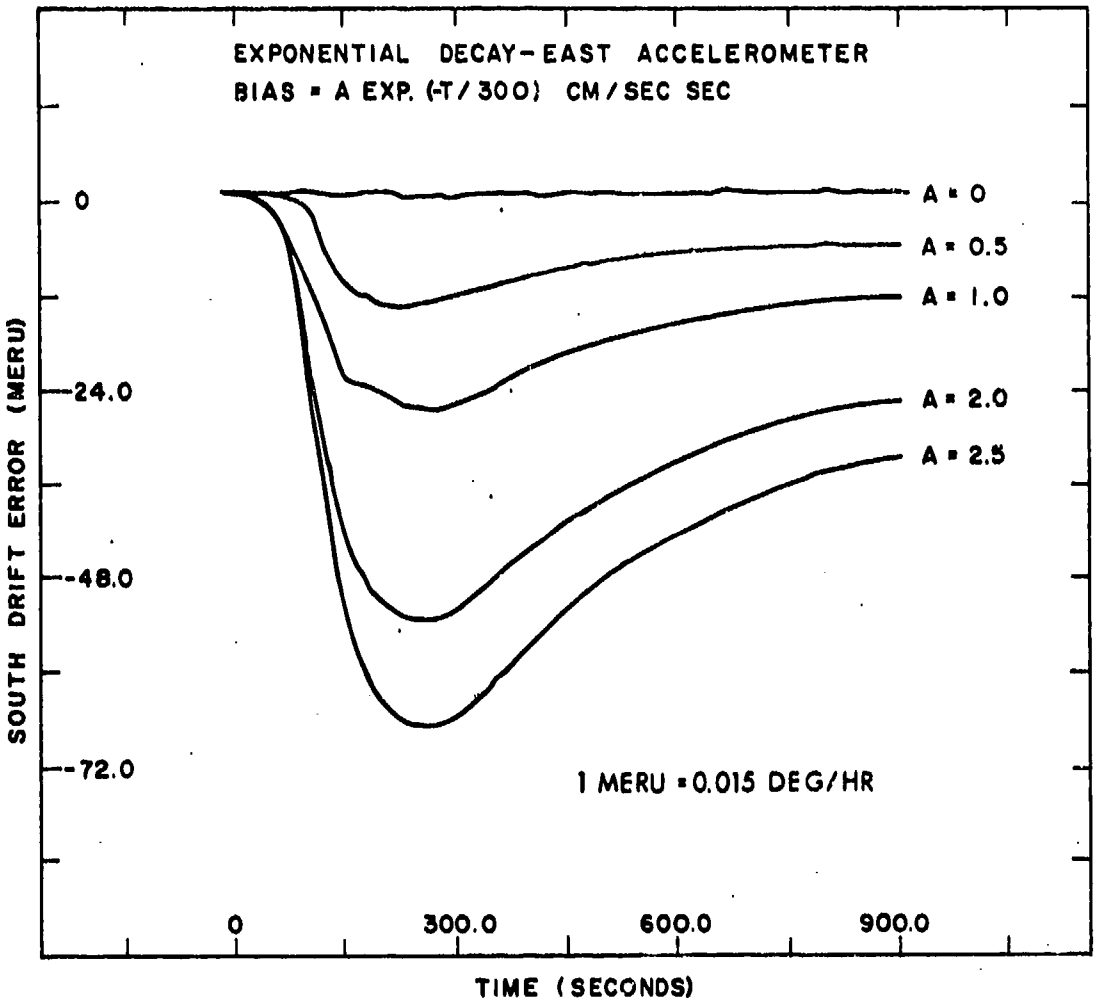


Fig. 28 South drift error for changing east accelerometer bias

CHAPTER 11 - APPLICATION OF KALMAN FILTERING THEORY TO AUGMENTED INERTIAL NAVIGATION SYSTEMS

by

J.E. Huddle

Litton Systems, Inc.
Guidance and Control Systems Division
8500 Canoga Avenue, Woodland Hills, California, USA

CHAPTER 11 - APPLICATION OF KALMAN FILTERING THEORY TO AUGMENTED INERTIAL NAVIGATION SYSTEMS

J.R.Huddle

1. INTRODUCTION

In the past two decades different navigation sensors have been combined with an inertial platform to obtain various multi-sensor systems for performing the navigation function. The main objective in integrating sensors into a system has been to increase the navigational accuracy obtainable when information from independent sensors is properly mixed. Another objective has been to increase the probability of success of a given mission that can be achieved through the existence of independent failure modes in a system. These modes are realized by different combinations of sensors, each capable of providing useful information during a mission. Reduction in total system cost and weight can also result in such designs being obtained implicitly through a relaxation in individual sensor performance requirements. The reduction in sensor requirements in multi-sensor systems usually means a decrease in design complexity with attendant reductions in sensor size, weight and cost, increased reliability and simpler maintenance.

Before the development of Kalman filtering theory, it was realized in the navigation field that integrated system operation offered the opportunities already outlined for obtaining improved navigation system characteristics. However, the analysis and optimization of the error behavior in these systems is a complex task, since the models describing these errors are stochastic, time-variant and have high dimension. The designs for the control of error in multi-sensor systems were obtained mainly by using the classical frequency domain and root locus techniques coordinated with trial and error design experimentation on analog or digital computers. The advent of the Kalman theory and the availability of more sophisticated digital computing hardware and software for investigating complex system error behavior has revolutionized the design of the error control mechanism for many multi-sensor navigation system applications. The impetus for this change results from the existence of the least squares controller solution for linear, stochastic systems, as provided by Kalman, and the means by which detailed behavior of complex systems can be examined, as provided through the use of high speed, large capacity digital computers. The significance of these points for practical design problems will become more evident later.

This chapter begins by describing the navigation function and the requirements which are usually imposed on navigation systems. The major classifications within which most navigation sensors can be grouped are defined. To illustrate how these different navigation sensors have been employed in conventional designs, the theory of operation of three fundamental augmented inertial navigation systems is presented. With this discussion, a perspective is obtained for considering a Kalman mechanization for these systems in a later section.

Since the computational allocation for error control in an actual system is limited, constraints must be imposed on the complexity and configuration of any Kalman design proposed. Some of the practical problems affecting the configuration of the error controller in actual systems are discussed, and the basic operations which are implemented in multi-sensor navigation system applications are defined. The question of the complexity of the Kalman design is also considered, and a procedure which has been used in error controller synthesis is described. This design approach is then used to obtain a Kalman error control mechanization which is suitable for applying each of the three augmented inertial navigation systems previously examined to a particular system mission.

The chapter concludes by illustrating the system performance for this mission, as obtained with the derived Kalman mechanization, and discussing the unique features of the three navigation systems.

2. NAVIGATION AND THE SENSORS EMPLOYED IN NAVIGATION SYSTEMS

2.1 General

The function of a navigation system is to determine the position of a vehicle such that it can be directed to other positions defined in the selected frame of reference. Terrestrial navigation is usually performed with respect to an orthogonal coordinate system affixed to the earth which has one axis coincident with the terrestrial pole and two axes in the equatorial plane. Vehicle position in this frame of reference is denoted in terms of the two angular measures of latitude and longitude and the linear measure of altitude. As seen in Figure 1, longitude is measured eastward or westward about the polar axes from the zero longitudinal meridian which passes through Greenwich, England. Latitude ϕ is measured north or south from the equatorial plane about an axis orthogonal to the meridional plane in which the vehicle lies. Altitude h is measured positively or negatively with respect to sea level along the local vertical at the latitude and longitude of the vehicle.

In addition to vehicle position, the navigation system is often required to provide velocity of the vehicle with respect to the earth. The components of the velocity vector are usually desired along the orthogonal axes of a coordinate system located at the vehicle position. This reference is called the local geographic coordinate system and has one axis coincident with the local vertical and two axes in the local level or horizon plane, one along the local north line and the other directed eastward.

Another type of data often required from a navigation system is vehicle attitude with respect to the local horizon plane and north line. In sophisticated avionics systems, attitude information is required to accomplish automatic aircraft guidance and control for weapon delivery, cruise to pre-selected destinations, and landing and take-off. Attitude information is available when the navigation system incorporates a stabilized reference such as an inertial platform.

To satisfy some or all of the navigation system requirements discussed, different types of sensors are employed. Most of these navigation sensors can be placed in one of three categories based on the types of data that they provide. These three classes encompass navigation sensors which provide inertial, speed, or position information in some form. Some of these sensors (e.g., inertial, LORAN) are capable of achieving some or all of the navigation system requirements autonomously. Others (e.g., Doppler radar, range measuring equipments) require information from other navigation sensors, primarily an inertial platform to accomplish the navigation function. Each of the three classifications is now described to introduce ensuing material on the principles of operation of some basic augmented inertial navigation systems.

2.2 Inertial Sensors

The basic element of an inertial sensor is the platform. The platform consists of a mechanical structure on which the inertial instruments are mounted. Up to three accelerometers are incorporated in the platform for measuring orthogonal specific force magnitudes. Gyroscopes are incorporated for the attitude measurement or stabilization of the accelerometer triad to a specified orientation.

With knowledge of the local mass-attraction vector, an inertial sensor can be used to determine the vector acceleration of the carrying-vehicle with respect to the inertial space. When this information is provided to a computer, subsequent integrations can be performed to yield vehicle velocity and position. In addition to this basic capability to provide sufficient information for autonomous navigation, inertial sensors are used as attitude references providing the angular orientation of the vehicle with respect to the local horizon plane and north line. This attitude data is used in the operation of other navigation sensors, as will be seen.

2.3 Speed Sensors

These sensors can provide speed of the carrying-vehicle along different directions. The interpretation of this information as a vector, however, requires the use of an attitude reference system. Two types of speed component measuring instruments exist. The first type determines speed with respect to the medium in which the vehicle travels, as achieved by air-speed indicators or ships' logs. The second type employs the Doppler frequency shift phenomenon which is exemplified in Doppler sonar or radar. These Doppler instruments provide speed with respect to a reflecting surface, usually the earth's surface. The attitude data provided to stabilize this information can yield velocity components along the local geographic coordinate axes. Subsequent integration of the velocity component data in a computer can yield the present latitude, longitude and altitude of the vehicle.

2.4 Position or Angular Sensors

A number of classifications exist for sensors which provide information useful in determining position. One type includes tracking systems, such as astro-trackers and optical sights, which provide the angular orientation of a point of known inertial or geographic location with respect to a set of coordinate axes of known orientation in the vehicle. This type of sensor requires an attitude reference and computer to obtain useful navigation data.

Another basic classification for position sensors includes those which are based on timing mechanisms. Included are systems which measure:

(a) Time between the transmission and receipt of a pulse of energy from objects of known location. This is done with ranging radar, laser range-finders and distance measuring equipments.

(b) Difference between time of arrival of pulses from different transmitters with known location. This is done with existing LORAN, and omega and proposed satellite navigation systems.

(c) Integrated Doppler shift data to provide the change in range to a reflecting or transmitting target.

To obtain precision in these systems, high frequencies are employed. Multiple frequencies can be used to resolve ambiguities. In systems, these sensors are usually employed with other types of navigation sensors and a digital computer. With an inertial sensor, for example, dead-reckoning of vehicle position can be performed between the times at which position information is available, with the advantage that navigation data is available continuously.

Angular information can be provided with radio equipment having directional radiation characteristics. Examples of such equipment are found in instrument landing systems at airports, which provide glide-slope and

bearing angle indications to landing aircraft. Another example is provided by military TACAN and civilian omnirange stations, which provide indications of vehicle bearing with respect to the respective transmitters. The most common vehicle heading indicator available is the magnetic compass, which provides heading information relative to the magnetic poles of the earth.

3. CONVENTIONAL AUGMENTED INERTIAL NAVIGATION SYSTEMS

3.1 Some Al

This section reviews the manner in which different sensors have been employed in more conventional augmented inertial navigation systems. The discussion will permit a better understanding of the advantages gained through the use of system error controllers based on Kalman theory, which are discussed later. To provide this introduction, the theory of operation for three augmented inertial navigation systems is discussed. The three navigation systems that have been selected are commonly referred to as Doppler-inertial, LORAN-inertial and astro-inertial.

An inertial sensor has been selected for each of these systems since, in applications of any complexity, the navigation system will include an inertial sensor. This occurs because the inertial sensor provides the most dynamic form of information possible through a direct measurement of acceleration and attitude often necessary for the operation of other navigation sensors. The three navigation systems selected illustrate distinct methods of system operation through the character of the sensor augmenting the inertial sensor. In each of the systems, a computer must be assumed to perform the required processing of measurement data to obtain the navigation variable outputs. Since the inertial sensor plays a basic role in each system, the theory of its autonomous operation will be discussed first.

3.2 Autonomous Inertial Navigation System Operation

Before an inertial platform can provide outputs useful for navigation, the accelerometer axes must be made coincident with the orientation assumed in the navigation computer. In most terrestrial navigation applications a Schuler-tuned platform is employed. In such an inertial system, two orthogonal accelerometers are initially aligned and then maintained in the local-level plane. A third orthogonal accelerometer, if used, will then be coincident with the local gravity vector. The maintenance of any fixed platform orientation with respect to the earth requires rotation of the platform with respect to inertial space to compensate for the rotation of the earth with respect to inertial space and any relative angular rate of change of the geographic position of the carrying-vehicle. The azimuth of the platform is defined as the angular rotation in the level plane of one of the level accelerometers from the local north line. We will assume that one of the level accelerometers is always pointed northward, resulting in what is known as the north-slaved platform mechanization.

Once the platform is aligned, the accelerometer orientation assumed in the navigation computer allows the actual specific force measurements to be processed to yield useful navigation information. For terrestrial navigation, vehicle velocity and position with respect to the coordinate frame affixed to the earth is required. To derive this information from the accelerometer inputs, acceleration with respect to inertial space is first obtained by subtracting the effect of known local gravity from the specific force measurements. The derivative of vehicle velocity with respect to the earth, taken with respect to the accelerometer coordinate system, is then obtained by removing computed Coriolis acceleration components from these inertial accelerations. The Coriolis acceleration components are dependent on the velocity and rotation rate of the accelerometer triad with respect to the earth and the rotation rate of the earth with respect to inertial space. The resulting three derivatives are

$$\left.\begin{aligned}
\frac{d}{dt}\,[V_{Ec}] &= a_{Ep} + [(\rho + 2\Omega)_{Ec}V_{dc} - (\rho + 2\Omega)_{Nc}V_{Nc}]\,, \\
\frac{d}{dt}\,[V_{Nc}] &= a_{Np} + [(\rho)_{Ec}V_{Ic} - (\rho + 2\Omega)_{Ec}V_{Ec}]\,, \\
\frac{d}{dt}\,[V_{Ic}] &= a_{Ip} + [(\rho + 2\Omega)_{Nc}V_{Ec} - (\rho)_{Ec}V_{Nc}] - g_{oc}\,,
\end{aligned}\right\} \qquad (3.1)$$

where

$(a_{E,N,I})_p$ are specific force magnitudes measured by the east, north and vertical accelerometers, respectively;

$(V_{E,N,I})_c$ are computed components of vehicle velocity with respect to the earth along the local east, north and vertical geographic axes, respectively;

$(\rho)_{E,N,I})_c$ are the instantaneous computed components of angular rate of the local geographic axes with respect to the earth about the east, north and vertical axes, due to vehicle velocity with respect to the earth;

$(\Omega)_{N,I})_c$ are the computed components of the angular rate of the earth with respect to inertial space about the local north and vertical axes (note $\Omega_{Ec} = 0$);

g_{Dc} is the computed magnitude of the local plumb-bob gravity vector along the local vertical.

The bracketed terms on the right side of Equations (3.1) are the Coriolis acceleration components. Integration of these derivatives, which are with respect to the local geographic coordinate system, yields the computed values of vehicle velocity with respect to the earth expressed in local geographic coordinates.

The computed components of relative angular rate of the local geographic coordinates when added to computed local earth-rate components are employed to rotate the gyros in inertial space to maintain the local-level, north-slaved orientation of the platform. The relative rates are obtained from the computed velocity components as

$$\left.\begin{aligned} \rho_{Ec} &= -V_{Nc}/R_{Mc} , \\ \rho_{Nc} &= V_{Ec}/R_{Nc} , \\ \rho_{Zc} &= \rho_{Nc} \tan \phi_c . \end{aligned}\right\} \tag{3.2}$$

where R_{Mc} is the computed meridional radius of the earth and R_{Nc} is the computed normal radius of the earth. The computed local earth rate components are obtained by projecting the earth rate about the pole onto the local north and vertical axes:

$$\left.\begin{aligned} \Omega_{Nc} &= \Omega \cos \phi_c \\ \Omega_{Zc} &= \Omega \sin \phi_c . \end{aligned}\right\} \tag{3.3}$$

The computed angular rate of change of the platform orientation with respect to inertial space is then

$$\left.\begin{aligned} \omega_{Ec} &= \rho_{Ec} \\ \omega_{Nc} &= (\rho + \Omega)_{Nc} \\ \omega_{Zc} &= (\rho + \Omega)_{Zc} . \end{aligned}\right\} \tag{3.4}$$

The actual angular rate of rotation of the platform with respect to inertial space differs from Equations (3.4) by a stochastic function due to errors in the platform gyros:

$$\left.\begin{aligned} \omega_{Ep} &= \omega_{Ec} + b_E , \\ \omega_{Np} &= \omega_{Nc} + b_N , \\ \omega_{Zp} &= \omega_{Zc} + b_Z , \end{aligned}\right\} \tag{3.5}$$

where

$(\omega_{E,N,Z})_c$ is the computed angular rate of change of the platform with respect to inertial space expressed about the local geographic axes;

$(\omega_{E,N,Z})_p$ is the actual angular rate of change of the platform with respect to inertial space expressed about the local geographic axes;

$(b_{E,N,Z})$ are stochastic functions representing generalized gyro drift rate with respect to inertial space expressed about the local geographic axes.

The final step in determining the navigation variables is the computation of vehicle position in the earth coordinate frame. Latitude and longitude are computed as integrals of the local relative angular rate vector components about the local east axis and terrestrial pole, respectively. Altitude is obtained as the integral of vehicle velocity with respect to the earth along the local vertical axis:

$$\left.\begin{aligned} \frac{d}{dt}|\phi_c| &= -\rho_{Ec} \\ \frac{d}{dt}|\lambda_c| &= \rho_{Nc} \sec \phi_c , \\ \frac{d}{dt}|h_c| &= V_{Zc} . \end{aligned}\right\} \tag{3.6}$$

where ϕ_c is the computed vehicle latitude, λ_c is the computed vehicle longitude and h_c is the computed vehicle altitude.

A block diagram illustrating schematically the mechanization of a Schuler-tuned, north-slaved, free-inertial navigation system defined by Equations (3.1)-(3.6) is shown in Figure 2.

The advantages of employing an inertial sensor derive from the availability of vehicle attitude information from the platform gimbals and dynamically-exact changes in vehicle velocity obtained from the accelerometer measurements. These types of information are usually necessary for the operation of other avionics subsystems in the vehicle. The main drawback in employing an inertial system for autonomous navigation is that the errors in platform attitude and computed vehicle position and velocity have oscillatory behavior. Stochastic noise sources will therefore cause unbounded divergence in these errors as the operating time increases. To retain the valuable features of the inertial system operation while counteracting the divergent and oscillatory error characteristic, other sensors have been incorporated in navigation systems to obtain bounds on the inertial system errors. These auxiliary sensors measure some component of position, the rate of change of position or angular orientation. Since the error in this measured variable is generally bounded, at least a bound is obtained for the relevant inertial system error in the integrated navigation system.

An appreciation of the complexity of the augmented inertial navigation system design problem is gained through an examination of the equations describing the errors in the inertial system outputs. A description which is adequate for most applications is obtained with a set of nine time-variant, linear differential equations with stochastic inputs. These equations can be obtained by differencing the sets of Equations (3.1), (3.5) and (3.6) from these same equations expressed in terms of only true values of the variables involved, by then introducing the effects of errors in the accelerometers and gyros, and then by ignoring second-order terms in the error variables.

The variables used to represent error in the inertial navigation system are defined as follows, where the true value of the variables in question are double subscripted:

$$\delta\lambda_c \triangleq [\lambda_c - \lambda]$$

$$\delta\phi_c \triangleq [\phi_c - \phi]$$

$$\delta h_c \triangleq [h_c - h]$$

are the errors in the computed vehicle longitude, latitude and altitude, respectively;

$$\delta v_{Ec} \triangleq [v_{Ec} - v_E]$$

$$\delta v_{Nc} \triangleq [v_{Nc} - v_N]$$

$$\delta v_{Zc} \triangleq [v_{Zc} - v_Z]$$

are the errors in the computed vehicle east, north and vertical velocity, respectively. g_{zc} is the error in computed apparent gravity magnitude.

From these definition the following errors in the computer rate of rotation of the local geographic axes with respect to the earth and computed earth rate about the north and vertical axes are derived to first order:

$$\delta\rho_{Ec} \triangleq [\rho_{Ec} - \rho_E] = \frac{-\delta v_{Nc}}{R_N}$$

$$\delta\rho_{Nc} \triangleq [\rho_{Nc} - \rho_N] = \frac{\delta v_{Ec}}{N}$$

$$\delta\rho_{Zc} \triangleq [\rho_{Zc} - \rho_Z] = [\delta\rho_{Nc} \tan\phi + \rho_N \sec^2\phi\, \delta\phi_c]$$

$$\delta\Omega_{Nc} \triangleq [\Omega_{Nc} - \Omega_N] = -\Omega_Z \delta\phi_c$$

$$\delta\Omega_{Zc} \triangleq [\Omega_{Zc} - \Omega_Z] = \Omega_N \delta\phi_c .$$

The first set of differential equations given below describes the angular deviation of the accelerometer coordinate axes from the local geographic coordinate axes. This platform attitude error is obtained by integrating the difference in the instantaneous rate of rotation of the actual platform and local geographic coordinate axes with respect to inertial space. Since it is convenient to perform this integration in the local geographic frame, these errors are defined by the three differential equations:

$$\left.\begin{aligned}
\frac{d}{dt}[\phi_E] &= (\omega_{Ec} - \omega_E) + b_E + (\Phi_{y'u_2} - \Phi_{y'u_N}) \\
\frac{d}{dt}[\phi_N] &= (\omega_{Nc} - \omega_N) + b_N + (\Phi_{y'u_E} - \Phi_{y'u_Z}) \\
\frac{d}{dt}[\phi_Z] &= (\omega_{Zc} - \omega_Z) + b_Z + (\Phi_{y'u_N} - \Phi_{y'u_E}) .
\end{aligned}\right\} \qquad (3.7)$$

The second set of differential equations describes the propagation of error in the computed values of east, north and vertical velocity:

$$\left.\begin{aligned}
\frac{d}{dt}[\delta V_{E_c}] &= [\Phi_E g_N - \Phi_N g_E] + \nabla_E - [\delta\rho + 2\delta\Omega]_{N_c} V_E - [\rho + 2\Omega]_N \delta V_{E_c} + \\
&\quad + [\delta\rho + 2\delta\Omega]_{E_c} V_N + [\rho + 2\Omega]_E \delta V_{N_c} , \\
\frac{d}{dt}[\delta V_{N_c}] &= [\Phi_E g_E - \Phi_E g_E] + \nabla_N - [\delta\rho + 2\delta\Omega]_{E_c} V_E - [\rho + 2\Omega]_E \delta V_{E_c} + \\
&\quad + \delta\rho_{E_c} V_E + \rho_E \delta V_{E_c} , \\
\frac{d}{dt}[\delta V_{E_c}] &= [\Phi_N g_E - \Phi_E g_N] + \nabla_E - \delta\rho_{E_c} V_{N_c} - \rho_E \delta V_{N_c} - g_{E_{ac}} + \\
&\quad + [\delta\rho + 2\delta\Omega]_{N_c} V_E + [\rho + 2\Omega]_N \delta V_{E_c} ,
\end{aligned}\right\} \qquad (3.8)$$

where $\nabla_{E,N,Z}$ are stochastic functions representing generalized measurement errors in the east, north and vertical accelerometers.

The final set of differential equations describes the propagation of error in the computed values of vehicle longitude, latitude and altitude:

$$\left.\begin{aligned}
\frac{d}{dt}[\delta\lambda_c] &= [\delta\rho_{N_c} + \rho_N \tan\Phi\, \delta\Phi_c] \sec\Phi \\
\frac{d}{dt}[\delta\Phi_c] &= -\delta\rho_{E_c} \\
\frac{d}{dt}[\delta h_c] &= \delta V_{E_c} .
\end{aligned}\right\} \qquad (3.9)$$

A block diagram illustrating the linear error model of the level channels of a Schuler-tuned, north-slaved, free-inertial navigation system defined by Equations (3.7)-(3.9) is shown in Figure 3.

Due to the stochastic, time-variant, multi-variable form of these error equations, the application of classical frequency domain and root-locus technique is not effective for error controller design unless assumptions can be made which simplify the equations. In most applications, such assumptions result in constraints on the manner in which the system is operated. For example, platform alignment mechanizations have been designed with classical approaches, assuming time-invariance and minimal coupling between error variables. Such assumptions are only valid if the vehicle undergoes no acceleration, which constrains system operation significantly. Since Kalman theory accommodates the time-variant character of navigation system behavior, only operational conditions which require a non-linear description of error behavior need be avoided when this approach is employed. Most design efforts using Kalman concepts for error control deal with the minimization of mechanization complexity that can be obtained for the given system performance requirements. Digital computers are usually required to deal with this synthesis problem effectively.

The introduction of Kalman theory has helped to change the viewpoint of analysts in augmented inertial navigation system design from one in which various other sensors simply aid the inertial sensor in performing the navigation function to one where each sensor is regarded as a source of information available to the navigation computer to be used in deriving values of the required navigation variables. With this new viewpoint, the system outputs at any instant of time, regardless of vehicle flight dynamics or intermittent availability of data from any sensor, can be optimal in the sense that all information has been fully utilized. This complete absorption of sensor data is realized through the use of correlation information contained in the system error covariance matrix. The use of this matrix as a basis for measurement data mixing in Kalman filtering constitutes the unique and novel aspect of this design approach.

When Kalman theory is applied to terrestrial navigation systems, the covariance matrix is propagated in real-time in the system computer, due to the unpredictable nature of vehicle maneuvers and navigation sensor data availability. As a consequence, the instantaneous statistical state of the system, reflected by the covariance matrix, is always available to enable optimal utilization of new sensor measurements. In following sections we will describe the manner in which Kalman filter concepts are employed in error controller design for augmented inertial navigation systems. For the present we describe the manner in which different navigation sensors have been used to control inertial navigation system errors with more conventional system mechanizations.

3.3 Doppler-Inertial Navigation System Operation

In a Doppler-inertial navigation system, the sensor which augments the inertial sensor is the Doppler radar. The Doppler radar measures the frequency shift between a transmitted and reflected radio wave. This Doppler shift is proportional to the speed of the vehicle with respect to the reflecting surface along the direction in which the beam of energy is transmitted. Doppler radars have been developed which can provide three independent components of vehicle speed. These components are, however, expressed in a coordinate system related to the

Doppler antenna, whose orientation is not known with respect to the earth. Consequently, a Doppler radar cannot perform the navigation function autonomously, but requires that the Doppler antenna be stabilized mechanically or computationally using attitude information provided by an inertial platform or some other orientation reference. Once the stabilization of the Doppler speed measurements is achieved, vehicle velocity with respect to earth along the local east, north, and vertical axes is known, allowing changes in vehicle position with respect to the earth to be obtained from integration in the system computer.

The errors in the determination of vehicle speed in antenna coordinates from the Doppler frequency shift are due to various equipment calibration errors (e.g., antenna temperature shifts, beam misalignments, etc.) and the statistical character of the beam reflecting surface over which the vehicle travels. Since the beam has finite width, for a given vehicle speed a continuous spectrum of Doppler frequency shift information is extracted from the reflected beam. The centroid of this return is usually tracked and a scale assigned based on nominal operating conditions. If the character of the reflecting terrain changes markedly, the assigned scale factor will be in error. Most of this type of difficulty in using Doppler radar occurs over water where the surface smoothness distorts the return spectrum. Such problems have been solved to some extent by beam lobing. Over very smooth surfaces, however, the Doppler radar output cannot be used, as insufficient return energy is received. Autonomous navigation with Doppler radar is also degraded in over-water operation when surface currents exist. Without an independent reference, navigation with respect to the moving surface rather than the earth takes place. The important characteristic of the Doppler errors is that they can be expressed as bounded fixed fractions of the vehicle speed components. Consequently, the errors in the Doppler speed components, expressed along the platform coordinate axes after data stabilization, are bounded. Proper blending of these data with the velocity components obtained from the inertial system will result in a system which has error components along these axes which are less than the bounds.

In a conventional coupling of the Doppler and inertial data[11], the velocity components obtained from each sensor, or other suitable independent functions of these variables, are differenced from each other. For sufficiently small errors, these differences can be expressed as a linear combination of various errors associated with the Doppler and inertial sensors. These signals can then be fed back to correct or compensate the different sources of error in the navigation system. The eventual desired result is the reduction of error in all the navigation system outputs. A block diagram illustrating a conventional ground speed and drift angle damped north-slaved inertial navigation system mechanization is shown in Figure 4. It can be seen in this figure that the differences between predicted and Doppler measured ground speed and drift angle are fed back through gains to correct the computed north and east velocities and alter the computed platform spatial rate components.

3.4 LORAN-Inertial Navigation System Operation

In a LORAN-inertial navigation system the data used to augment the inertial system in performing the navigation function is derived from a LORAN time-difference receiver. LORAN employed as an autonomous navigation system provides position information only. The position is derived from measurements of the difference between the local time of receipt of pulses transmitted from stations at known position on the surface of the earth. At least three stations are required if a complete position fix is to be obtained from LORAN. One of the stations, called the master, transmits the first pulse. This pulse when received by the two other stations, which are called slaves, is then retransmitted. Upon receiving the master pulse, the LORAN radio receiver in the vehicle starts a precise clock which measures the times until the pulses are received from the slave stations. The constant difference of the distances from two fixed points defines a hyperboloid in space with the station positions as the foci. Consequently, the time-difference measurement obtained from a master-slave station pair will locate the vehicle on such a surface. Assuming the vehicle is near the surface of the earth, a line-of-position is defined. When a second time-difference measurement creates a line-of-position which intersects with the other, the position of the vehicle is defined. The reason that time-difference is measured in LORAN, as opposed to the time it takes a pulse to travel from a station to the vehicle, is that there is then no requirement for precise synchronization of the vehicle and station clocks with an absolute time standard. If precise synchronization were achieved, the derivation of range to two stations, using an absolute time standard and a time-of-receipt receiver on the vehicle, would suffice to obtain a complete position fix on the earth's surface.

Errors exist in the determination of true time-difference at the vehicle for a variety of reasons, including synchronization error at the slave stations in pulse retransmission, ambient receiver and atmospheric noise and local radio interference near the vehicle. Another important source of error is that associated with the lack of precise knowledge of the speed of propagation of electromagnetic energy over the earth's surface. This is caused by stochastic variations in the propagation path definition. The result is an actual time-difference grid which is distorted relative to that predicted using geodetic data only.

The important characteristic of these errors in the use of LORAN in aiding the inertial system in navigation is that they are bounded. Hence the error in position obtained from the time-difference measurements in areas of defined LORAN coverage will serve as an upper bound on the position error for the navigation system. Proper mixing of position information from the LORAN and inertial sensors can result in less error in all system outputs than could be achieved with either sensor operating autonomously.

In a conventional coupling of information from the LORAN and inertial sensors, time-difference predictions derived from the inertial system, or predicted geographic position information derived from the LORAN measurements, are subtracted from the appropriate variable from the other sensor. For sufficiently small errors, this difference can be expressed as a linear combination of errors associated with each of the sensors. Consequently, as in the Doppler-inertial system, the same basis exists for the design of a linear feedback controller

to reduce the errors in the outputs of the system. This basis has been exploited in the past, where time-invariant error control mechanism[9] have been designed. These mechanizations are usually far from optimal owing to the aforementioned limited applicability of the time-invariance assumption for the inertial error model, as well as to the varying quality of time-difference data as the vehicle moves through a LORAN coverage area. The use of Kalman filtering concepts in the design of such a navigation system offers great benefits, as the time-variant behavior of the error can be accounted for automatically on a real-time basis.

3.5 Astro-Inertial Navigation System Operation

The final navigation system to be discussed is one in which a star-tracker is employed to measure, at the vehicle, the angular orientation of astronomical bodies with known celestial position. Like the Doppler, an astro-tracker cannot be employed autonomously to determine vehicle position. An attitude reference with known orientation with respect to an earth-fixed coordinate system must be provided to allow an interpretation of the sighting angles. Normally, the optical axis of the tracker is oriented by slewing it through bearing and elevation angles, respectively, from an initial coincidence with the local north line. An astro-tracker usually incorporates a detection and servo-mechanism system which drives the telescope elevation and bearing gimbals until a detected light source resides at the center of the field-of-view. When this is achieved, measurements of the bearing angle from the local north line in the local-level plane, and the elevation angle out of the local-level plane, can be read directly from telescope gimbals.

In an automated tracker system, the optical axis is first directed to a point in the sky where, based on the computed vehicle position and known celestial position of a star, the star would be expected to be found. Because of errors in the computed vehicle position and the attitude reference used, the star will deviate from the center of the field-of-view. The errors in the initial pointing angle can be determined by centering the star and measuring the changes in the telescope's angular orientation. If the angular errors in vehicle position and platform attitude are sufficiently small, the pointing angle errors can be expressed as linear combinations of these errors. For example, elevation and bearing angle errors could be linear functions of the errors in computed vehicle latitude and longitude, platform azimuth error, and the platform tilt errors about the local north and east axes. These causes of pointing angle error can be expressed about axes of an orthogonal coordinate system which has one axis coincident with the optical axis of the tracker. Since the initial telescope orientation error can be expressed about axes orthogonal to the line-of-sight axis, the angular error component along the telescope axis is unobservable. The errors in the elevation and bearing angle measurements due to the tracker system itself are bounded. These errors result from mechanical misalignment of the telescope, electrical bias and noise in the nulling control loops, and environmental effects, such as the index of refraction of the atmosphere.

In a conventional astro-inertial system, where the attitude reference is of high quality, it is usually assumed that no error exists in the platform level. Corrections are then made only to the computed vehicle position and the azimuth orientation of the platform. Usually in a star-tracker system, sequences of stars are shot to determine vehicle position. After each shot, residual error exists about the most recent line-of-sight, unless stars with orthogonal lines-of-sight are available. Conventional mechanizations usually do not capitalize on the relative orientation of the lines-of-sight that exist in a star shot sequence due to the bookkeeping involved. By automatically accounting for the optimal use of all previous measurement data, Kalman theory offers a much more powerful approach for obtaining error separation for this navigation system.

4. PRACTICAL CONSIDERATIONS IN USING KALMAN THEORY IN NAVIGATION SYSTEM DESIGN

4.1 General

The successful application of any method to achieve real-time error control in multi-sensor navigation systems requires an awareness of the constraints imposed by the actual hardware available for the system mechanization. In particular, the sensitivity of system performance, based on theoretical assumptions, to such phenomena as finite measurement rates and computation error resulting from computation equipment and/or computation algorithm characteristics must be determined to establish the efficacy of any error control concept. In the following paragraphs, the implications of the characteristics of navigation computers and sensors on the implementation of an error controller based on Kalman linear correlation theory are outlined.

4.2 Digital Computation

In the past, conventional augmented inertial navigation systems have been mechanized using analog computing elements. Such an approach is adequate for very simple mechanizations where stringent requirements for computational accuracy do not exist. More recently, the majority of augmented inertial navigation systems have employed various types of digital computing elements (e.g., general purpose and digital differential analyzer) to accomplish the required computations. Some of the advantages encountered in the use of these digital computers are capability for increased computational precision and flexibility in implementing logical functions. The current trend toward increased processing rate and greater storage at reduced cost is an additional attraction. The disadvantages incurred in digital computation stem primarily from the finite rate at which logical and arithmetic operations can be performed. These finite rates generate requirements for the use of computational algorithms to accomplish mathematical operations. Computational algorithms can, however, achieve only imperfect realizations of the desired mathematical operations. The most familiar, and probably most important, of these operations in practice is that of integration. Generally, an error control mechanization for virtually any

augmented inertial navigation system based on Kalman linear correlation theory will require that a considerable amount of precise arithmetic computation and logical switching be implemented. Consequently, the use of general purpose digital computers is required in these applications.

4.3 Advantage of Kalman Mechanizations in Augmented Inertial Navigations Systems

The essential difference between Kalman and conventional navigation system error control methods results from the use of a covariance matrix, with the Kalman technique as a basis for computing the gain to be used for system correction when a measurement is made. The information contained in this matrix provides the means by which superior system performance can be achieved in comparison to that obtained with conventional approaches. Conventional error control generally involves the use of matrices of pre-stored gains, which are determined using *a priori* assumptions. The use of the different gain sets is dictated as the real-time operating conditions of the system are ascertained.

As already noted, an error model for a navigation system is represented by sets of time-variant equations. Since there is infinite variety of possible flight and sensor data profiles for a navigating vehicle, the actual realization of these equations can only be made in real-time. Therefore, the nature of navigation system error propagation requires that a considerable number of gain matrices be stored in the digital computer to realize a conventional mechanization which accounts to any extent for the many sequences of system operating conditions possible. Usually this is not done in conventional designs, with the result that restrictions on operational capability of the vehicle are imposed to accommodate the mechanization inadequacies. Such restrictions on the operational capability of a system are extremely undesirable and can be avoided by using more sophisticated Kalman error control techniques.

The Kalman error control mechanization avoids constraints on vehicle operation through real-time propagation and update of the error covariance matrix. The actual realization of arbitrary functions of time (e.g., acceleration and velocity components) present in the differential equations describing system error are used in the error covariance matrix propagation, resulting in a correct representation of the system error state. Also, the arbitrary availability and variation in quality of sensor measurement data during system operation is properly accounted for through real-time update of the error covariance matrix on system correction. Real-time availability of sensor data is arbitrary, as the sensors may be intermittently operable. For example, accurate tracking of LORAN time-differences can be interrupted in the presence of high strength interference. Sensor data quality also varies with time as the vehicle moves into regions affecting sensor operation. With LORAN, the signal to noise ratio associated with information sent from a station usually drops as the distance between the vehicle and the transmitter increases. The quality of Doppler data changes with the terrain over which the vehicle travels, the antenna attitude, and the vehicle speed and altitude, all of which vary with time. Astro-tracking capability depends on the effective brightness of available stars at the position of the tracker.

4.4 Sensor Measurement Processing

Different formulations of the Kalman filtering equations have been obtained. The initial derivation[1] assumed that the observation or measurement process was continuous and that each measurement was contaminated by an additive white noise. In practice neither of these assumptions can actually be realized.

In the first case, continuous measurements cannot be assumed when the equations are implemented on a digital computer, as the computer can achieve only a finite rate of processing and will, additionally, be required to perform functions other than that of error control. Consequently, in applications, only a finite rate of absorption of sensor data can be realized.

The requirement for additive white noise contamination of the measurements in the initial formulation also presents difficulties. In all cases the assumption that the measurement-contaminating noise has short correlation time with respect to the system time constants cannot be made. A sensor in the process of extracting information in the presence of short-term noise employs some filtering. The filtering associated with this detection process will have the effects of correlating and reducing the magnitude of the contaminating noise process and introducing dynamic error in the information being measured. Equipment biases are further examples of long-term errors present in measurements.

The practical aspects resulting from digital implementation and sensor measurement smoothing have usually led to the use of the augmented state formulation of the equations with discrete observations in applications[2]. In this approach, coupling between the individual sensor subsystems is realized only at time instants when the computer processes sensor data to make system corrections. Between these update instants, the covariance matrix describing modeled system errors is propagated, reflecting the non-interaction of errors in the individual sensor subsystems. At the update time, measurements are made and compared with predicted values thus obtaining observed errors which are weighted with the Kalman gain to provide corrections to the modeled system errors.

Since the original derivation of the Kalman filtering equations, an alternate formulation by Bryson and Johansen[3] has been obtained which does not require the assumption that measurements are contaminated by white noise. With this formulation, continuous absorption of sensor data contaminated with correlated noise is theoretically possible. The dimension of this filter is less than that of the augmented state form by the number of measurements contaminated by first-order correlated noise. The amount of computation required, however, is greater than that of the original formulation for filters of the same dimension. The Bryson-Johansen formulation was extended to the discrete case by Bryson and Henrikson[4]. This formulation is equivalent to the augmented state

formulation with the discrete measurements made at a finite rate. However, in the limit, the Bryson-Henrikson filter converges to the Bryson-Johansen filter, while the augmented state formulation is not defined. With continuous absorption of sensor data, possible for the case where only correlated noise exists in the sensor measurements, the augmented state formulation with discrete observations will be comparatively less optimal. Results obtained by Bryson, Johansen and Henrikson imply that, as the sensor sampling rate is increased, performance achieved with the augmented state formulation will tend to the theoretical optimum obtained by Bryson and Johansen, until computational instability is reached. This result has been demonstrated by Rischel[4].

If the frequencies associated with the difference of a sensor measurement and the predicted value of that measurement based on system computed navigation variables are long with respect to the measurement sampling frequency, then there will be little benefit in faster measurement sampling. In cruise navigation system applications, the sensor sampling intervals selected are seldom much shorter than the correlation time associated with the highest frequency noise. Shorter sampling periods have little utility except, perhaps, during the time when the inertial platform is being aligned. During this period, error deviations in system position and velocity will occur relatively more rapidly due to the large platform attitude errors. Therefore, a higher rate of system correction would be desirable for obtaining tighter bounds on these errors.

The augmenting sensors in augmented inertial navigation systems have time constants associated with their individual information processing procedures. If these time constants are long, significant dynamic error can result in the sensor output measurements. The dynamic inaccuracies are usually due to vehicle maneuvers. In most systems, acceleration information from the inertial platform can be used to compensate the dynamic error associated with the augmenting sensor. This has been demonstrated in conventional LORAN-inertial systems[9], where the rate of change of time-difference has been supplied to the LORAN receiver from the inertial system. Such a mechanization relieves the receiver from the task of tracking dynamic changes in time-difference, at least to the limits of the inertial errors, thereby committing it primarily to the task of separating transmitted information from noise. When this is done, long receiver time constants on the order of many seconds can be retained.

In some cases, the additional design complexity implied by such an approach may not be desired. What is done in these situations is to select time constants which achieve a balance between stochastic and dynamic error in the sensor processing design. In Doppler-inertial navigation systems, the time constant of the frequency shift tracker in the Doppler radar can be short, on the order of 0.1 second. The purpose of employing short-time constants in the detection process is to minimize the dynamic error in the variable being measured. When short-time constants are used in the sensor filtering networks, a system correction utilizing the Kalman filtering equations at comparable rates may not be achievable due to allocation of computer time to error control. When this occurs, methods must be devised to minimize the information loss resulting from the slow sampling rate.

The most popular technique for minimizing this information loss[10] requires a high rate of comparison of the sensor measurements and the predicted values of these measurements. These differences, taken at the desired sampling rate, can be inserted into a pre-filter for smoothing between system correction instants. The pre-filtering will reduce the random sensor noise in this error difference, thereby improving the relative information content at the system update time. Note that this scheme does not introduce dynamic error in the sensor measurement, as only the error difference information, not the sequence of whole-value measurements, is being filtered. Consequently, the advantage gained in minimization of dynamic error in the measurements through the use of short sensor filtering time constants is retained, while the information content of the observed error used for system correction is improved.

A first-order representation of this pre-filtering approach is indicated schematically in Figure 5. Here predicted values, $\hat{m}$, are differenced with the raw measurements, m. These differences are weighted with a filtering gain, k, and used to update a summing register. At the system update time, the smoothed difference, y, is weighted with the Kalman gain to obtain the system correction. The summing register is usually zeroed whenever its contents are used.

The principal characteristic to investigate before employing such a mechanization, is the degree of distortion imposed on information passing through the pre-filter. Unless the pre-filtering is modeled in the Kalman formulation, the actual error characteristics differ from the modeled error characteristics. In augmented inertial navigation system applications, serious distortion will usually only occur during platform alignment modes. For these cases, an analysis of the distortion of error signals due to platform attitude error due to passage through the pre-filter should be made. Workable levels of distortion can be obtained with an appropriate selection of sensor and pre-filter time constants and the system update interval.

In this section we have considered some of the practical problems which arise in actually implementing a given error controller in the navigation computer. In augmented inertial navigation systems, digital computers are required for the complex arithmetic and logical operations associated with a Kalman error controller. Real-time propagation of the system error covariance matrix is required to accommodate the real-time dynamic maneuvers of the vehicle and the possible intermittency and variation in quality of sensor data. Finally, the reasons for finite discrete sampling of sensor data for system correction were discussed and a method for minimizing information loss was noted. In the next section, an approach is presented which has been used to alleviate some of the computational burden imposed on the digital computer. This approach deals with the synthesis of the error controller rather than the techniques described for its implementation.

5. SYNTHESIS FOR LINEAR ERROR CONTROLLERS USING KALMAN THEORY

5.1 General

The control of error in any system requires a definition of the sources of error and their dynamic behavior. If the error model for the system is described by linear differential equations with random initial conditions and white-noise driving functions, it is a simple matter to directly apply Kalman's filtering theory for the control of these errors. If the model is non-linear, either a linearization must be made or a different control algorithm must first be employed until the errors are reduced sufficiently for linear behavior to be realized.

The linear differential equations describing navigation system error are forced by stochastic functions. These stochastic driving functions have various degrees of auto-correlation, ranging from biases to essentially white noise terms. To devise dynamic models for such noise errors, correlation functions are derived from empirical data obtained about the noise sources. Linear dynamic systems excited by white noise are then synthesized from an analysis of these describing correlation functions to represent these error sources in the total system error model. Thus for multi-sensor navigation systems, information which is available about errors usually results in the definition of an extensive linear error model. The application of Kalman theory employing such error models results in significant computer computation time and storage requirements. To reduce these computational requirements, various methods have been tried in order to find simpler error control schemes which still obtain the beneficial correlation property. One of the most successful of these methods[1] involves the study of the complex error model itself to determine what simplifications can be made to reduce the computational burden. In this approach, one employs to advantage the fact that certain error components and dynamic interactions tend to dominate the error behavior in the system outputs. This approach of simplifying the error model to reduce the computational requirements is important, because it enables the analyst to use his experience directly in determining the error controller design.

5.2 Simplified Linear Error Controller Synthesis and Evaluation

Since the error model plays the central role in the Kalman linear error controller, its simplification yields a natural method of reducing the controller complexity. Generally, the idea is to retain the more dominant state variables and their dynamic interaction to define a simpler error model. To be specific, three steps are involved in this design simplification. The first step is to reduce the dimension of the refined error model by deleting those components of the error state vector which are felt to contribute least to overall error in the usable system outputs. The second step is to eliminate dynamic inter-couplings between these retained state variables which contribute least to the definition of their behavior. These two steps amount to obtaining simplified differential equations describing the more dominant state variables of the original error model. Once these steps are accomplished, refinement of the proposed filter based on the simplified error model is made. This final step of refinement involves the adjustment of white noise observation or disturbance noise parameters. These parameters are selected so as to compensate to some extent for the simplifications made in the preceding steps.

To determine whether an error model simplification or parameter adjustment is appropriate, the effect on overall system performance must be ascertained. Since the systems considered in navigation applications are highly complex, little progress in making such determinations can be made in a simple way analytically. Consequently, digital computer programs which provide either simulation results or solutions to complex arrays of analytical equations must be employed in studying such systems. For general system studies, Monte Carlo simulation programs have been applied with success. Usually, however, in applying such programs a large number of runs must be made to obtain precise results. Precision is required in this synthesis problem, as the exact statistical effect on system performance due to a design change must be known for design decisions to be made.

For the case in which an adequate representation of the system can be obtained from a linear error model, the following approach[1] has proved useful in determining linear controller designs which are simpler than the optimal. This technique will be discussed for the design of a Kalman error controller based on a simpler error model of the system than the reference error model. The Kalman error controller impacts on the system error propagation at update times, which occur when a reliable comparison of data derived from any two independent sensors can be made and the system computer has sufficient time to perform all computations required for updating. Between the update instants, all system errors propagate freely from the initial values resulting after the most recent system update. Between the system correction times we assume that the actual system error behavior is described by the linear vector differential equation

$$\frac{d}{dt}[x(t)] = A(t)x(t) + \epsilon(t) , \qquad (5.1)$$

where

$x(t)$ is an n-dimensional vector describing all system errors which have initial values described by the covariance matrix:

$$\Sigma_0 \triangleq E[x(0)\rangle\langle x(0)] ;$$

$A(t)$ is an $n \times n$ matrix defining the linear dynamics of the errors in the refined or reference error model;

$\epsilon(t)$ is a zero-mean, white noise process, where

$$Q(t)\delta(t-\tau) \triangleq E[\epsilon(t)><\epsilon(\tau)] .$$

With this notation, a well-known equation[7] describing the propagation of the covariance matrix corresponding to the system error $x(t)$ is

$$\frac{d}{dt}[\Sigma(t)] = A(t)\Sigma(t) + \Sigma(t)A^T(t) + Q(t) . \tag{5.2}$$

Whenever an update is to be made, it is assumed that the comparison of data from different sensors results in a linear combination of the errors $x(t)$, plus an additive white noise component. Thus the error which is actually observed is defined, in terms of the variables denoted below,

$$y(t) \triangleq H(t)x(t) + \eta(t) , \tag{5.3}$$

where

$y(t)$ is the m-dimensional error difference which is actually observed;

$H(t)$ is an $m \times n$ matrix describing the linear inter-relationship of the n-error components of $x(t)$ entering into the observation error;

$\eta(t)$ is a zero-mean, white noise process, where

$$R(t)\delta(t-\tau) \triangleq E[\epsilon(t)><\epsilon(\tau)] .$$

When an update is executed, an $n \times m$ gain matrix, $K(t)$, is used to weight the vector, $y(t)$, to form a correction vector, $c(t)$. This is then introduced into the system to result in the following change in the current system error:

$$x^+(t) = x^-(t) - c(t) , \tag{5.4}$$

where

$x^{-,+}(t)$ is the current system error before and after the update is made, respectively;

$c(t) \triangleq K(t)y(t)$ is an n-dimensional correction vector;

$K(t)$ is the $n \times m$ gain matrix used to weight the observed error difference, $y(t)$.

The transition of the covariance matrix describing system error upon update is given by the well-known equation

$$\Sigma^+(t) = [I - K(t)H(t)]\Sigma^-(t)[I - K(t)H(t)]^T + K(t)R(t)K(t)^T , \tag{5.5}$$

where $\Sigma^{-,+}(t)$ is the covariance matrix before and after update, respectively.

For that case in which the optimal linear controller with respect to the error model is implemented, the optimal gain is given by the well-known equation[1]

$$K^*(t) = \Sigma^-(t)H(t)^T[H(t)\Sigma^-(t)H(t)^T + R(t)]^{-1} , \tag{5.6}$$

The use of the optimal gain $K^*(t)$, allows Equation (5.5) to be written in the simplified form

$$\Sigma^+(t) = [I - K^*(t)H(t)]\Sigma^-(t) . \tag{5.7}$$

Our interest here is not the use of this optimal mechanization, but rather one which is less complex, where the reduction in complexity results from a direct examination of the system error in Equation (5.1). With this approach, the analyst first examines the error vector $x(t)$, and deletes those components which he feels contribute in a relatively minor way to total error in the system application. Once this is done, a new error vector $\mathbb{X}(t)$, of reduced dimension $\bar{n}$, $m \leq \bar{n} \leq n$, is obtained. The analysis then examines the subset of elements of the error dynamics matrix $A(t)$, which describe the interaction of the error components in $\mathbb{X}(t)$. Elements of this abbreviated portion of the matrix $A(t)$ not felt to describe significant interaction of the components of $\mathbb{X}(t)$ are then eliminated. A new dynamic matrix $\bar{A}(t)$, of dimension $\bar{n} \times \bar{n}$, can then be formed of the retained elements of $A(t)$. The next step is to determine whether the disturbance white noise process, $\epsilon(t)$, might be modified to aid in describing the actual system error propagation. Usually an increase in the magnitudes of the retained white noise components is beneficial, in that additional uncertainty is artificially introduced into the simpler error model, compensating to some extent for the reduction in uncertainty that is associated with the elimination of error components in obtaining $\mathbb{X}(t)$ from $x(t)$. The final result is then a simpler error model described by

$$\frac{d}{dt}[\mathbb{X}] = \bar{A}(t)\mathbb{X}(t) + \bar{\epsilon}(t) , \tag{5.8}$$

where

$\mathbb{X}(t)$ is the $\bar{n}$-dimensional error vector ($m \leq \bar{n} \leq n$);

$\bar{A}(t)$ is the $\bar{n} \times \bar{n}$ dimensional dynamics matrix for the vector $\mathbb{X}(t)$;

$\bar{\epsilon}(t)$ is an $\bar{n}$-dimensional zero-mean, white noise process, where

$$Q(t)\delta(t-\tau) \triangleq E[\bar{x}(t)\bar{x}(\tau)] \ .$$

Clearly, the equation describing the propagation of the covariance matrix corresponding to Equation (5.8) is

$$\frac{d}{dt}[\bar{\Sigma}] = \bar{A}(t)\bar{\Sigma}(t) + \bar{\Sigma}(t)\bar{A}(t)^T + \bar{Q}(t) \ . \tag{5.9}$$

The description of the observation process resulting from the use of the error model in Equation (5.8) is necessarily changed. The simplified observation matrix $\bar{H}(t)$ is obtained by including only those elements in $H(t)$ involving the error components retained in $\bar{x}(t)$. Again to compensate the reduction in the uncertainty due to elimination of error vector components, the white noise process, $\eta(t)$, might be modified in obtaining a model of observed error. As was done in the case of the disturbance noise process, the magnitudes of the observation noise components are usually assumed larger to introduce, artificially, uncertainty into the description of the simpler model. The final result is a representation of the error actually observed in terms of the variables denoted below:

$$y(t) \triangleq \bar{H}(t)\bar{x}(t) + \eta(t) \ , \tag{5.10}$$

where

$y(t)$ is the actual m-dimensional observation vector;

$\bar{H}(t)$ is an $(m \times \bar{n})$ matrix specifying the relationship desired between components of the error vector $\bar{x}(t)$;

$\eta(t)$ is a zero-mean, m-dimensional white noise process describing additive noise in the observed errors $(m \leq \bar{n})$, where

$$E[\eta(t)\eta(\tau)] \triangleq \bar{R}(t)\delta(t-\tau) \ .$$

If Kalman's algorithm for the computation of the gain matrix is now employed:

$$\bar{K}(t)^* = \bar{\Sigma}(t)\bar{H}(t)^T[\bar{H}(t)\bar{\Sigma}(t)^-\bar{H}(t)^T + \bar{R}]^{-1} \ , \tag{5.11}$$

and used to weight the actual observed error, the correction obtained for the simplified error model at update is

$$\bar{x}^+(t) = \bar{x}^-(t) - \bar{v}(t) \ , \tag{5.12}$$

where $\bar{x}^{-,+}(t)$ is the modeled system error before and after update, respectively, and $\bar{v}(t) \triangleq \bar{K}^*(t)y(t)$ is the correction vector. The transition of the covariance matrix corresponding to the simplified error model is expressed from Equation (5.12) as

$$\bar{\Sigma}(t)^+ = [I - \bar{K}^*(t)\bar{H}(t)]\bar{\Sigma}(t)^- \ . \tag{5.13}$$

Now that the technical details have been presented, a review is in order. First, by assumption, the analyst has been given a complex linear model, defining error propagation in Equation (5.1), and the observed errors in Equation (5.3), for some system. The objective is to design an error controller which, while having the statistical correlation attributes of Kalman theory, has a lesser computational burden than that implied by the optimal mechanization. The optimal mechanization requires the use of Equations (5.2), (5.4), (5.6) and (5.7). The choice of approach involves the examination of the complex error model to determine what simplifications can be introduced into Equation (5.1) and their impact on Equation (5.3). The results are the simplified expressions (5.8) and (5.10). From these equations a less complex Kalman mechanization is implied, which would be implemented with Equations (5.9) and (5.11)-(5.13).

Clearly, the simpler error controller realizes poorer performance than that achieved with a mechanization based on the complete error model. However, the performance loss incurred by using the simpler mechanization might be insignificant in terms of the performance required of the system. The question of practical importance which remains is that of determing the performance loss. The method by which this question is answered is also the manner by which the analyst will test the validity of the various simplifications and alterations of the original error model that are made.

The method by which the answers to these questions are obtained is quite simple. First the ultimate performance obtainable is determined by propagating the covariance matrix $\Sigma(t)$, using Equation (5.2), with optimal updating through the use of Equations (5.6) and (5.7). The solution for $\Sigma(t)$, denoted $\Sigma_R(t)$, and, in particular, for the diagonal elements or error variances then serves as the reference against which performance obtained with any other mechanization can be compared. The second step involves the propagation of the two covariance matrices $\Sigma(t)$ and $\bar{\Sigma}(t)$. $\Sigma(t)$ is again propagated using Equation (5.2), while $\bar{\Sigma}(t)$ is propagated by Equation (5.9). At the update instants, the gain $\bar{K}^*(t)$, using Equation (5.11), is computed and used in Equation (5.13) to update $\bar{\Sigma}(t)$. To determine the effect of using $\bar{K}^*(t)$, which is sub-optimal compared to $K^*(t)$ obtained by Equation (5.6) on the actual system errors, an $(n \times m)$ matrix $K(t)$ is defined by augmenting $\bar{K}(t)$ suitably with zeros. Then $K(t)$ is employed in Equation (5.5), which describes the effect on the actual system error covariance matrix that results from the use of an arbitrary gain matrix. Comparison of the solution for $\Sigma(t)$, denoted as $\Sigma_E(t)$ and determined in the second step, with the reference solution $\Sigma_R(t)$, obtained in the first step, then indicates the degree of performance degradation that results through the use of the proposed computationally-less burdensome mechanization.

There is a significant increase in cost of the computer analysis with an increase in the dimension of the reference model in the synthesis approach described. Therefore, the analyst must often make certain approximations even in the reference error model. To obtain a suitable reference model of manageable dimension, multiple sources of error which enter the system at the same point can be merged and represented as single noise error. To maintain the integrity of the design results, the statistical characteristics of this lumped noise source can be selected to more than account for the multiple sources of error represented. For example, in the design of augmented inertial navigation systems, a finely-detailed structure can be developed to model the error in the accelerometer measurements or the stability of the gyro orientation. Upon examining this detailed error structure and considering the environment in which the instrument must operate, the analyst can represent these errors as single sources of first-order correlated noise. With these assumptions used to obtain a simple stochastic gyro drift rate or accelerometer bias shift, the synthesis technique described could then be used to determine precisely the relationships between proposed design innovations and the system performance.

Once a navigation system error controller is derived with this technique, some question might exist as to the validity of the assumptions made in using simple descriptions of the inertial instruments. This question can now be addressed in a meaningful manner by employing a Monte Carlo program to simulate the system using refined error models and the newly-derived error controller. Accepting the imprecision of results obtainable with a few trials using the Monte Carlo program, gross deviations from expected system performance can be adequately identified at a reasonable cost.

In summarizing this section, a straightforward application of Kalman theory to augmented inertial navigation system design using all information about the sensor error characteristics can result in a design which is impractical in terms of the requirements imposed on the system digital computer. A method of obtaining an error control mechanization which has the desirable correlation property of a Kalman filter was described in detail. This approach is direct in the sense that it deals with the character of the modeled error characteristics of the sensors. A digital computation procedure was outlined whereby this synthesis approach could be utilized in designing error controllers for complex systems. This procedure has been used in the avionics systems industry, particularly in obtaining the augmented inertial navigation system error control mechanization described in Section 7.

6. IMPLEMENTATION OF KALMAN ERROR CONTROLLERS IN AUGMENTED INERTIAL NAVIGATION SYSTEMS

6.1 General

In a previous section, practical considerations influencing the implementation of a Kalman error controller in an augmented inertial navigation system were discussed. We now consider the detailed operations which are implemented to mechanize the derived error controller in the system digital computer. When the synthesis of the Kalman controller is completed, the design error model for the system is described by the following linear differential equation, which is represented schematically in the block diagram of Figure 6:

$$\frac{d}{dt}[\bar{x}(t)] = \bar{A}(t)\bar{x}(t) + \bar{\varepsilon}(t) . \tag{6.1}$$

In augmented inertial navigation systems, measurements are made at discrete instants of time and compared with estimates of their values computed from the system navigation variables. For sufficiently small errors, the difference of the measurement and the predicted value can be represented as a linear combination of error states. This difference is called the observation error and is written

$$y(t_i) = \bar{H}(t_i)\bar{x}(t_i) + \bar{\eta}(t_i) . \tag{6.2}$$

When a measurement is made the observation error is multiplied by a Kalman-derived gain matrix to form a correction vector to update the system:

$$\bar{c}(t_i) = \bar{K}(t_i)y(t_i) . \tag{6.3}$$

The updated value of the modeled error state vector of the system is expressed by

$$\bar{x}(t_i)^+ = \bar{x}(t_i)^- - \bar{c}(t_i) . \tag{6.4}$$

When corrections are used to alter the values of the navigation variables within the system, servo-mechanism control of the modeled errors is realized. The implications of such a mechanization as opposed to the formal Kalman filter implementation are noted later.

A description of the error behavior of the system for the operations defined above can be represented schematically by the block diagram shown in Figure 7. This diagram is analogous to that for a servo-mechanism which attempts to null the system outputs by feeding linear combinations of these outputs back through the gain matrix. Kalman theory contributes to error controller design primarily through the algorithm used to obtain the gain matrix. The gain matrix obtained with Kalman theory takes cognizance of the statistical state of the modeled errors as represented by the covariance matrix and the quality and type of observation available.

As the error state vector changes with time between the system correction instants as described by Equation 6.1, the covariances of these error states also change. These changes in the statistical state of

the system error can be determined in real-time through the propagation of the system error covariance matrix. Alternate methods have been derived for performing the computations required for propagating this information. These methods include direct integration of the differential equation for the covariances and discrete propagation through use of the transition matrix. Other schemes[4] deal with root mean square values rather than the variances of the errors to reduce the precision requirements in the digital computation. In any case, the particular application under consideration and the equipment available for system implementation will determine which of these computational techniques is the most advantageous to employ. When a measurement is made with a sensor, the correction of the system requires the computation of the gain matrix. When a system correction is made, changes occur in the statistical inter-relationships between the error state variables. At the update time the covariance matrix elements are altered to reflect the fact that the system correction was made.

A review of the above paragraphs concludes that the following basic operations must be implemented in the system digital computer to accomplish the mechanization of a Kalman error controller:

1. Propagation of covariance matrix variables between instants of time at which measurements are made.
2. Selection and sampling of the measurement data to be processed at the system update time.
3. Computation, based on the values of system variables at the update time, of predicted values for the selected measurements and formation of the differences of these predictions with the actual measurement data to obtain observed errors.
4. Computation of the Kalman gain matrix, based on covariance matrix values at the update time and the type and quality of the measurement data.
5. Weighting of the observed error with the Kalman gain matrix to obtain the correction term used to update the system variables whose errors are represented in the design error model.
6. Updating of the covariance variables at the update time to reflect the execution of the system correction.

The logical relationship of these basic operations is summarized schematically in Figure 8.

6.2 Error Controller Configurations

Two major approaches have been proposed for the configuration of the error control mechanism. In the first and most popular approach, the system correction vector determined at a measurement time is used to update values of the navigation variables and to bias noise sources within the system. This is the basic feedback scheme already discussed. In the second approach, only the system outputs are corrected using the values of propagated estimates of the system errors.

6.3 Feedback Correction

In the feedback correction scheme, the system is organized schematically as shown in Figure 9. Here corrections are made to those errors in each subsystem which are included in the design error model. The feedback of the corrections derived by weighting the observed error with the Kalman gain matrix has two advantages. Since the Kalman controller is based on a linear error model of the system, any process which keeps this assumption valid is advantageous. Since proper feedback correction will tend to decrease error within the system, it is advantageous to maintain the linearity assumption made for the behavior of the errors. Classically, feedback has been used to neutralize automatically the effect of variations within a system. This insensitivity property inherent to feedback systems leads to the second advantage in employing the servo-mechanism approach. The more error states which are included in the design error model, the better should be the performance obtained by the system. However, the basic insensitivity realized by feedback to the system variations implies that the effect of many sources of error will automatically be reduced even when they are not included in the design error model.

6.4 Output Correction

In the output correction scheme, the system is organized schematically as shown in Figure 10. Here no direct corrections are made to assumed sources of error in the subsystems. The linear error model equations for the system are used to propagate the error state vector estimate. A correction vector is then derived as a linear transformation of the propagated error vector and used to update only the system outputs.

The chief disadvantage of such an approach is that error is introduced through imperfect modeling. This occurs through two effects. Since error in the system is not corrected toward the null, as with the feedback method, the assumption of linearity can become invalid. Also, the insensitivity characteristic of the feedback scheme is lost. Consequently, a higher-dimensional linear error model will probably have to be employed as a basis of the Kalman filter to obtain performance comparable with that obtained using the feedback scheme. Another disadvantage of this approach is that an estimate of the modeled system error states must be propagated. This is avoided in the feedback mechanization. An advantage of this approach is that improper updates resulting from undetected spurious subsystem operation will affect only the outputs. In the feedback method an improper update can drive a subsystem to an improper operating region, resulting in a loss of overall system performance. However, such behavior can be avoided in the feedback method by developing effective checks on subsystem operation and the measured error difference.

7. A KALMAN ERROR CONTROLLER DESIGN FOR MULTI-SENSOR AUGMENTED INERTIAL NAVIGATION SYSTEMS

7.1 General

The Doppler radar, astro-tracker, LORAN receiver and inertial platform are representatives of different classes of navigation sensors having complementary characteristics which can be employed to advantage in an integrated navigation system. The autonomous and conventional integrated operation of these sensors has been discussed. The Doppler-inertial, astro-inertial and LORAN-inertial-navigation modes are re-examined below, using the more sophisticated Kalman linear correlation theory for more optimal mixing of the sensor data. The mechanization obtained is suitable for implementing any combination of the foregoing modes, including the fully-integrated astro-Doppler-inertial-LORAN navigation mode.

As previously noted, a straightforward application of Kalman theory may not be desirable or necessary in obtaining a navigation system error control mechanization: not desirable, owing to the magnitude of the computer requirements implied for an implementation based on a detailed error model of each sensor; not necessary, as the mission to which the system is to be applied may not warrant a highly-complex mechanization. A thorough examination of the navigation system requirements generated by the intended mission must be made if the resulting design is to be cost-effective.

The mission considered in synthesizing the navigation system of this section is termed fast-reaction, meaning that accurate values of vehicle position, velocity and attitude are desired within a short period of time after the system is activated. In this mission no operational restrictions on the vehicle should result from any characteristic of the navigation system operation. For this mission the system error control mechanization described below was derived using synthesis procedure described in Section 5. The performance of the navigation system, implemented with the derived Kalman mechanization, is illustrated for the fast-reaction mission in the principal integrated modes of operation.

7.2 Derivation of the Kalman Mechanization

The straightforward application of Kalman theory to a detailed error model of an astro-Doppler-inertial-LORAN system could easily involve computations dealing with fifty or more error state variables having quite complex dynamic interactions. If a Kalman mechanization were based on such a model, corrections would be obtained for the accounted sources of equipment and environmental error, as well as the system's navigation outputs. In addition, through a detailed description of the dynamic interaction of variables, error propagation over the long-term would be accurately predicted.

An examination of the application of the navigation system to meet the requirements of the fast-reaction mission results in the primary requirement that the inertial platform be rapidly aligned regardless of the operational environment. Rapid platform alignment is necessary as the mission requires dynamically accurate attitude and velocity information, which are natural outputs only of the inertial system. Accurate values of these variables will not be available until platform alignment is achieved. If platform alignment and control of position and velocity errors are the principal objectives of the Kalman error control mechanization, it remains only to determine which of the system errors and interactions has especial importance in meeting this objective. Since position, velocity and attitude are the required system outputs, it is clear that errors in these variables must be controlled. Clearly then, a minimal mechanization would incorporate at least the errors in these variables as components of the state vector of the simple error model.

Platform alignment is essentially a process of eliminating high initial values of attitude error and any error in computed velocity or position occurring as a result of attitude error. The existence in the system error model of biases or noise sources associated with equipment or induced by the environment will then only be relevant if such sources seriously affect alignment. The optimal control of errors resulting from stochastic inputs to a system is usually a problem which occurs more often in the steady-state. In navigation systems, such error control is of importance in applications involving long-term cruise missions requiring extremely high navigation accuracy. For the alignment application, many such sources are not likely to be very important and consequently, can be eliminated in the simple error model. It will be recalled that the loss in uncertainty, resulting from an elimination of these variables from the error model, can be partially accounted for by artificially increasing the variances of appropriate white noise sources in the simple model. However, exceptions to this philosophy exist. These exceptions can occur when the error in question is a long-term noise which constitutes a significant fraction of the difference between a measurement value and the prediction of that value during platform alignment. When this situation exists, the modeling error, resulting from an exclusion of the noise source in question, can cause a significant loss in system performance which cannot be offset by white noise parameter adjustments. One case of this sort occurs during gyro compassing of the inertial platform with Doppler speed data containing large biases. Another example is afforded by the case in which there exists significant distortion of the actual electronic LORAN time difference grid from the theoretical grid based on geodetic data. This distortion is caused by variations in the propagation speed of the electromagnetic energy over the earth's surface resulting from unaccounted variations in the nature of the paths of propagation.

In addition to the elimination of most noise sources in the alignment application, dynamic interactions between the retained error state variables can be reduced to a minimum by eliminating those defining error propagation characteristics which are long-term with respect to the fast-reaction mission. Such couplings are associated primarily, with the inertial sensor defining the characteristic 24-hour oscillations and modulating Foucault frequencies which are present in a detailed error model.

7.3 Kalman Error Controller Description

After examining a detailed error model of the navigation system and the requirements of the fast-reaction mission along the lines outlined, the following minimal error control mechanization can be derived. This mechanization will incur only a small loss in system performance during platform alignment, if relatively accurate sensors are employed in the navigation system. The mechanization assumes the north-slave, Schuler-tuned platform implementation, but can accommodate any local-level platform mechanization with only slight modification. It is a minimal design in the sense that only the errors in the required system outputs are represented in the simplified error state vector, consequently reducing the model dimension to a minimum. The seven elements of this state vector are

$$\underline{x} = \begin{bmatrix} \delta\theta_N \\ \delta\theta_E \\ \delta V_N \\ \delta V_E \\ \phi_z \\ \phi_N \\ \phi_E \end{bmatrix}, \tag{7.1}$$

where

$\delta\theta_{E,N}$ are the angular errors in system computed vehicle position, represented as rotations about the local geographic east and north axes, respectively;

$\delta V_{E,N}$ are the errors in system computed vehicle velocity along the local north and east axes, respectively;

$\phi_{E,N,Z}$ are the angular errors in platform attitude, represented as rotations about the local east, north and vertical axes, respectively. The east and north components are usually referred to as platform tilts while the third component is called platform azimuth misalignment.

The matrix defining the dynamic interaction of these seven error state variables is

$$\bar{A}(t) = \begin{bmatrix} 0 & 0 & 0 & R^{-1} & 0 & 0 & 0 \\ 0 & 0 & R^{-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -s_{Ep} & 0 & s_{Np} \\ 0 & 0 & 0 & 0 & s_{Np} & -s_{Ep} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & R^{-1} & \omega_{Ec} & 0 & 0 \\ 0 & 0 & -R^{-1} & 0 & -\omega_{Nc} & 0 & 0 \end{bmatrix}, \tag{7.2}$$

This matrix has been reduced to include only those terms essential to alignment followed by cruise navigation for a short time period. The result is that only ten of the elements are non-zero. Of these, four are identical constants R^{-1}, R being a nominal radius for the earth and four are derived from the specific force measurements, s_{Ep}, s_{Np} and s_{Zp}, along the local east, north and vertical axes, respectively. The final two elements are spatial rate terms, ω_{Ec}, ω_{Nc}, being the computed rates at which the platform must rotate about the local east and north axes, respectively, to remain locally-level. Clearly, this representation of inertial error behavior is much simpler than that described in Section 3.

The zero-mean, white disturbance noise vector, corresponding to Equation (7.1), can include up to seven components,

$$\bar{\xi} = \begin{bmatrix} \bar{\xi}_{\delta\theta_N} \\ \bar{\xi}_{\delta\theta_E} \\ \bar{\xi}_{\delta V_N} \\ \bar{\xi}_{\delta V_E} \\ \bar{\xi}_{\phi_E} \\ \bar{\xi}_{\phi_N} \\ \bar{\xi}_{\phi_E} \end{bmatrix}, \tag{7.3}$$

to account to a certain extent for the deletion of noise state variables from the detailed system error model. The variances of the elements $\bar{\xi}_{\phi_{E,N,Z}}$ are of principal importance in accounting for errors which have an

effect on the platform as does a gyro drift rate about the corresponding axis. The variances of the elements $\bar{\xi}_{\delta V_{N,E}}$ are introduced to account for errors which have a stochastic effect on the velocity computation as does accelerometer noise, quantization error, misalignment, etc. Finally, the variances of the elements $\bar{\xi}_{\delta\theta_{N,E}}$ reflect additional uncertainty in the position computation.

The description of the error model given above corresponds to propagation between instants of time at which system updates are made. To employ Kalman's algorithm at an update time, a model of observation error must be constructed. In an astro-Doppler-inertial-LORAN navigation system, the measurements that can be made with the sensors which augment the inertial sensor are Doppler ground speed, Doppler drift angle, LORAN time-difference(s), tracker elevation angle, and tracker bearing angle.

When one of these measurements is made, it is differenced with a predicted value of the measurement based on the computed values of the system navigation variables. If the associated errors are sufficiently small, the measurement difference can be represented as a linear combination of the errors involved. The coefficients of the error terms in such an expression are simply gradient components evaluated at the local operating point. Since noise terms associated with the tracker, Doppler and LORAN sensors are not incorporated in the simple error model, the observation matrices involve only gradient components relative to the errors in the navigation variables. These matrices are now defined for the different measurements.

The vehicle ground speed observation matrix is expressed as

$$H_{\Delta V_g} = [0 \;\; 0 \;\; \cos\theta_c \;\; \sin\theta_c \;\; 0 \;\; 0 \;\; 0] , \tag{7.4}$$

where θ_c is vehicle track angle and $\Delta V_g \triangleq [V_{gc} - V_{gD}]$, denotes the difference between computed, V_{gc}, and Doppler-measured, V_{gD}, ground speed.

The vehicle drift angle observation matrix is expressed as

$$H_{\Delta\delta} = \left[0 \;\; 0 \;\; \frac{\sin\theta_c}{V_{gc}} \;\; \frac{-\cos\theta_c}{V_{gc}} \;\; 1 \;\; 0 \;\; 0\right] , \tag{7.5}$$

where $\Delta\delta \triangleq [\delta_c - \delta_D]$ denotes the difference between computed, δ_c, and Doppler-measured, δ_D, drift angle.

The time-difference observation matrix is expressed as

$$\tilde{H}_{\Delta t} = \frac{R}{C}[(\sin\beta_M - \sin\beta_S)_c \;\; (\cos\beta_M - \cos\beta_S)_c \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0] , \tag{7.6}$$

where $\beta_{M,S}$ are the computed bearing angles to the LORAN master and slave stations relative to the local north line, respectively, C is the velocity of light, and $\Delta t \triangleq [t_c - t_L]$ denotes the difference between computed, t_c, and LORAN-measured, t_L, time-difference.

The elevation angle observation matrix is expressed as

$$H_{\Delta E} = [\sin E_c \;\; -\cos E_c \;\; 0 \;\; 0 \;\; 0 \;\; -\sin E_c \;\; \cos E_c] , \tag{7.7}$$

where E_c is the computed star bearing angle with respect to the local north axis, and $\Delta E \triangleq [E_c - E_T]$ denotes the difference of computed, E_c, and tracker-measured, E_T, star elevation angle.

The bearing angle observation matrix is expressed as

$$H_{\Delta B} = [(-\cos B \tan E + \tan\phi)_c \;\; (-\sin B \tan E)_c \;\; 0 \;\; 0 \;\; -1 \;\; (\cos B \tan E)_c \;\; (\sin B \tan E)_c] , \tag{7.8}$$

where ϕ_c is the computed latitude, E the computed star elevation with respect to the local-level plane, and $\Delta B = [B_c - B_T]$ denotes the difference between computed, B_c, and tracker-measured, B_T, bearing.

The final elements of the error model to be defined are the zero-mean, additive white noise processes associated with each of the measurements cited above. As has been stated previously, uncertainty can be artificially associated with the observation process in the simple error model by increasing the variances of these noise components from values associated with the detailed error model. When the sensors used to align the inertial platform are relatively accurate, the components deleted in constructing a simple error model on which to base the Kalman mechanization can be accounted for adequately by this method. For the above system, the vector denoting additive white observation noise would include:

$$\eta = \begin{bmatrix} \eta_{\Delta V_g} \\ \eta_{\Delta\delta} \\ \eta_{\Delta t} \\ \eta_{\Delta E} \\ \eta_{\Delta B} \end{bmatrix} . \tag{7.9}$$

The variances of the elements $\eta_{\Delta V_E, \Delta\delta}$ are used to account for noise error in the Doppler ground speed and drift angle measurements which vary with vehicle speed, altitude, attitude and terrain. The variance of the element $\eta_{\Delta t}$ represents LORAN-measured time-difference noise error, which varies with distance from the transmitting stations, local interference and some anomalies associated with the paths of propagation. Finally, the variance of the elements $\eta_{\Delta\epsilon, \Delta\beta}$ account for noise error associated with tracker measurements of star elevation and bearing, respectively, which vary with star brightness and atmospheric conditions.

7.4 Description of a Fast-Reaction Mission

We next illustrate the relative performance achievable in the different navigation modes when the Kalman feedback error controller described is employed. This is done for a nominal fast-reaction mission for which this controller was designed. The performance of the system for the major modes of navigation is discussed in detail later.

The vehicle flight dynamics result in the profile indicated schematically in Figure 11. In more detail, it was assumed that the aircraft spent two minutes on the ground after the inertial system was activated. Alignment of the platform was initiated when the inertial system was activated. The initial r.m.s. attitude errors of the platform were assumed to be 300 arc-seconds in tilt and 3600 arc-seconds in azimuth. The vehicle was assumed to remain stationary having no velocity during the two-minute ground alignment period. This results in the null-velocity mechanization commonly employed for ground alignment of inertial platforms. The assumption in reality is in error because of platform motion caused by wind buffeting and disturbances, resulting from the loading of fuel, cargo, armament etc., of the airframe. After the two-minute ground alignment period at coordinates of latitude $\Phi = 12$ deg , and longitude, $\lambda = 106$ deg , the vehicle was accelerated on a heading $\theta = 180$ deg southward and vertically at 28 and 16 ft/sec^2, respectively. When a ground speed of 1000 ft/sec was achieved, a turn of 180 deg at a rate of 3 deg/sec was initiated such that, at 1 min 40 sec into the mission, the aircraft was directed northward. The aircraft proceeded at a climb rate of 80 ft/sec until the cruise altitude of 10,000 ft was achieved. After one half hour, the aircraft turned 90 deg at a rate of 3 deg/sec and proceeded eastward for another half hour. The mission terminated after one hour of operating time.

7.5 System Performance for the Nominal Fast-Reaction Mission

The graphs shown in Figures 12-23 illustrate system performance obtained in the free-inertial, Doppler-inertial, LORAN-inertial, and astro-inertial navigation modes for the fast-reaction mission. The performance measures displayed are the error in computed system position and velocity and the azimuth misalignment at the inertial platform. The performance was obtained through the use of a digital computer Monte-Carlo simulation program. The curves reflect sample root mean square (r.m.s.) values of the indicated errors obtained from 30 trial runs. This sample size does not give precise results but indicates the appropriate order of magnitude of trends. Extensive models incorporating random bias, correlated noise and white noise were employed to represent the error characteristics of each navigation sensor. The bias errors present in each sensor are constant over any one simulation run. They vary from run to run by virtue of being selected randomly from a zero-mean, uniform distribution of prescribed standard deviation. These distributions were derived using digital random number generators. The essentially white noise sources in each sensor are simulated as sequences of samples selected randomly from zero-mean, uniform distributions of prescribed standard deviations derived in the same manner as for the bias errors. The correlated noise sources in each sensor are represented as exponential noise. Such noise is simulated by passing simulated white noise through a first-order digital smoother whose analog time constant is equal to the desired correlation time of the noise. The standard deviation of the exciting white noise sequence is selected to obtain the prescribed standard deviation desired for the correlated sequence.

In the following paragraphs, performance obtained in each of the navigation modes, as shown in the figures, is discussed. The mechanization used for error control in all the navigation modes was the Kalman feedback design, based on the extremely simple system error model described above. In each case the quality and quantity of data available from each sensor is defined and any unique feature of the particular mode of operation is discussed.

7.6 Free-Inertial Navigation Mode

The system performance obtainable in the free-inertial mode of navigation depends critically on the amount of time allotted to the inertial platform alignment on the ground. If the aircraft takes off with large platform attitude errors, relatively large errors in computed system position and velocity will result. This obtains as the acceleration information measured in platform coordinates is processed in the computer on the assumption that the accelerometer axes are coincident with the local geographic axes. In the augmented inertial navigation modes, completion of ground alignment is not as critical, as the desired coincidence of the actual accelerometer axes and the geographic coordinates can be obtained through the use of data from the augmenting sensor.

During stationary ground alignment, platform attitude error is removed in two stages, referred to respectively as leveling and gyro compassing. When the level accelerometers deviate from the local horizon plane, their sensitive axes detect components of the local gravity vector. These measured components result in errors in computed system position and velocity through subsequent integrations in the computer which can then be used to remove the platform tilts. Note, however, that, if the platform is assumed stationary, exact determination of the attitude error cannot be achieved when platform motion is caused by local stochastic disturbances.

If there exists a misalignment in the platform azimuth once the initial level error is approximately removed, the platform still tends to deviate from the level plane. This obtains as the platform is rotated to maintain

the local-level orientation as the earth rotates. Deviations of the platform axes from the assumed coordinates result in erroneous rotation of the platform coordinate system. The resulting tilt caused by the platform azimuth misalignment will again generate errors in computed system position and velocity. These errors then provide the basis to remove approximately the azimuth error. It should be noted that effective determination of the error in computed system position or velocity due to azimuth misalignment can occur only after platform level has been obtained. During ground alignment, platform leveling is achieved much more rapidly than the gyrocompassing of the platform azimuth. Further, as the latitude at which alignment occurs is increased, the time required for the azimuth gyrocompassing also increases. Both these effects obtain as the coefficient in the linear error model of the inertial system which couples platform azimuth misalignment to platform tilt is small with respect to unity and decreases in magnitude as the latitude is increased. The speed at which overall alignment is achieved is also dependent on the system correction rate and the r.m.s. value and correlation characteristics of the stochastic motions of the platform. For the examples of this section, a system correction was made every second, and the null reference in velocity was assumed to be characterized by a white noise with an r.m.s. value of 1/8 ft/sec.

From this discussion we see that the platform will be aligned only after the azimuth error is minimized. In Figures 12-14, system performance is illustrated for the free-inertial modes for ground alignment times of 2, 4 and 10 minutes. At 10 minutes, ground alignment has essentially been completed and subsequent errors shown in computed system position and velocity are due primarily to stochastic errors in the inertial equipment. From these particular results, the inertial system would be referred to as a 0.75 mile per hour inertial system. This inertial sensor error characteristic applies for all the performance curves shown in this section. At 2 minutes, a relatively larger azimuth misalignment exists. Consequently, when the aircraft executes the flight dynamics described above, the acceleration measurement error due to azimuth misalignment dominates the propagation of error in computed system position and velocity.

7.7 Doppler Inertial Navigation Mode

The system performance obtainable in the Doppler-inertial mode of navigation does not depend critically on the ground alignment time, as once the system is airborne, speed measurements made with the Doppler radar can be used to continue the platform alignment. During the period of airborne alignment, system performance will obviously be degraded with respect to the performance which results when a longer platform ground alignment time is assumed. Since the inertial and Doppler sensor both employ the inertial platform for stabilization, no deviation between the inertial and Doppler velocity vectors occurs instantaneously on take-off due to platform attitude error. The detection of attitude error results therefore from the differing longer term error characteristics of the two sensors. The characteristic errors in the inertial system have frequencies associated with the Schuler, 24-hour and Foucault oscillations. Since the presence of these oscillations only in the platform attitude error affects the stabilized Doppler data, a basis for error separation at the velocity level exists. It turns out that airborne platform alignment Doppler data is equivalent in form to ground alignment assuming a null velocity reference.

The rapidity at which the velocity error characteristics can be separated depends on the quality and quantity of the Doppler data. For two of the performance curves shown in Figures 15-17, the Doppler was assumed to provide horizontal speed components with an r.m.s. bias error of 0.1% and an r.m.s. white noise error of 1%. Experiments made for this mission, as indicated in Figures 15-17 show that doubling of the r.m.s. bias values results in significant loss of system performance if the minimal Kalman mechanization is employed. Such sensitivities can be determined precisely through the use of the correlation analysis technique described in Section 5. In the cases considered in Figures 15-17, different white observation noise covariance matrices, R, were employed. These matrices differed over the time period from 2 to 30 minutes in which most of the platform azimuth misalignment is achieved with the Doppler data. When the R matrix is increased by a factor of three, the tuning factor, errors in computed velocity and the platform azimuth alignment are significantly reduced. This occurs because the Doppler bias errors, unmodeled in the Kalman controller, can be accounted for to some extent, artificially by increasing the white observation noise variances.

Using the Doppler data, corrections to the modeled errors in computed system position, velocity and the platform attitude were made with the Kalman error controller every 5 seconds. These updates were initiated at the vehicle take-off. An important characteristic of system error using the Kalman mechanization is the correction of position realized with the velocity data. In conventional design, since the correlation information relating measured velocity error and position error is not available, no basis for updating the error in computed system position exists. Consequently, only a bounding effect on velocity error is realized in contrast to effective control of all the modeled errors with the Kalman design.

7.8 LORAN-Inertial Navigation Mode

System performance obtained in the LORAN-inertial navigation mode is essentially limited only by the quality and quantity of the geodetic position data derived from the LORAN time-differences. Unlike the Doppler updating process, extremely rapid alignment of the inertial platform can be realized with the receiver data for the fast-reaction mission. This obtains as the platform does not play the data stabilization role with LORAN data that it does with Doppler data. When the aircraft takes off, transient acceleration errors and, subsequently, velocity and position errors, exist in the inertial system as a result of the platform attitude error. These components of inertial system position error are readily detectable using time-difference data obtained in a region of defined LORAN coverage. Correction of platform attitude and computed velocity is achieved with the correlation data relating measured position error to error in computed system velocity and the platform misalignment angles. The real-time availability of this information accounts automatically for any vehicle flight profile or sequence

of measurements of LORAN time-differences. In conventional designs, a mechanization which accounts for the infinite variety of flight and aid data profiles possible would require an extensive amount of complex bookkeeping in the airborne computer, while being inherently sub-optimal.

The principal complication involved in employing LORAN is presented by any warpage error which exists between the actual time-difference grid and a predicted time-difference grid based on accurate computation using geodetic data. The warpage is due to unaccounted variations along the transmission paths between the LORAN stations and the aircraft, which affect the speed of propagation of the electromagnetic energy. Modeling of these errors is highly complex, as the changing propagation paths continually introduce new conductivity and reflecting surface variations. In the fast-reaction mission, the effect of grid distortion can be minimized by accepting an initial system position at the airfield which is based on actual time-difference data. This approach results in corrections to the system which are degraded only by the changes in distortion which occur between the airfield and point of initial use of LORAN measurements. Consequently, if the distortion correlation distance is large with respect to distance traveled before the LORAN is used, system performance loss due to grid warpage modeling error will be small. Experiments have indicated that, if accurate stochastic models of the warpage phenomena are employed, only small corrections to the grid distortion result. Consequently, a LORAN-inertial system provides geodetic position information whose accuracy is limited by the LORAN time-difference grid distortion. Accurate navigation can still occur, however, when the aircraft and destination positions are specified in terms of measured time-difference coordinates.

For the performance curves shown in Figures 18-20, the LORAN time-difference data were assumed to be contaminated by first-order, 50 second correlated noises with r.m.s. values of 0.14 micro-seconds and white noises with r.m.s. values of 0.005 micro-seconds. The LORAN time-difference data was used to make a system correction every 10 seconds, these updates being initiated 60 seconds from the take-off time. In addition, constant bias or warpage error in the time of propagation from each transmitter to the aircraft are assumed in two of the cases shown. The r.m.s. values used were 1 and 5 micro-seconds. In these cases, the system position was not initialized using the LORAN time-difference data with the result that highly degrading updates to platform azimuth alignment and computed system velocity are made when the LORAN data is first used after take-off. Once computed vehicle position has been updated with the LORAN time-differences, error due to the transmission path biases appears in the computed system position. Except for the changes in geodetic position errors due to the changing effects of the time-difference biases as the vehicle moves in the LORAN coverage area, the subsequent observation errors are caused primarily by the other system error sources. Consequently, with further updating, computed vehicle velocity and platform attitude error are brought under control as indicated in the figures.

The geometry involving an assumed LORAN chain of three transmitters and the aircraft profile is indicated in Figure 11. The master station is at 12^{0} 42′ latitude and 101^{0} 6′ longitude, the two slaves are at 18^{0} 18′ and 8^{0} 45′ latitude and 99^{0} 30′ and 106^{0} 40′ longitude, respectively. As the aircraft travels through the coverage area, the capability of deducing position from the two time-differences varies due to changes in the gradients relating time-difference error to geodetic position error. In this example, position can be determined to a r.m.s. precision of 1712 ft/μsec at the airfield and 3901 ft/μsec at the mission termination point for independent r.m.s. errors of a micro-second in both time-differences.

7.9 Astro-Inertial Navigation Mode

System performance in the astro-inertial navigation mode is limited by the capability of the tracker to lock onto a light source in the ambient noise environment and the accuracy with which the tracking axis can be related to the local geographic axes. The tracking precision is dependent on electronic component noise of the tracker and the brightness of the star relative to its immediate background. The accuracy in determining the angular orientation of the tracker axis with respect to the local geographic coordinates is dependent on the accuracy of the mechanical alignment of the tracker gimbal axes to the reference inertial platform axes and to the accuracy of the alignment of these platform coordinates to the geographic frame.

In the fast-reaction mission, the dominant system error at the first star shot is the platform azimuth misalignment. This error is very observable at this time, as the error in computed system position and error associated with the tracker are relatively small. Consequently, an accurate alignment of the platform azimuth is achieved with the first system correction. Subsequent system corrections serve primarily to bound the three platform misalignment angles and the errors in computed system position.

Significant errors can be made in conventional astro-inertial system mechanizations, as these mechanizations usually assume that the platform reference coordinate system is locally-level. This assumption results in erroneous updates to computed position and platform azimuth to offset any existing tilt. An arc-second of position error created by an arc-second of platform tilt is approximately 100 feet. In a Kalman error controller design, the contribution of tilt to observed star position error can be modeled and accounted for properly in system updating. The sensitivity between attitude error and system position error implies, however, that tilt inducing errors require adequate modeling in the Kalman design. The main advantage of a Kalman mechanization in astro-inertial navigation systems is the automatic accommodation of the infinite variety of star-to-vehicle position relationships possible. Data resulting from each star shot in an arbitrary sequence can be optimally employed, as consideration is automatically made for the residual error coincident with the optical axis and the current relationship between attitude error and error in computed system position when the correction is determined.

For the performance curves shown in Figures 21-23, the star elevation and bearing angle measurements are contaminated by a first-order correlated noise with r.m.s. value of 5 arc-seconds and a correlation time of 4 hours. In addition, a white noise of r.m.s. value 2 arc-seconds was assumed. Two stars were considered in making the system corrections. These stars had elevation angles of 30 and 60 degrees and bearing angles of zero and 120 degrees, respectively. The shots alternated between the two stars, being made every 60 seconds and initiating 60 seconds after take-off. In Figures 21-23, selected curves from the preceding sets of figures have been displayed to permit convenient comparisons of system performance.

8. CONCLUSION

This chapter has considered the application of Kalman filtering theory to error control in multi-sensor navigation systems by detailing a particular error control design. The reader has been introduced to the basic types of existing navigation sensors and how their complementary characteristics have been exploited in the past in augmented inertial navigation systems. The practical aspects of implementing the much more sophisticated Kalman error controllers were discussed as well as practical synthesis techniques which have been employed to realize computationally feasible designs. The design example which was considered illustrated system performance in the principal navigation modes to portray the particular characteristics of the navigation sensor augmenting the inertial system.

The state-of-the-art in the application of highly-complex data processing techniques, such as Kalman filtering theory, is rapidly changing. The use of these sophisticated procedures is limited by the capability of digital computation equipment available for the synthesis of appropriate designs and the mechanization of these designs in actual systems. Limitations in some applications exist owing to the lack of information adequately identifying and describing all consequential system error characteristics.

As the cost of computation decreases and more information defining system error behavior becomes available, more marked improvements in navigation system accuracy can be achieved with existing sensors. A relaxation of individual sensor accuracy requirements permitted by more complex software design will, in turn, permit simpler sensor design. Sensor simplification coupled with natural design improvements can result in less costly, lighter, more reliable and more easily maintained sensors. These more sophisticated error control techniques imply a shift from complexity in hardware to complexity in software, resulting in navigation system designs with higher accuracy and reliability, simpler maintenance, and reduced size, weight and cost.

REFERENCES

1. Kalman, R.E. — *New Methods and Results in Linear Prediction and Estimation Theory.* Research Institute for Advanced Study, Baltimore, Maryland, Tech. Report 61-1, 1961.

2. Bryson, A.E. Jr; Johansen, D.E. — *Linear Filtering for Time-Varying Systems Using Measurements Containing Colored Noise.* Institute of Electrical and Electronic Engineers, Transactions on Automatic Control, Vol. AC-10, Jan 1965, pp. 4-10.

3. Bryson, A.E. Jr; Henrikson, L.J. — *Estimation Using Sampled Data Containing Sequentially Correlated Noise.* AIAA Guidance, Control and Flight Dynamics Conference, Aug 1967.

4. Rishel, R.W. — *Convergence of Sampled Discrete Optimum Filters.* Published in the Proceedings of the Joint Automatic Control Conference, June 1968, pp. 863-870.

5. O'Day, J. et al. — *Study and Analysis of Selected Long-Distance Navigation Techniques,* Vols I and II. Institute of Science and Technology, University of Michigan, Dec 1962.

6. Bellantoni, J.F.; Dodge, D.W. — *A Square Root Formulation of the Kalman-Schmidt Filter.* AIAA Journal, Vol.5, No.7, July 1967, pp. 1309-1314.

7. Meditch, J.S. — *Sub-Optimal Linear Filtering for Continuous Dynamic Processes.* Aerospace Corporation, Los Angeles, California, Technical Report TOR-469(5107-35)-2, July 1964.

8. Huddle, J.R. et al. — *Systems Analysis and Mechanization Studies, DIL Navigation System Definition Appendix I.* US Air Force Contract AF33(615)67C-1184, Wright-Patterson Air Force Base, Dayton, Ohio, Oct 1967.

9. Benbrook, H.W. et al. — *Final Engineering Report LORAN C/D Navigation System Integration.* US Air Force Contract AF(657)-14388, Items 2 and 3, Wright-Patterson Air Force Base, Dayton, Ohio, Jan 1966.

10. Klementis, K.A. et al. — *Synergistic Navigation Systems Study, Phases I and II.* US Office of Naval Research Contract N00014-66-C0192, Oct 1966 and Nov 1967.

11. Dworetsky, L.H.; Edwards, A. — *Principles of Doppler-Inertial Guidance.* American Rocket Society Journal, Dec 1959, pp. 967-972.

12. Richman, J.; Friedland, B. — *Design of Mixer-Filter for Aircraft Navigation Systems.* Published in the Proceedings of NAECON Conference, Dayton, Ohio, 1967, pp. 429-438.

13. Friest, D.T. — *Optimization of a Celestial-Inertial Navigation System Using Single Body Tracking.* PhD Dissertation, Iowa State University, 1963.

14. Johler, J.R.; Berry, L.A. — *LORAN D Phase Corrections Over Inhomogeneous, Irregular Terrain.* ESSA Technical Report IER 59-ITSA 56, Nov 1967.

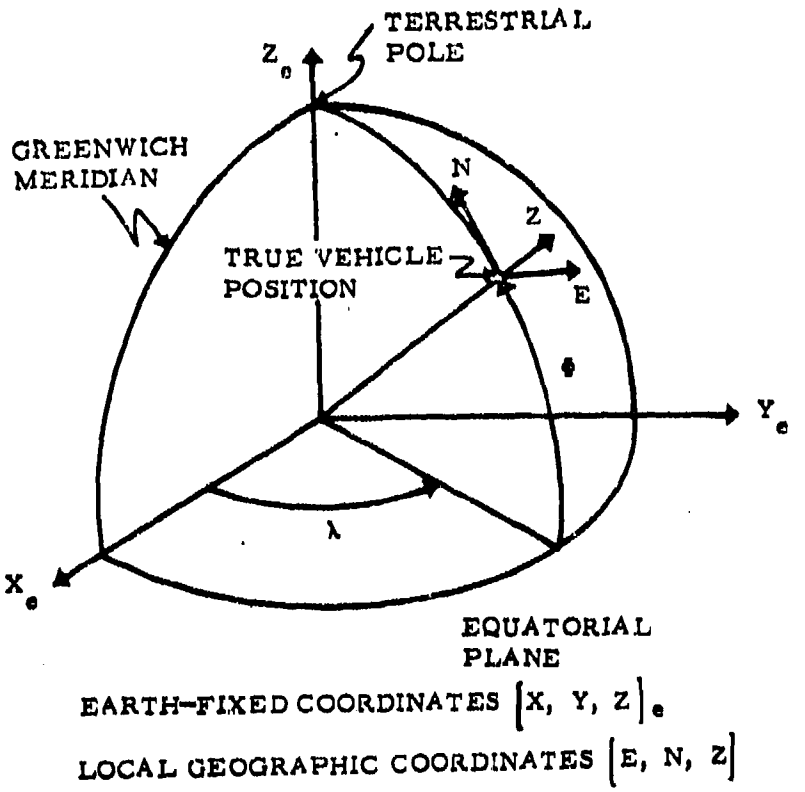


Fig.1 Earth-fixed and local geographic reference coordinate systems

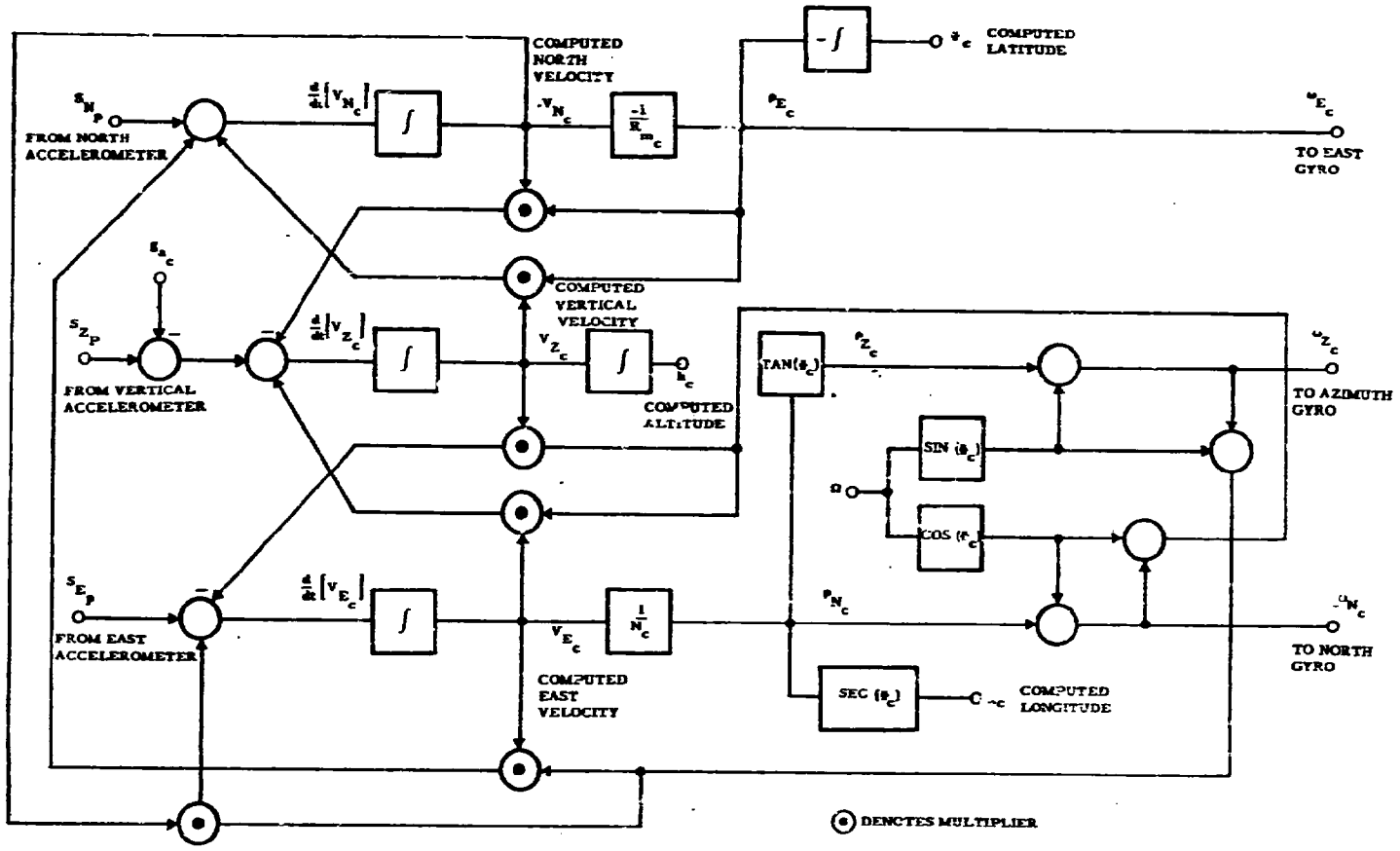
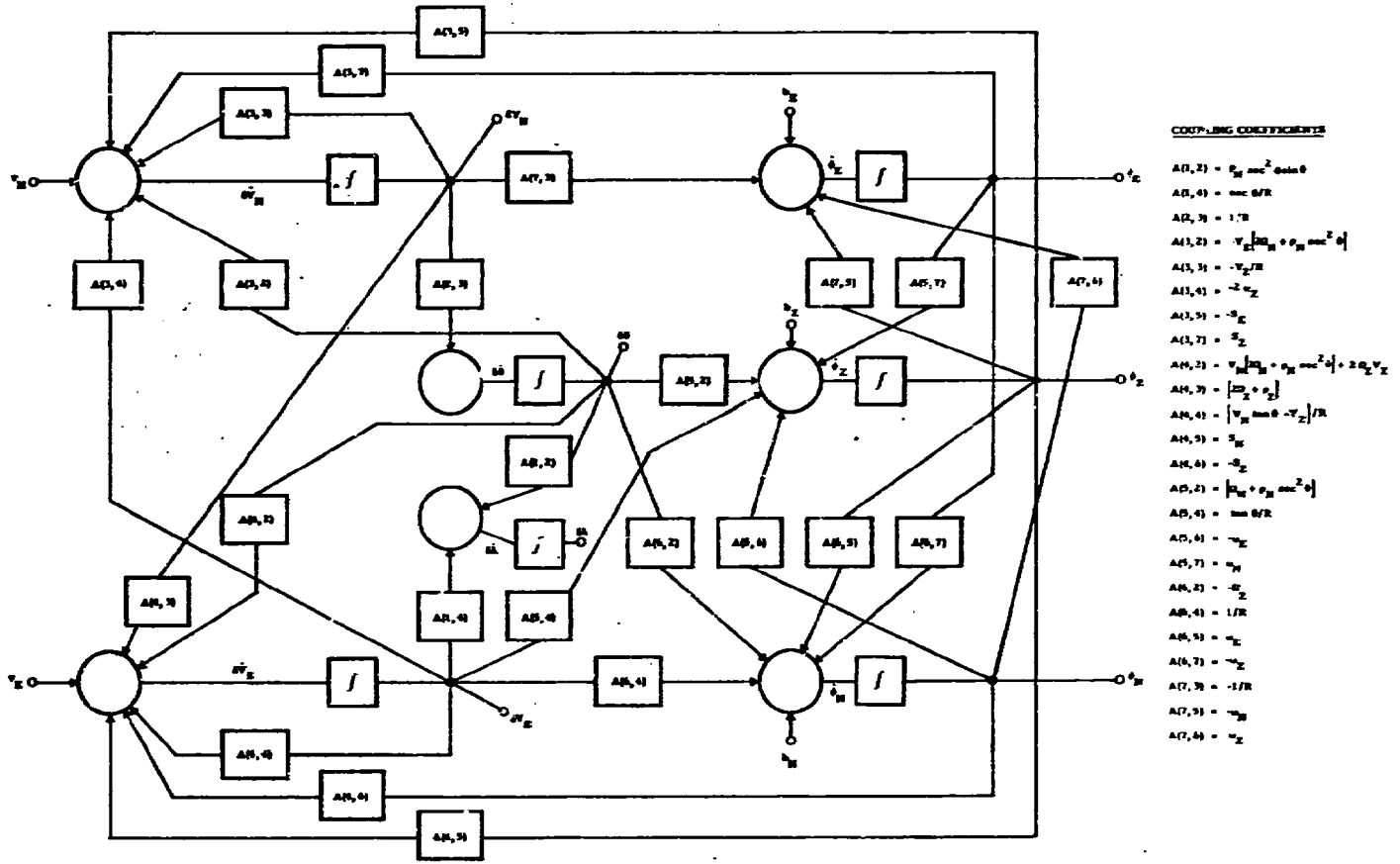


Fig.2 Schuler-tuned, north-slaved, free-inertial mechanization



COUPLING COEFFICIENTS

$A(0,2) = g_N \sec^2\theta \tan\theta$

$A(0,4) = \sec\theta/R$

$A(2,3) = 1/R$

$A(3,2) = -V_E\left[2\Omega_N + g_N \sec^2\theta\right]$

$A(3,3) = -V_E/R$

$A(3,4) = -2\,\omega_E$

$A(3,5) = -\delta_E$

$A(3,7) = \delta_E$

$A(4,2) = V_N\left[2\Omega_N + g_N \sec^2\theta\right] + 2g_E V_E$

$A(4,3) = \left[2\Omega_E + g_E\right]$

$A(4,4) = \left[V_N \tan\theta - V_E\right]/R$

$A(4,5) = \delta_N$

$A(4,6) = -\delta_E$

$A(5,2) = \left[\Omega_E + g_E \sec^2\theta\right]$

$A(5,4) = \tan\theta/R$

$A(5,6) = -\omega_E$

$A(5,7) = \omega_N$

$A(6,2) = -g_E$

$A(6,4) = 1/R$

$A(6,5) = -\omega_E$

$A(6,7) = -\omega_N$

$A(7,3) = -1/R$

$A(7,5) = -\omega_N$

$A(7,6) = -\omega_E$

Fig.3 Linear error model of level channels of Schuler-tuned, north-slaved, free-inertial system

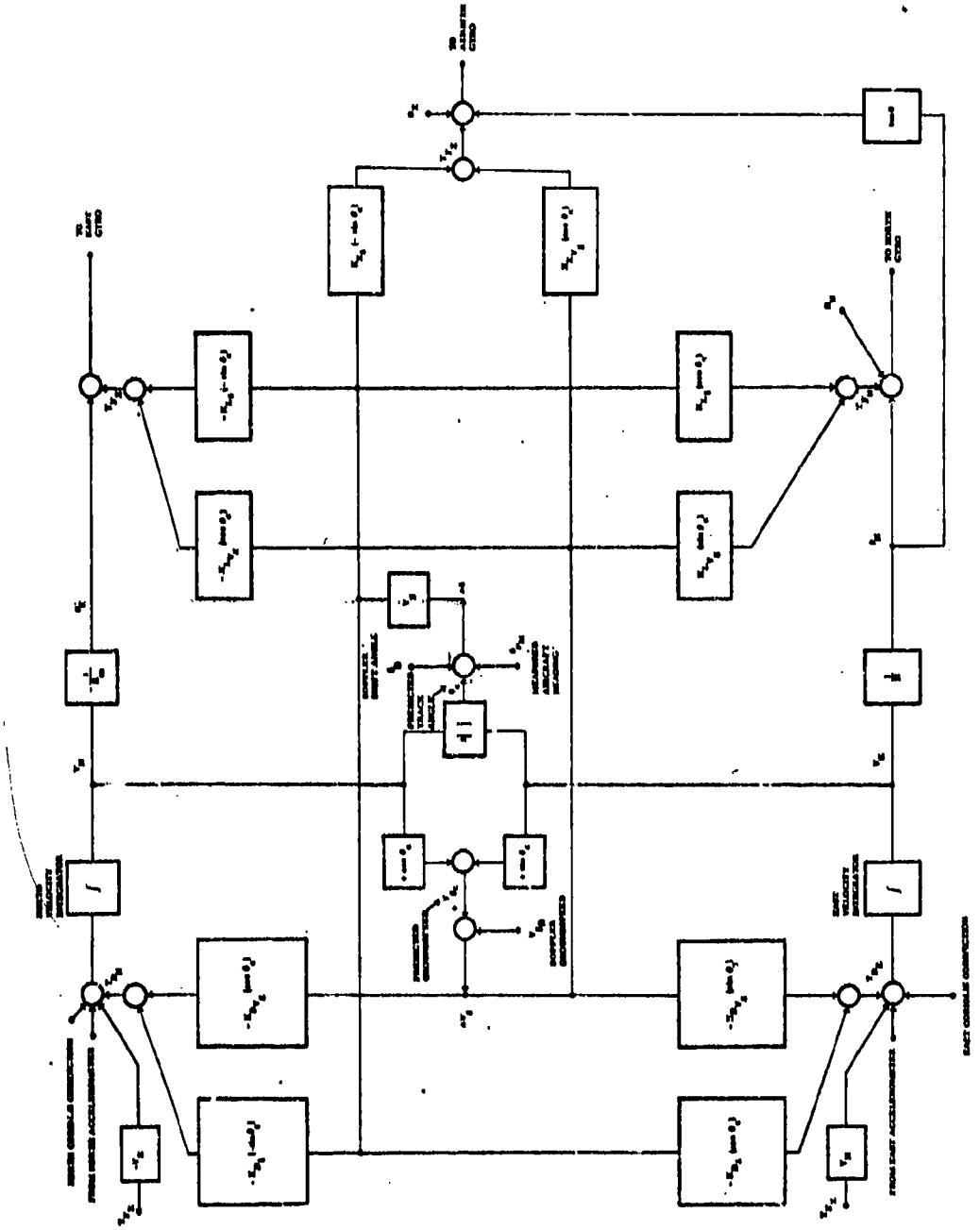


Fig.4 Conventional Doppler ground speed and drift angle damped-inertial navigation system mechanization

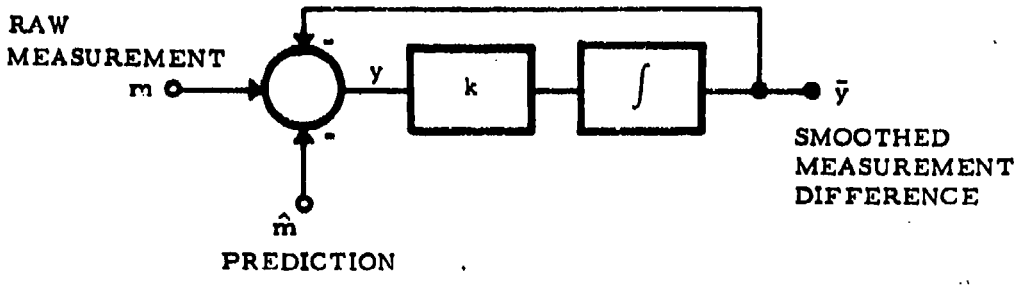


Fig.5 Pre-filter for smoothing the differences between actual measurements and predicted values before system correction

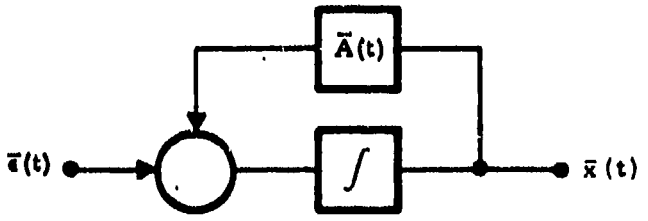


Fig.6 Linear system error model serving as basis for Kalman error controller

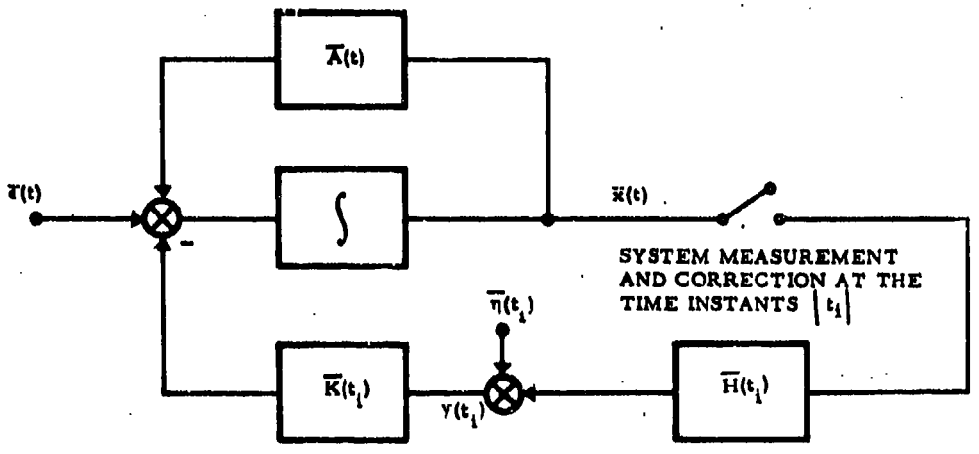


Fig.7 Schematic representation of system error propagation with discrete update

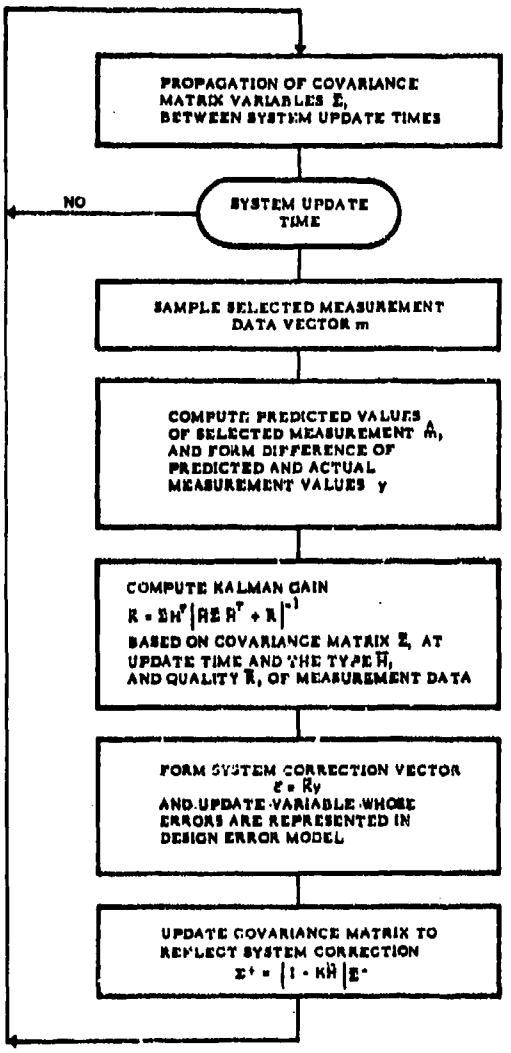


Fig.8 Schematic diagram of basic operations required to implement Kalman feedback error controller

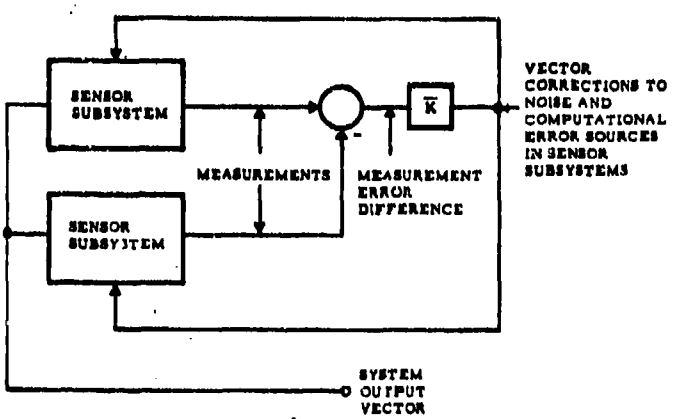


Fig.9 Schematic information flow in system employing feedback error control

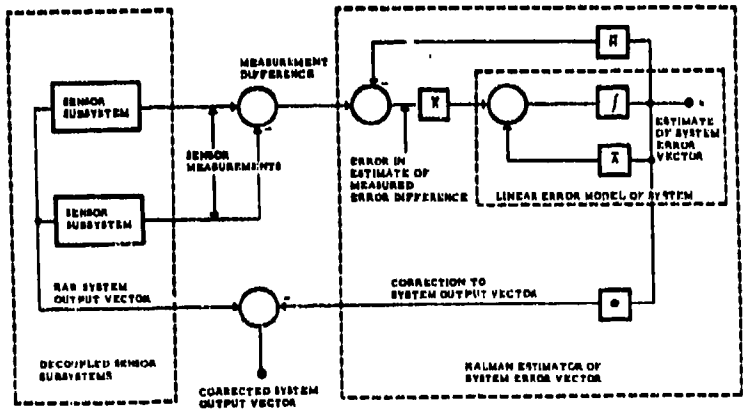


Fig.10 Schematic information flow for system using output correction

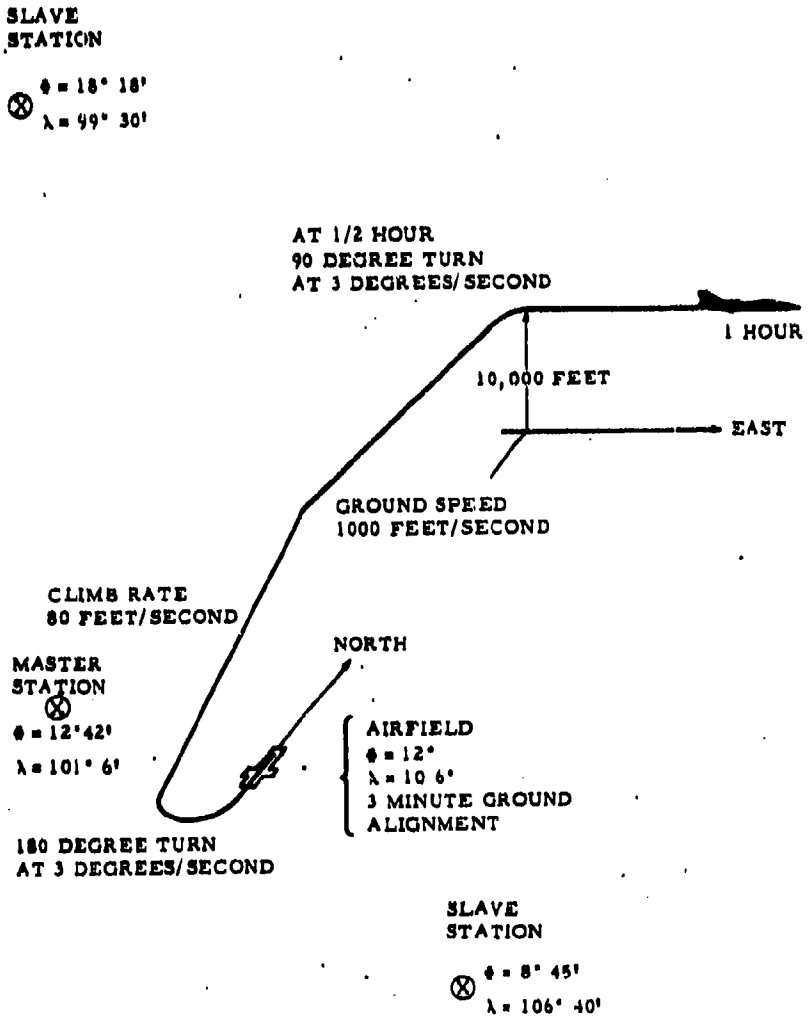


Fig.11 Schematic illustration of aircraft flight profile and LORAN transmitter locations

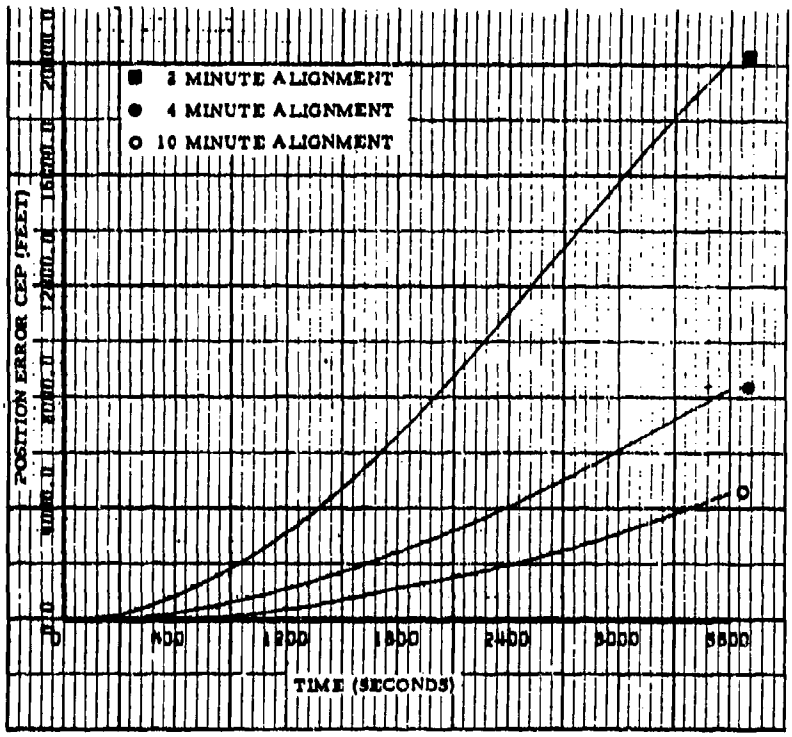


Fig.12 Position error CEP versus time. Free-inertial mode

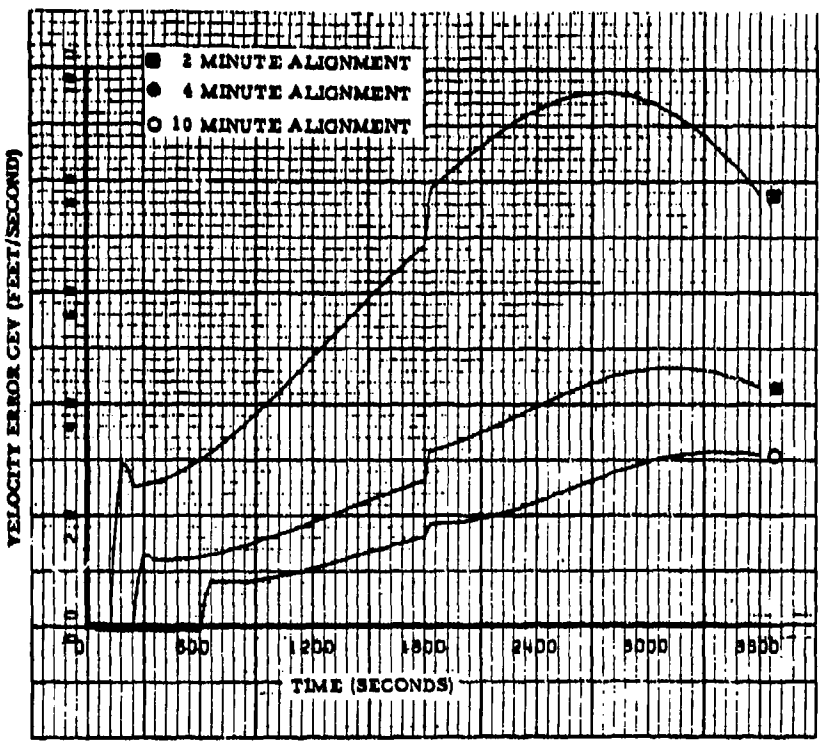


Fig.13 Velocity error CEV versus time. Free-inertial mode

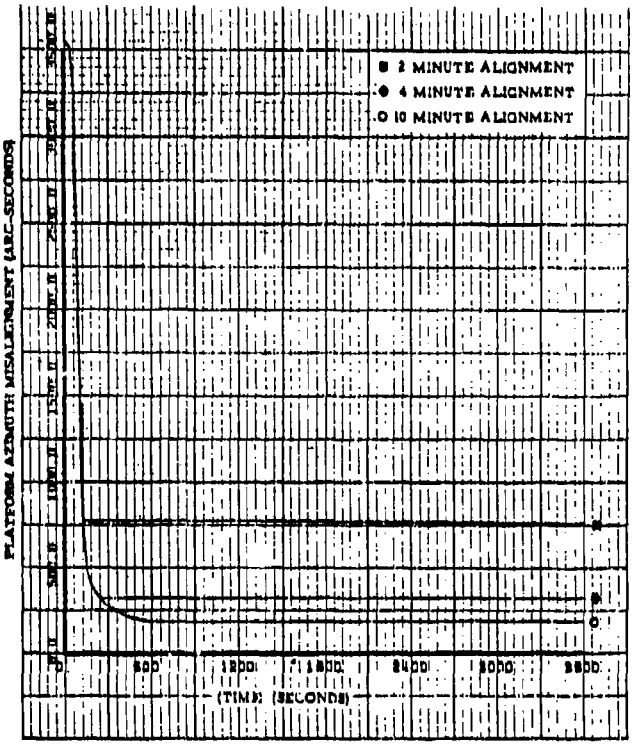


Fig.14 Platform azimuth misalignment versus time. Free-inertial mode

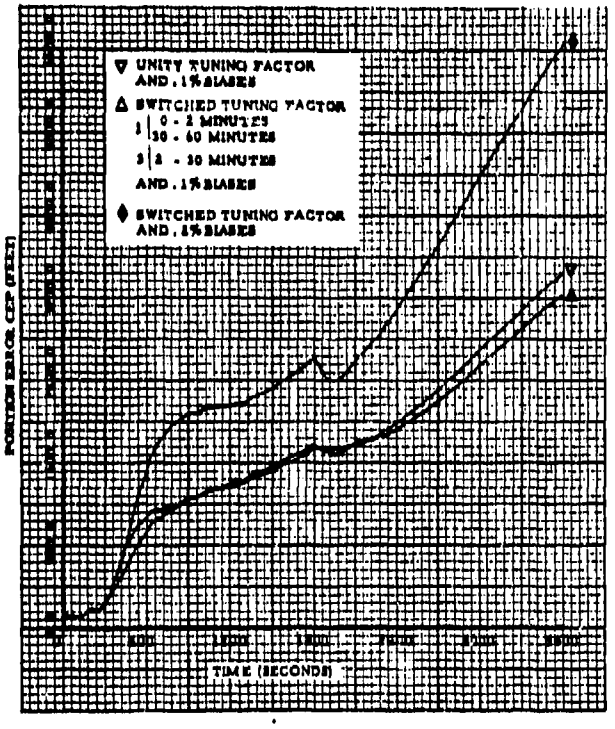


Fig.15 Position error CEP versus time. Doppler-inertial mode

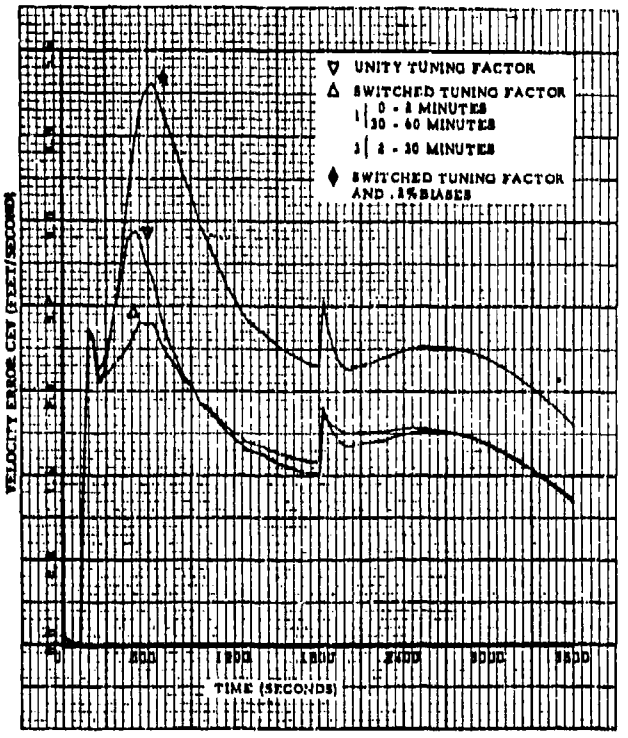


Fig.16 Velocity error CEV versus time. Doppler-inertial mode

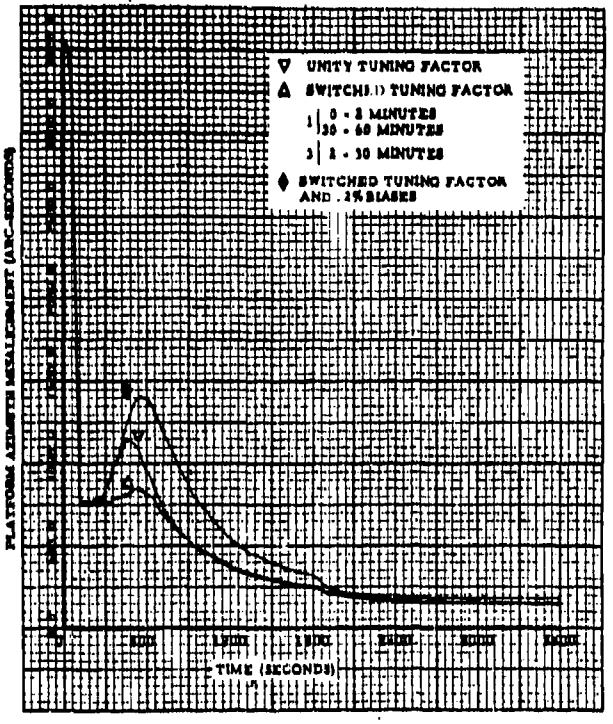


Fig 17 Platform azimuth misalignment versus time. Doppler-inertial mode

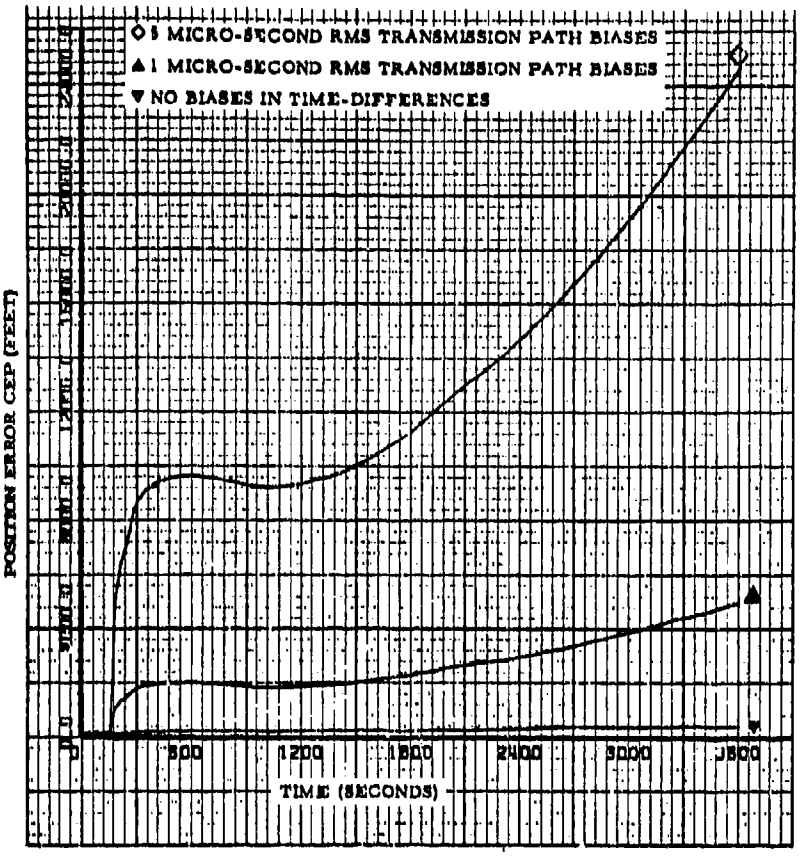


Fig.18 Position error CEP versus time. Loran-inertial mode

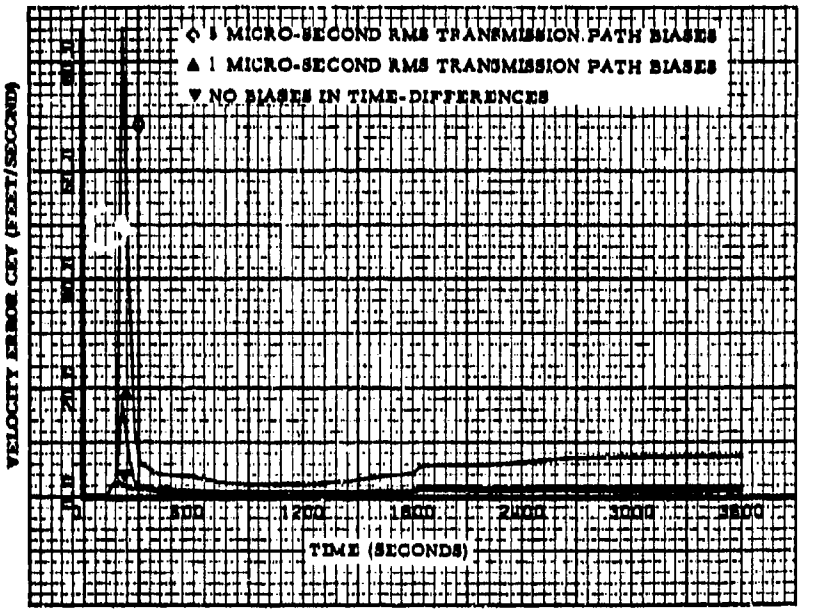


Fig.19 Velocity error CEV versus time. LORAN-inertial mode

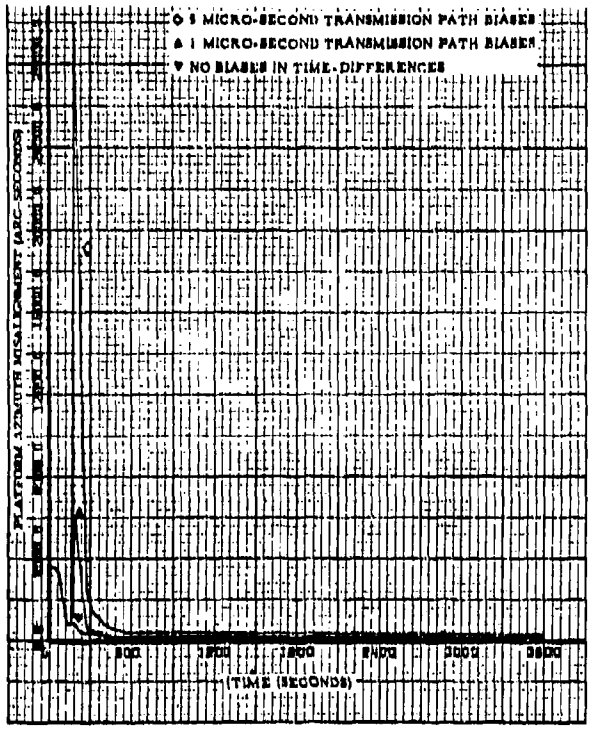


Fig.20 Platform azimuth misalignment versus time. LORAN-inertial mode

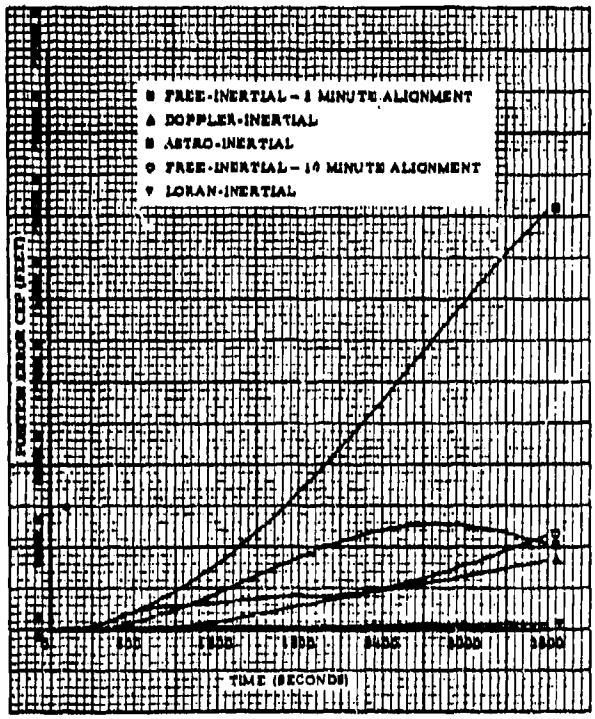


Fig.21 Position error versus time. All modes

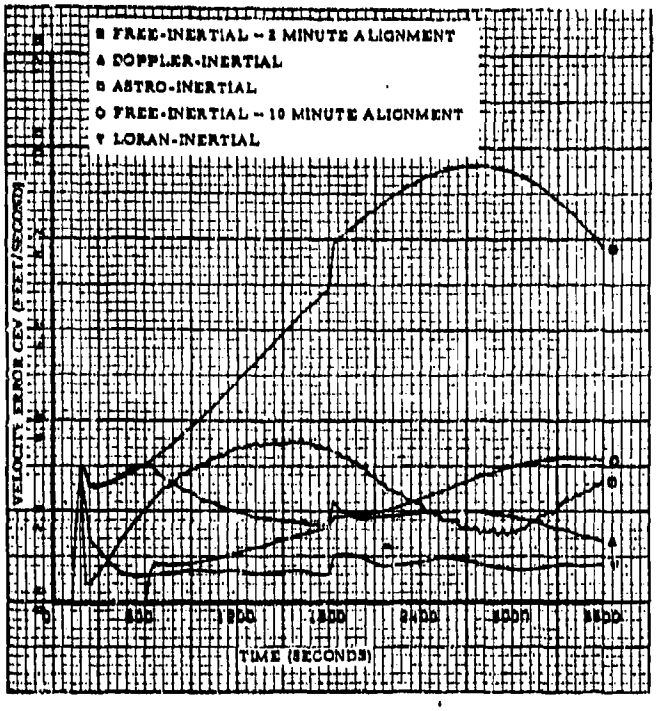


Fig.22 Velocity error CEV versus time. All modes

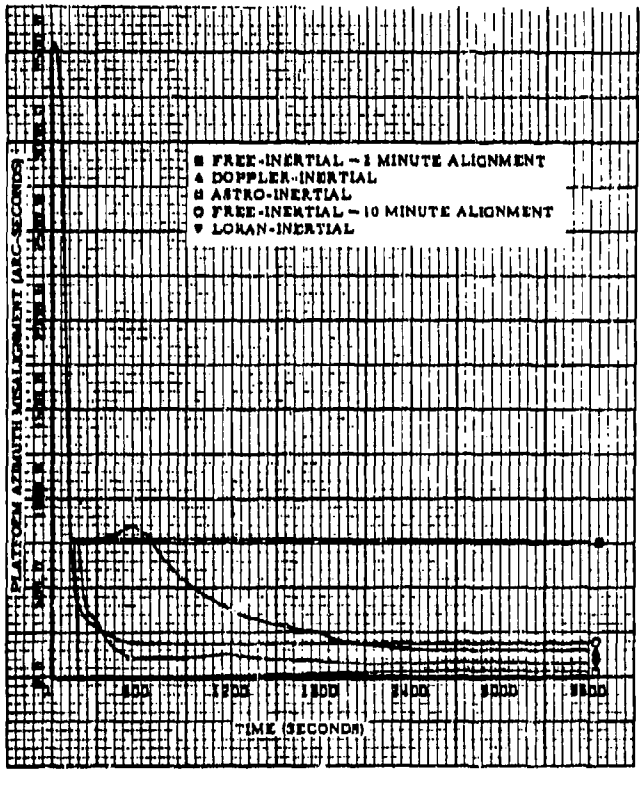


Fig.23 Platform azimuth misalignment versus time. All modes

CHAPTER 12 - APPLICATION OF KALMAN FILTERING TO BARO/INERTIAL HEIGHT SYSTEMS

by

P.M.Barham and P.Manville

Avionics Department, Royal Aircraft Establishment,
Farnborough, Hampshire, England.

CHAPTER 12 - APPLICATION OF KALMAN FILTERING TO BARO/INERTIAL HEIGHT SYSTEMS

P.M.Barham and P.Manville

1. INTRODUCTION

The aviator has had no alternative, until recently, to the use of barometric pressure as the source of his height and height rate information. Despite the ingenuity of the designer, which has resulted in the development of some extremely sensitive and accurate instruments, the fact remains that the use of pressure necessitates assumptions about the structure of the atmosphere in order to derive height or height rate. This has led to the concept of a standard atmosphere (e.g., ICAN). Errors can occur due to variations in the lapse rate, the height of the tropopause and in sea-level temperature and pressure. Even though a datum is established before take-off, this means that climbing or diving through a non-standard atmosphere will result in height rate errors of up to 8%, with corresponding errors in height. The two "actuals" shown in Figure 8 were recorded on the same day (1st March 1967), one over Cornwall and the second 600 nautical miles further north, over the Shetland Isles. In each case the pressure error is seen to correspond to a mean slope (of 0.055 and 0.015 respectively) with a random noise superimposed. With variations in the sign and magnitude of the mean slope, these are typical records. Effects such as lag and hysteresis make it very dangerous to rely on height rate accuracy at times when height rate itself is changing rapidly (such as during an overshoot when landing).

It was natural therefore that, with the advent of practical inertial platforms, the avionics system designer should have looked with interest to see what help he could get from the vertical inertial channel. Unlike the horizontal channels, which are comfortably bounded by the Schuler loop, the vertical channel is unstable, so that any accelerometer errors will cause the height error to increase as the square of time. Moreover the effect of gravity on the vertical accelerometer has to be cancelled. Since this correction is height-dependent, an error in assumed height will result in an acceleration error in the sense which will increase the original error.

However, it is easily seen that barometric and inertial height systems are largely complementary to each other. The former provides good height rate information in nearly level flight (subject only to the slope of the isobars in the vertical plane, which is typically 1 foot per mile standard deviation), but is poor in climbs or dives or in the presence of significant vertical acceleration. On the other hand the inertial system needs to be bounded by an external reference (such as barometric data) for sustained periods, but provides direct information about vertical acceleration and a good short-term reference for use during climbing or diving.

The use of such baro/inertial mixed systems is considered in more detail below.

2. TRADITIONAL BARO/INERTIAL MIXING

The operation of a Kalman filter in a baro/inertial height system cannot be fully appreciated without first considering the more traditional mixing systems and their limitations.

2.1 Fixed Gain Systems

The basic mixing loop is shown in Figure 1. There is no drift compensation for the inertial channel and, assuming that the system is trimmed before flight, the dominant error at high altitude would probably be poor "g" compensation due to the use of erroneous height. In a closed-loop climb inertial height (h_I) will be forced to follow changes in pressure height (h_p) with resultant height rate errors of up to 8%. On the other hand, during an open-loop climb the system would soon become unstable because of the uncompensated acceleration errors. The transient which would normally occur on re-closing the loop could be reduced by a "fast re-set" phase, in which the feedback through K_1 is temporarily increased.

Addition of drift compensation, as shown in Figure 2, would improve the open-loop climb as the inertial channel would now start in a well trimmed state. However, even with a fast re-set the system cannot settle without a lengthy transient on re-closing the loop as the act of forcing h_I to equal h_p will necessitate additional drift compensation to accommodate the new error in "g" correction.

Adding a pressure height correction loop, as shown in Figure 3, is a big improvement as it removes the transient on re-closing the inertial loop and is making an attempt to operate in "true" height. It operates in two different states based on the following assumptions:

(a) In level flight the barometric data gives perfect height rate (ignoring slope of isobars) and can be used to trim accurately the inertial channel.

(b) In a short period of climbing, the inertial channel may be considered perfect and used to determine the error in change of height as indicated by the barometric unit.

In practice some difficult compromises are necessary. Perfectly level flight can never be achieved and, to keep the inertial channel bounded, it is necessary to remain in the closed-loop state up to moderate height rates, in which case the pressure height rates can have significant errors. If, on the other hand, the switching takes place at lower height rates, the inertial channel could be left open-loop for excessive periods during slow climbs, with resultant instability.

2.2 Variable Gain Systems

The above considerations lead one to examine the possibility of using a variable gain system, as shown in Figure 4, rather than direct switching between the third and fourth integrators. Both K_3 and K_4 can be made height-rate-dependent. In level flight $K_3 = 1$ and $K_4 = 0$, so that the system becomes equivalent to the system in Figure 2. With increasing height rate K_4 will increase and K_3 will decrease. The error signal is now distributed between the inertial and barometric systems according to the values of K_3 and K_4. This is similar to the multi-dimensional correction to the state vector after a measurement with a Kalman filter, and the values of K_3 and K_4 are being used in a similar way to the appropriate terms in the gain matrix. Despite the increase of complexity the system in Figure 4 is still very imperfect compared with the Kalman filter described in Section 3.

With a Kalman filter the error between forecast and measurement results, in general, in corrections being fed back to all the elements in the state vector in accordance with the current values in the variance matrix. This is roughly analogous to K_1 and K_2 in Figure 4 also being made variable. Of greater importance is the fact that, with the Kalman filter, the terms in the variance matrix reflect the previous history of the system and are not simple functions of height rate. For example, if a prolonged period of level flight had resulted in a very well trimmed inertial system, and hence correspondingly small values of variances associated with it, almost the whole discrepancy arising at the start of a climb would be attributed to an abnormal atmosphere, regardless of height rate. Finally, the filter described below has a built-in capability to learn the structure of the atmosphere and make short-term use of this knowledge.

3. THE KALMAN FILTER HEIGHT SYSTEM

3.1 The State Vector

Since the purpose of the filter is to control the vertical channel of an inertial navigator it is obvious that the state vector should include terms representing the best estimates of the errors of inertial acceleration (ϵa_i), velocity (ϵv_i) and height (ϵh_i). Measurement will consist of a comparison of inertial and pressure height, so a term representing error in pressure height (ϵh_p) must be included. With the subscripts i, p and t representing inertial, pressure and true values respectively, these are defined as follows:

$$\epsilon a_i = a_t - a_i \tag{3.1}$$

$$\epsilon v_i = v_t - v_i \tag{3.2}$$

$$\epsilon h_i = h_t - h_i \tag{3.3}$$

$$\epsilon h_p = h_t - h_p \,. \tag{3.4}$$

To enable the filter to learn something of the departures from standard of the atmosphere, a further element (δ) is included in the state vector. This is defined by the relationship

$$\dot{h}_p = \dot{h}_t(1+\delta) \tag{3.5}$$

and is, in effect, the slope of the plot of pressure height error against true height. The complete state vector thus becomes

$$X = \begin{bmatrix} \epsilon h_i \\ \epsilon v_i \\ \epsilon a_i \\ \delta \\ \epsilon h_p \end{bmatrix}. \tag{3.6}$$

3.2 The Transition Matrix

Examination of the chosen state vector shows that the first three terms are concerned with the open-loop inertial system, and the last two terms with the atmospheric model. These two sets only interact as a result of the measurement process which, of course, builds up a strong correlation between inertial and pressure height estimates. However, in developing the transition matrix, which controls the updating process by the equations of motion, the two sets can be considered separately.

Now the classical equations of motion for a uniformly accelerating system without any feedback can be expressed as

$$\begin{bmatrix} \epsilon h_i \\ \epsilon v_i \\ \epsilon a_i \end{bmatrix}_{k+1} = \begin{bmatrix} 1 & t & \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon h_i \\ \epsilon v_i \\ \epsilon a_i \end{bmatrix}_k , \tag{3.7}$$

where t is the time interval between state k and state $(k+1)$, and must be chosen so that the acceleration error may be considered constant during one step. However, in the vertical channel of an inertial navigator there exists an additional feedback. A height error in calculating the value of "g" will give an acceleration error of 3×10^{-6} ft sec^{-2} for every foot of height error. If ϵa_o is the acceleration error from all causes other than erroneous "g" compensation (i.e. untrimmed accelerometer drift) which may be considered constant between two successive steps, then the total error is

$$\epsilon a_i = \epsilon a_o + (3 \times 10^{-6})\epsilon h_i .$$

Considering now the change in ϵa_i during one cycle, we have

$$\begin{aligned} \Delta(\epsilon a_i) &= 3 \times 10^{-6}\Delta(\epsilon h_i) \\ &= 3 \times 10^{-6}(t \epsilon v_i + \tfrac{1}{2}t^2 \epsilon a_i) \\ &= (3 \times 10^{-6}t)\epsilon v_i + (1.5 \times 10^{-6}t^2)\epsilon a_i . \end{aligned}$$

Hence

$$\begin{aligned} (\epsilon a_i)_{k+1} &= (\epsilon a_i)_k + \Delta(\epsilon a_i) \\ &= (3 \times 10^{-6}t)(\epsilon v_i)_k + [1 + (1.5 \times 10^{-6}t^2)](\epsilon a_i)_k . \end{aligned} \tag{3.8}$$

The transition matrix for the inertial channel thus becomes

$$\begin{bmatrix} 1 & t & \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 3 \times 10^{-6}t & [1 + (1.5 \times 10^{-6}t^2)] \end{bmatrix} . \tag{3.9}$$

Of the remaining terms in the state vector δ should ideally have both a time and a distance correlated exponential decay, since it is obvious that any data on atmospheric structure would become of little use after a prolonged period of time, or at a distant location. However, in view of the difficulty of modelling a correlation with distance it was decided to use only an exponential time correlation with an arbitrarily chosen 10,000 second time constant. This is represented as

$$\delta_{k+1} = (1 - 10^{-4}t)\delta_k . \tag{3.10}$$

From the definition of δ in Equation (3.5) and of ϵh_p in Equation (3.4), we have that

$$\begin{aligned} \Delta(\epsilon h_p) &= (\dot{h}_t - \dot{h}_p)t \\ &= -\dot{h}_t t\delta \\ &= -\Delta h\delta , \end{aligned}$$

where

$$\Delta h = \int_o^t \dot{h}_o \, dt , \tag{3.11}$$

$\dot{h}_o$ being the computed best estimate of height rate.

Thus

$$(\epsilon h_p)_{k+1} = (\epsilon h_p)_k - \Delta h\delta_k . \tag{3.12}$$

Combining (3.9), (3.10) and (3.12), the complete transition matrix for the state vector at (3.6) becomes

$$\Phi = \begin{bmatrix} 1 & t & \frac{1}{2}t^2 & 0 & 0 \\ 0 & 1 & t & 0 & 0 \\ 0 & 3 \times 10^{-6}t & [1 + (1.5 \times 10^{-6}t^2)] & 0 & 0 \\ 0 & 0 & 0 & (1 - 10^{-4}t) & 0 \\ 0 & 0 & 0 & -\Delta h & 1 \end{bmatrix} . \tag{3.13}$$

3.3 The Measurement Matrix

The measurement matrix required for baro/inertial mixing is

$$M = [1 \quad 0 \quad 0 \quad 0 \quad -1] \, . \tag{3.14}$$

This selects from the state vector the terms

$$\epsilon h_i - \epsilon h_p = (h_t - h_i) - (h_t - h_p)$$
$$= h_p - h_i \, .$$

3.4 Plant Noise

3.4.1 Acceleration

Plant noise on acceleration can arise from a number of causes, of which the following are of most significance:

(a) random changes in bias level, which typically occur every 10 minutes on average with changes with a standard deviation of about 10^{-5} g;

(b) g^2 distortion (with a force feedback accelerometer) with a typical level of 2×10^{-5} g/g^2 for a typical inertial quality accelerometer.

Considering each of these in turn, the random changes increase the variance by 10^{-10} g^2 every 10 minutes, or by 17×10^{-10} ft^2 sec^{-4} in one 10 second cycle.

The second cause is more difficult to assess. However, an examination of the acceleration power spectral densities for a typical strike aircraft enables us to derive some approximate figures. Assuming an anti-vibration mounting with a 10 c/s resonant frequency, and typical damping, the power spectral densities are first weighted by the square of the amplitude transmission curve for the mounting and then integrated over frequency to provide mean square acceleration power levels for typical flight conditions. These can be converted into equivalent bias levels by application of the assumed g^2 distortion coefficient of 2×10^{-5} g/g^2. It is found that, for the aircraft investigated, the highest level occurs while pulling high "g". This is followed in descending order by take-off and landing and then by high IAS dives at a factor of 20 down on the high "g" case.

We are not interested in the absolute level of vibration, or the resultant bias, but in the variance of the change of bias per computing cycle (10 seconds). Consider a simple example of a flight which has, on average, one high "g" manoeuvre every 10 minutes (with a bias level of 10^{-4} ft sec^{-2}), the rest of the flight being straight and level (negligible bias). This will result in two changes of bias level for each manoeuvre, or one change in every 30 computing cycles. This gives a variance of the change of acceleration bias per step of 3.3×10^{-10} ft^2 sec^{-4}, or about 20% of that due to random changes. Similar consideration of a high IAS dive occurring on average once every 5 minutes results in a negligible variance of 0.02×10^{-10} ft^2 sec^{-4}, which indicates that only the high "g" manoeuvres need be considered. Allowing a small increase in variance to account for other conditions, the plant noise consists of the addition each cycle of 21×10^{-10} ft^2 sec^{-4} to the variance of acceleration error.

3.4.2 Departure from Standard Atmosphere (δ)

It was assumed that, overall, the standard deviation of δ was 0.03, although there is some evidence that a higher value should be chosen. In prolonged level flight a plant noise contribution is required to maintain this level, despite the decay term introduced into the transition matrix. For a 10 second cycle this term is 0.999, which results in the variance p of δ varying as

$$p_{k+1} = (0.999)^2 p_k + n \, ,$$

where n is the plant noise contribution. If $p_{k+1} = p_k = (0.03)^2$ in the steady state, it is required that $n = 1.8 \times 10^{-6}$ during level flight.

During climbs or dives it is obvious that the probability of a change of slope of the atmospheric error is increased. Analysis of typical measured atmospheres (see Figure 5) indicates an additional contribution to the variance of δ of about $4.8 \times 10^{-6} \, |\dot{h}|$ for a 10 second cycle, where $|\dot{h}|$ is the modulus of the height rate in ft sec^{-1}.

The total plant noise to be added to the variance of δ each cycle is thus $(1.8 + 4.8|\dot{h}|) \times 10^{-6}$.

3.4.3 Pressure Height

On the assumption that all isobaric surfaces are horizontal planes, any uncertainty in pressure height must arise from a lack of knowledge of the true value of δ during a climb or dive. This is allowed for by the plant noise on δ described above. However, in practice the isobaric surfaces in any region will be slightly sloped (with a standard deviation of about 1 foot per nautical mile). The system has no means of determining such a slope, apart from making height fixes (which are discussed below). To implement such a system would require two extra terms in the state vector (northerly and easterly slopes) which would couple into the pressure height correction term through terms in the transition matrix involving displacements determined by the horizontal channels. This

would be directly equivalent to the coupling of δ into the height error by the Δh term. In practice such a refinement is hardly practical in the absence of a continuous external measurement, such as vertical Doppler, and would not be worth the extra complication. It must be appreciated, however, that without such terms the predicted height and height rate errors and their variances are relative to the isobaric surfaces and not to a horizontal datum. To allow for changes in slope, however, a small amount of plant noise is added directly to the pressure height. This amounts to the addition of 4×10^{-6} ft^2 to the variance of pressure height at each computing cycle.

3.5 Measurement Noise

The measurement noise is the inability of the system to determine the correct static pressure of the undisturbed atmosphere outside the aircraft. This is essentially peculiar to a particular system and would need to be carefully studied in each practical case. However, for the purposes of this study, some general observations were made for use in mathematical modelling. Sources of error considered included

(a) random instrument noise,

(b) hysteresis effects,

(c) instrumental lags,

(d) effects of wind gusts.

The first three errors were studied for a typical modern force re-balance air data system. The random instrument noise was found to have a standard deviation of about 0.11 millibar. Although relatively large under some conditions, hysteresis and lag effects were found to be very reproducible and, as such, could be allowed for by modelling in the digital computer of a practical system. In this case only the statistics of the error in modelling need be considered.

The actual pressure changes occurring due to gusts are, in general, small. Peak values of typically 1 millibar change in 1 kilometre can be recorded during thunderstorms and about an order less in ordinary "gusty" conditions. Hence the "all time, all weather" standard deviation must be very small. It should be noted, however, that a poor design of pitot-static will result in air velocity changes during gusts being recorded as pseudo-pressure-changes.

For the purposes of mathematical modelling the overall standard deviation of measurement noise was assumed to be 0.12 millibar for $|\dot{h}| < 25$ ft sec^{-1} and 0.32 millibar for $|\dot{h}| \geq 25$ ft sec^{-1}. These were converted to height errors by the usual relationship of $(0.6h^2 \times 10^{-7} + 27.8)$ feet per millibar. The two conditions were introduced on the grounds that lag effects are more significant during climbs and dives. In practice, of course, the measurement noise used could result from a study of the actual system.

3.6 Height Fixes

For a height fix using, for example, a radio altimeter, the required measurement matrix is

$$M = [1 \ 0 \ 0 \ 0 \ 0] .$$

This selects $\epsilon \, h_i$ from the state vector, which is compared with $(h_r + h_g - h_i)$, where h_r is the radio altimeter height and h_g is the height of the ground above datum. Since in the steady state the normal baro/inertial measurements result in a correlation between inertial and pressure height which approaches unity, the same correction will usually be made to both.

4. THE MATHEMATICAL MODEL

Orthodox modelling was used, making use of gaussian weighted random numbers of appropriate standard deviations to generate discrete errors for the "real world" model. One point of interest to this particular case is the model used for the atmosphere. Examination of a number of recorded atmosphere error plots, two of which are shown in Figure 5, had suggested a fairly well defined mean slope in each case, with random noise superimposed. The latter could be approximated by a series of straight line segments connecting points with a mean separation in height of about 5000 ft and with a standard deviation departure from the mean slope of 100 ft.

A series of random numbers, with rectangular distribution in the range 0 to 1, were drawn until one of them fell within the range 0 to 0.1. This was taken to represent a change point, each number in the series corresponding to a height change of 500 ft. Thus, if the seventh number drawn was the first to fall within the range 0 to 0.1, this represented a change point at 3500 ft. A gaussian weighted random number with standard deviation 100 ft was then drawn to represent the departure of this change point from the line of mean slope. A series of change points with mean separation 5000 ft were thus generated and were joined to provide a model of the atmosphere. The full lines in Figure 5 show typical models generated in this manner (with mean slopes of 0.05 and 0.15 respectively). It is seen that they compare well with the actual records.

Due to the much lower level and similar frequency of g^2 distortion changes compared with random accelerometer changes, it was not thought worthwhile to model both effects separately. Instead the frequency of random changes was slightly increased to compensate. Thus, on each cycle, a random number in the range 0 to 1 was drawn. If this fell within the range 0 to 0.02 a gaussian weighted random change of standard deviation 0.00032 ft sec^{-2} was added to the "real world" acceleration error.

5. RESULTS

Some typical results obtained with the mathematical model are shown in Figures 6 – 10. No results have been shown for short duration high rate climbs or dives, as the errors arising from these are negligible, due to the high degree of reliance placed on the inertial channel. It should also be appreciated that reversals of height rate also produce reversals of the "forcing function", that is the erroneous height rate derived from pressure, so that the Kalman filter is easily able to separate inertial and barometric errors. This is an important feature which, unlike the systems described in Section 2, enables the inertial channel to be trimmed accurately during typical terrain-following profiles. Theoretically the filter can distinguish between pressure and inertial errors whenever the height rate changes at all.

For this reason long climbs at low rates have been illustrated as representing the worst case for the filter. Figure 6 shows the height rate error and standard deviation for a climb to 60,000 ft at 150 ft sec^{-1} through an atmosphere with mean $\delta = 0.05$ after an initial settling period of level flight. A further five minutes of level flight is followed by a dive at the same rate, levelling out at ground level. The standard deviation just exceeds 0.2 ft sec^{-1} at the end of the descent, but rapidly drops to a negligible amount on regaining level flight. The actual error for this run is within the standard deviation curve and shows no systematic trends. The corresponding height error with its standard deviation is shown in Figure 7 and reveals a similar pattern, the standard deviation reaching about 80 ft at the end of the descent. This height rate is of considerable interest as it represents an especially difficult one for traditional height systems.

As the worst case for the Kalman filter height unit is a very prolonged climb at a constant rate, a similar run is illustrated in Figures 8 and 9, with a height rate of only 50 ft sec^{-1}, so that the ascent to 60,000 ft and the subsequent descent each lasted for 20 minutes. This results in a considerable increase in the variances of the inertial system so that the effect of the mean value of δ (0.05) shows as a definite systematic error build-up on both velocity (Fig.8) and height (Fig.9). It demonstrates a very important feature of the Kalman filter. During the climb the erroneous pressure height changes were able to affect the inertial system as the reliance that could be placed on the latter decreased. However, the Kalman filter builds up in the variance matrix a complete set of correlations in the off-diagonal terms. Thus, on regaining level flight, although no measurement of absolute height is made, the filter is not only able to correct the inertial velocity error, but is also able to correct the height error which had arisen as a result of this velocity error.

Figure 10 shows the Kalman filter estimate of δ, with the true value for comparison, during the run corresponding to Figures 6 and 7.

6. CONCLUSIONS

The Kalman filter can provide a useful tool for the optimal mixing of baro/inertial height information. For any practical system it is essential that the form of the plant and measurement noise should be carefully studied and modelled as accurately as possible. It is not claimed that the system described above is optimised, but it represents a useful starting point for the study of specific aircraft systems.

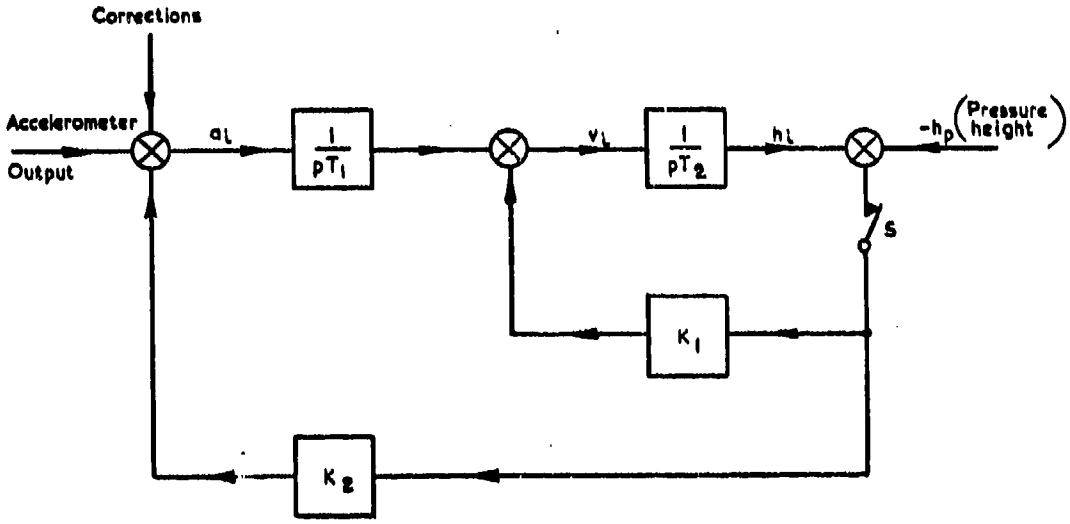


Fig.1 Basic baro/inertial mixing system

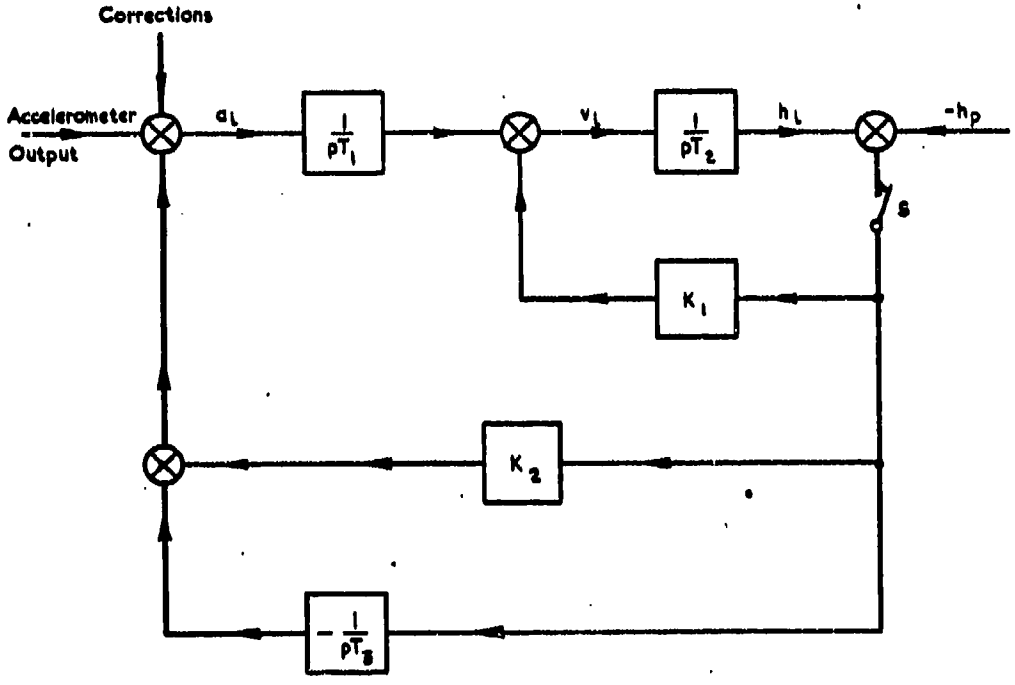


Fig.2 System with drift compensation

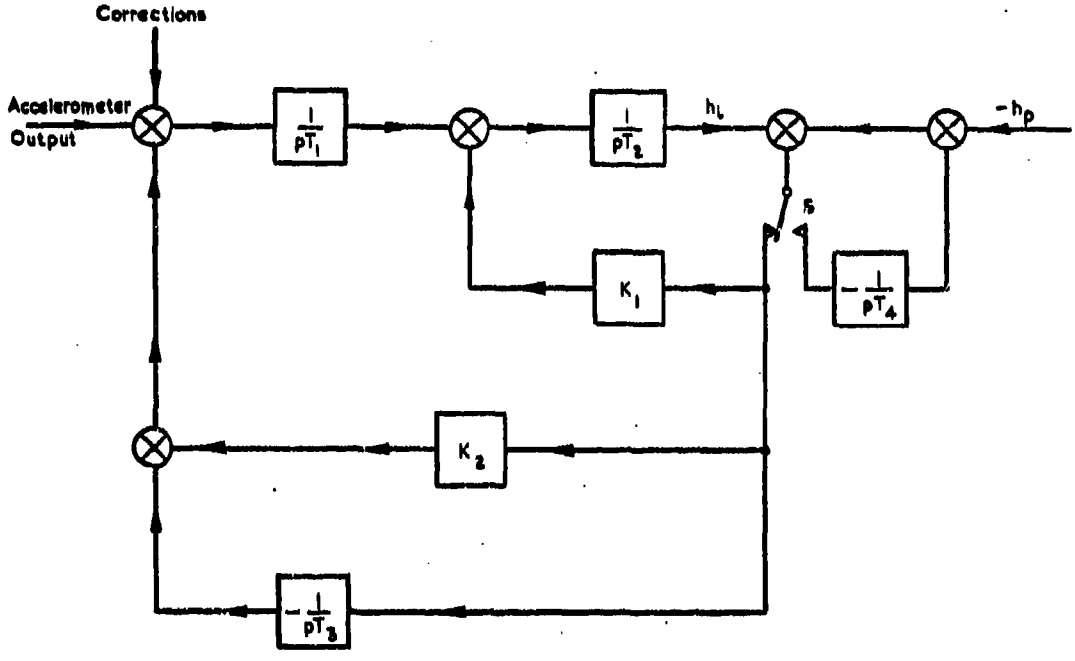


Fig.3 System with barometric height correction

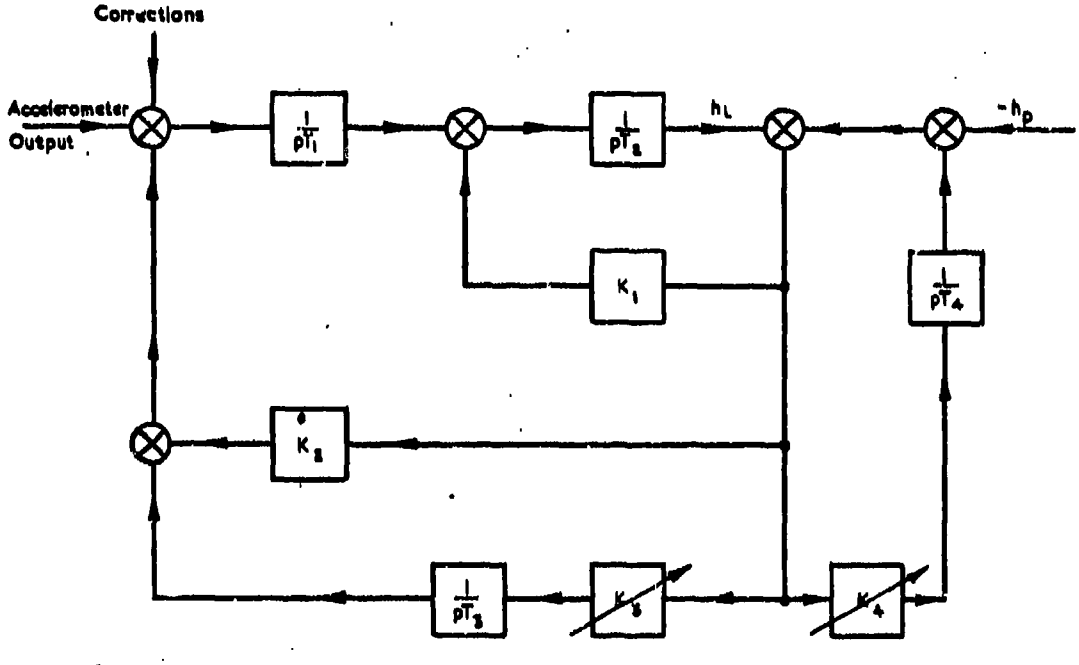


Fig.4 Variable gain system

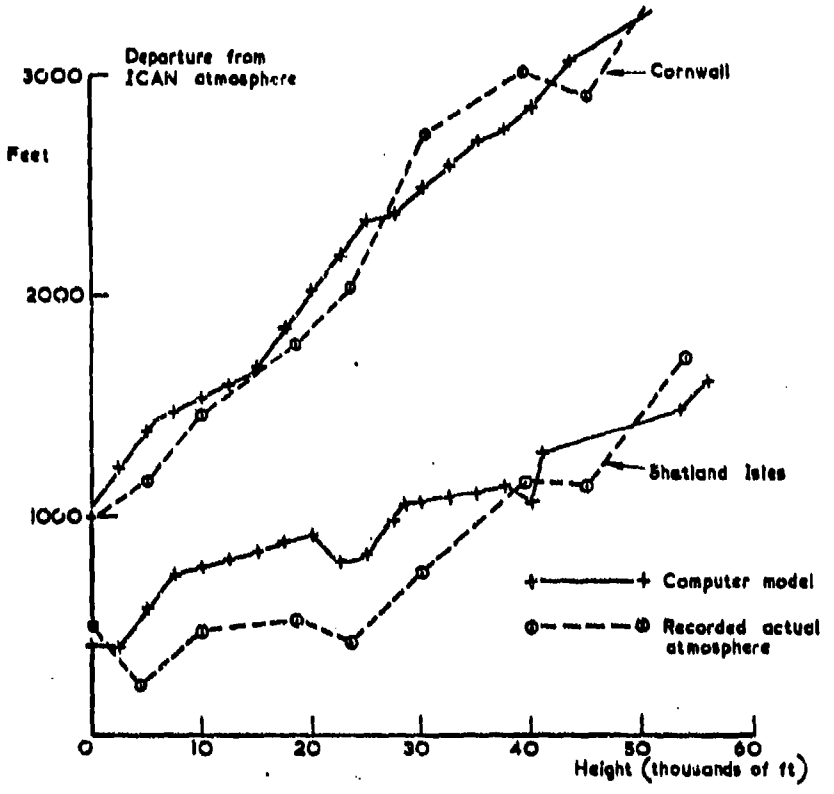


Fig. 5 Atmospheric model

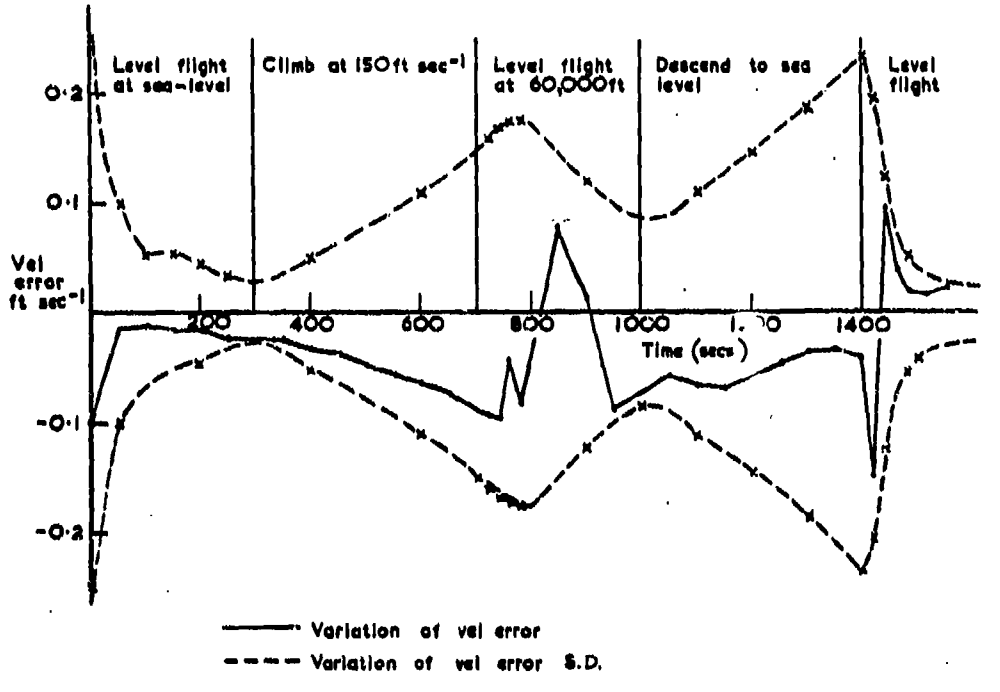


Fig. 6 Velocity error (150 ft sec^{-1})

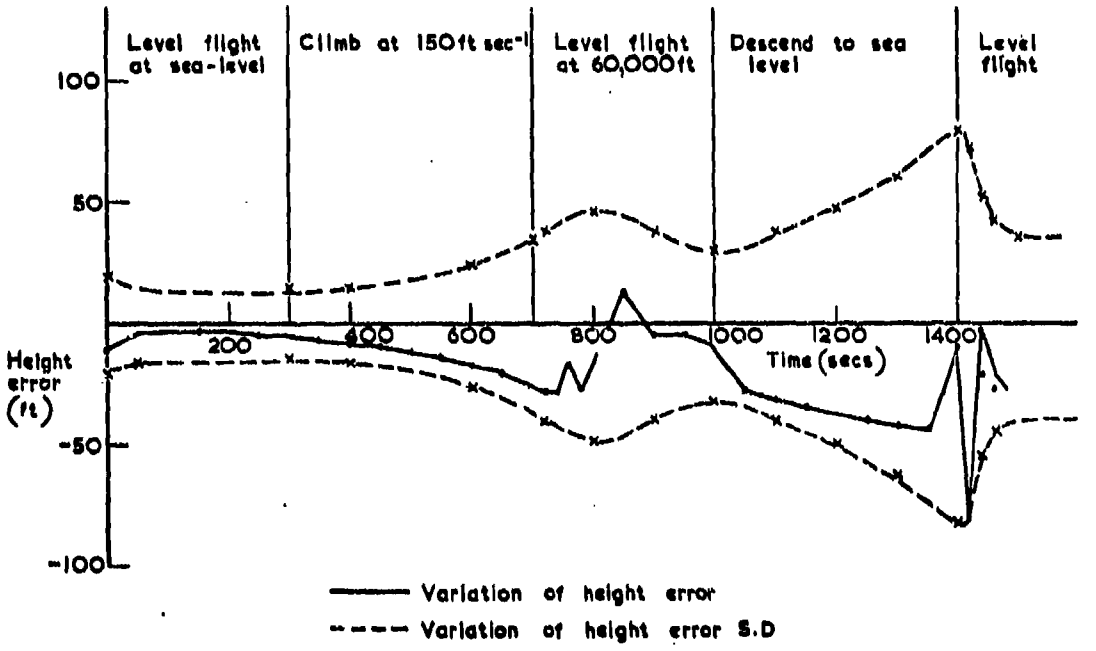


Fig.7 Height error (150 ft sec^{-1})

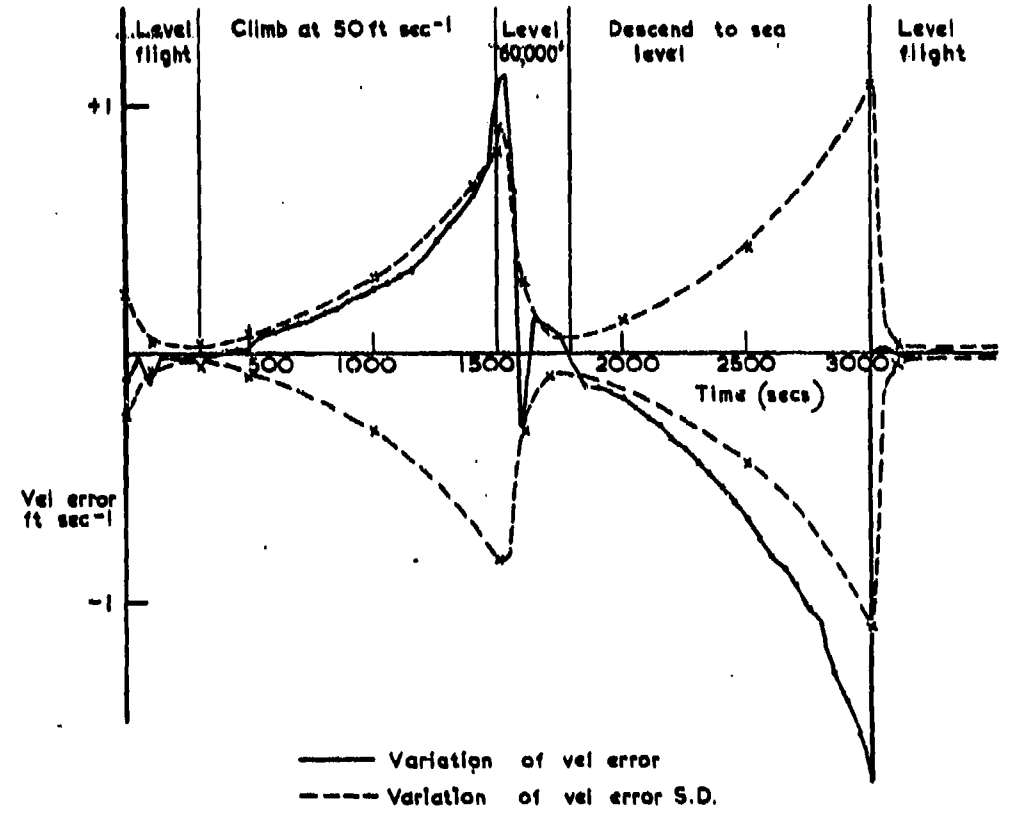


Fig.8 Velocity error (50 ft sec^{-1})

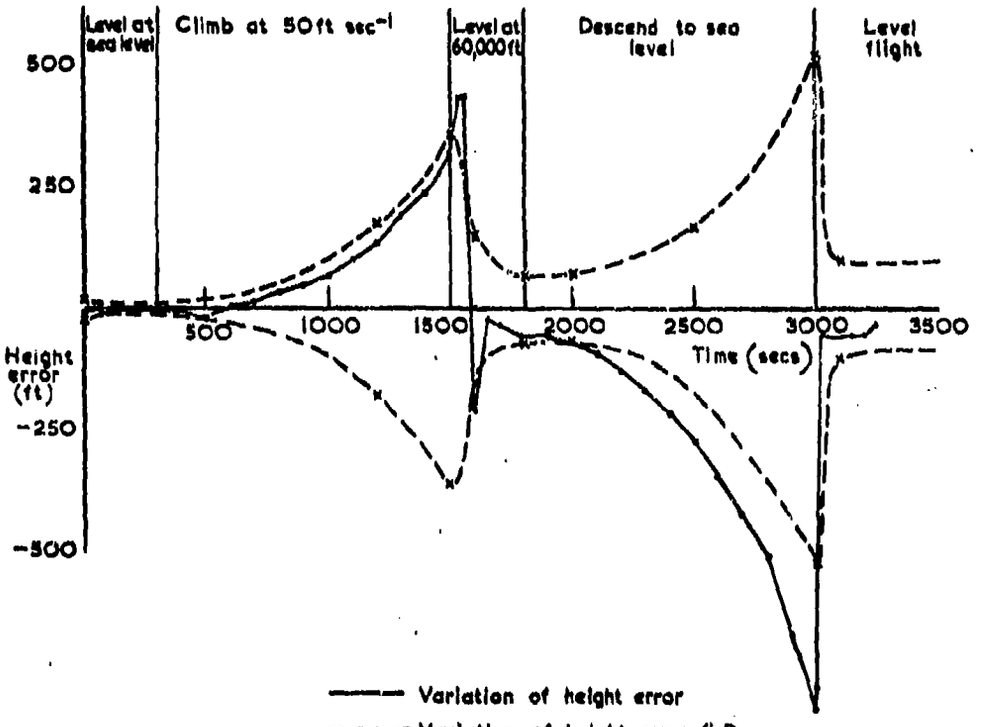


Fig.9 Height error (50 ft sec^{-1})

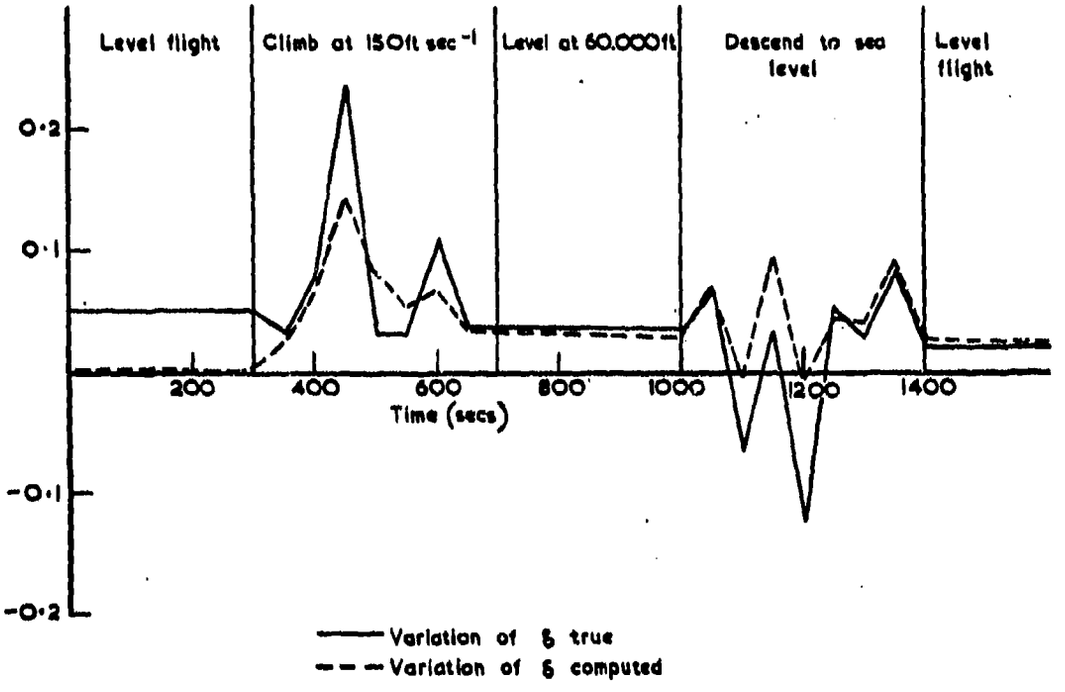


Figure 10

CHAPTER 13 - APPLICATION OF KALMAN FILTERING TO THE C-5 GUIDANCE AND CONTROL SYSTEM

by

Stanley F. Schmidt*, John D. Weinberg† and John S. Lukesh†

* Analytical Mechanics Associates, Inc.
Palo Alto, California, USA

† Northrop Corporation, Electronics Division
Hawthorne, California, USA

PRECEDING PAGE BLANK

NOTATION

Coordinate frames

EFF	orthogonal, earth-fixed reference frame
IF	orthogonal, wander azimuth, local vertical, ideal reference frame
CF	orthogonal, wander azimuth, local vertical, computational frame
PRF	orthogonal platform reference frame
AIF	non-orthogonal, accelerometer input axis frame
GIF	non-orthogonal, gyro input axis frame
OFi	orthogonal, i^{th} gyro input-output-spin axis frame

Operational and structural symbols

$\rightarrow$	(superior arrow) three-dimensional vector
$\cdot$	(superior dot) time differentiation
$\frac{d}{dt}$	time differentiation
T	(superscript) matrix transpose
-1	(superscript) matrix inverse
∇	vector gradient
$\simeq$	approximately equal
$\mathcal{L}^{-1}$	inverse Laplace transformation
s	Laplace variable
E[]	ensemble expectation operation

Navigational symbols

$\vec{f}$	specific force acting on ball
$\vec{f}_I$	$\vec{f}$ in IF
$\vec{f}_p$	$\vec{f}$ in PRF
$\vec{f}_A$	$\vec{f}$ in AIF
$\vec{f}_{Gi}$	$\vec{f}$ in GFi
$\vec{A}_s$	estimate of $\vec{f}_I$ in CF
$\vec{f}_o$	accelerometer outputs in AIF
$\vec{a}$	inertial acceleration of ball
$\vec{G}$	gravitational force (per unit mass) acting on ball
$\vec{A}$	aerodynamic force (per unit mass) acting on aircraft
$\vec{T}$	thrust force (per unit mass) acting on aircraft
$\vec{v}$	Doppler velocity vector
g	plumb-bob gravity
$\vec{A}_c$	Coriolis and centripetal acceleration estimate in CF

$\vec{\Omega}$	earth rate in CF
$\vec{V}_L$	aircraft level velocity with respect to earth in CF
$\vec{\omega}$	aircraft rate in CF
λ	aircraft latitude (computed)
L	aircraft longitude (computed)
θ_{pl}	platform heading (computed)
θ_{BR}	ball azimuth readout
θ_H	platform heading (indicated)
V_{GS}	aircraft groundspeed (computed)
θ_{GT}	aircraft ground track angle (computed)
C	earth's curvature matrix (computed)
W	angular rate matrix (computed)
h	aircraft altitude (computed)
V_I	aircraft vertical velocity (computed)
$\vec{\omega}_c$	commanded (uncompensated) gyro torquing rates in CF
$\vec{\omega}_d$	commanded (compensated) gyro torquing rates in CF
$\vec{\omega}_d$	gyro drift rates in GIF
$\vec{\omega}_p$	total inertial angular rate of platform in PRF
$T_{PI:P}$	IF to PRF transformation matrix
$T_{P:A}$	PRF to AIF transformation matrix
$T_{GI:P}$	GIF to PRF transformation matrix
$T_{P:GI}$	PRF to GFi transformation matrix
T	EFF to CF transformation matrix
$\vec{a}_b$	accelerometer bias errors in AIF
$\vec{a}_n$	accelerometer noise errors in AIF
$\vec{a}_{bc}$	accelerometer bias error calibrations in AIF
$[a_{ij}]$	accelerometer scale factor and input axis orthogonality error calibration matrix
$\vec{g}_{bc}$	gyro fixed drift rate calibration in GIF
$[g_{ij}]$	gyro scale factor and input axis orthogonality error calibration matrix
$[u_{ij}]$	gyro mass unbalance error calibration matrix
SFG	diagonal gyro scale factor error matrix
SFA	diagonal accelerometer scale factor error matrix
G_i	simulator error generating function for i^{th} gyro

Kalman filter symbols

Y	observation vector
X	augmented navigation state vector

$\tilde{x}$	augmented navigation error state vector
$\hat{x}$	estimated augmented navigation error state vector
x	error in $\hat{x}$
Φ	transition matrix for $\hat{x}$
P	navigation system error covariance matrix
R	navigation system forcing error noise matrix
K_i	Kalman gain vector, $= b$
K_ϵ	epsilon gain vector
Δy	observation residual
M	observation matrix (discrete observation)
ϵ	"epsilon" gain scalar
C	observation noise scalar
t_0	start of Kalman cycle
t_a	end of Kalman cycle
Δt	Kalman cycle time
t_i	time within a Kalman cycle
n	dimension of $\hat{x}$
u_{sc}	system control vector
$\hat{f}(\hat{x})$	estimate control algorithm
$\hat{g}(\hat{x})$	system control algorithm

Other symbols

Z	navigation system state vector
U	navigation system forcing function vector
p	vector defined by Equation (2.1), as indicated in text
H	navigation system description function
V	vector defined by Equation (2.2), as indicated in text
q	vector of random time-uncorrelated errors in the navigation observation
S	function defined by Equation (2.2)
F	augmented navigation system description function
u	vector of augmented navigation system white noise forcing functions
Q	functional, time-dependent relationship between navigation observations and augmented navigation system state
$\vec{R}$	aircraft position
Z	simulator true position and velocity vector
e	vector of principal augmented navigation system errors
e_f	vector of forcing augmented navigation system errors
α	matrix of linearized $\dot{e}$ to e coefficients

CHAPTER 13 - APPLICATION OF KALMAN FILTERING TO THE C-5 GUIDANCE AND CONTROL SYSTEM

Stanley F. Schmidt, John D. Weinberg and John S. Lukesh

1. INTRODUCTION

The Inertial-Doppler Navigation System developed by Northrop for the USAF/Lockheed C-5 heavy logistics transport[1] is designed to provide an accuracy sufficient for dropping supplies and/or personnel at desired locations, in any kind of weather, anywhere on earth. The requirements of high accuracy and all-weather capability are satisfied by combining information from self-contained and ground-based navigation aids. The measurements from the navigation aids are processed by the on-board navigational computer which uses algorithms derived from Kalman filter theory[2].

This chapter describes the development of the Kalman filter as utilized in the on-board computer of the C-5 navigation system. The Kalman filter development in itself represents only a small part of the overall system development. To aid in an understanding of the practical problems associated with application of the filter theory, considerable relevant information on the overall navigation system has been included. The discussion proceeds in substantially the same order in which the actual mechanization algorithms were developed: (a) System Description, Problem Definition and Approach; (b) Design Development and Description; and (c) Design Evaluation and Refinement.

Section 2 discusses the conceptual design of the overall navigation system, describes the system components, defines the system's functional capabilities and accuracy requirements, and discusses the development approach to the Kalman filter.

Section 3 formulates the mathematical bases for the design and for the error evaluation programs. There follows a discussion of (a) the major constraints that were imposed on the design by the hardware implementation, (b) the alternate design solutions that were considered and evaluated, and (c) the final solutions adopted. At the completion of this phase, design definition was sufficiently complete for the detailed real-time programming of the computer to begin.

The "Design Evaluation and Refinement" phase (Section 4) required a detailed simulation program to validate the design and to develop refinements in problem areas. Because extensive system laboratory tests and some system flight tests had already been conducted at the time of writing, it has been possible to include in this section some comparative results between simulated and real system behavior. These comparisons demonstrate the performance capability of the design, and also show how the simulator was used to locate and correct trouble areas.

Section 5 is a brief summary of the overall C-5 Kalman navigation filter development effort described in the earlier sections.

Finally, Section 6 takes a brief glance at future trends in the area of Kalman filtering application to real-time, digital computer implementation of augmented inertial and other types of navigational systems.

2. SYSTEM DESCRIPTION, PROBLEM DEFINITION AND APPROACH

Development of the Kalman filter for the C-5 navigation system started with the following areas already more or less completely defined:

(i) The functional design of the overall navigation system.

(ii) The functional capabilities and accuracy goals necessary to satisfy the various C-5 mission requirements.

(iii) The detailed specifications on the performance characteristics of the various subsystems.

The central objective was to develop the algorithms to be used in the on-board computer which would enable the achievement of these functional and accuracy capabilities. Pursuit of this objective involved much attention to mathematical error models - both functional and statistical - of all of the subsystems. The constraints of on-board computer memory, word length, and speed were also of fundamental importance.

This section presents: (a) a few fundamental concepts on inertial systems and the Kalman filter; (b) a description of the overall system configuration, sufficiently detailed for the reader to understand the role of the Kalman filter; (c) a description of the overall system design requirements and goals; and (d) a definition of the problem and a description of the approach adopted in the Kalman filter development.

pp. 287 & 288. check

2.1 Fundamental Concepts

As will be seen, the C-5 navigation system in its primary modes of operation uses an inertial measurement unit to maintain a continuous knowledge of the position and velocity of the aircraft. It will also be seen that the Kalman filter is used to update the inertial navigation system when data from other navigation aids is available.

The following is a brief description of the fundamentals of inertial navigation systems and the practical application of Kalman filter theory to them.

2.1.1 *Inertial Navigation Systems**

Figure 1 is a highly simplified block diagram illustrating the fundamental operation of an inertial navigation system. The principal components, the inertial measurement unit (IMU) and the computer are shown.

The IMU contains a stable platform on which three gyros and three linear accelerometers are orthogonally mounted. As implied by its name, the stable platform maintains an orientation, commanded by the computer, which is independent of the orientation of the case. This is accomplished, as indicated in Figure 1, by sensing error signals from the gyro output axis pick-offs and, through feedback, by providing the appropriate platform drive nulling torques. This "inner loop" feedback control system must have sufficient gain and dynamic response to keep the pick-off errors small, regardless of IMU case attitude motions.

If the inner loop functions properly, then the platform maintains the orientation dictated to the gyros by the computer. Each gyro, however, is not a perfect device, of course, and tends to exhibit a small "drift rate". That is, even when the gyro output axis pick-off is nulled, the gyro case (fixed to the platform) has a rotation rate about the input axis direction which differs slightly from the commanded rate.

These drift rates, and the unwanted output axis† torques which attend them, arise because of internal rotational unbalances in the gyro about its output axis, gyro output axis torquer scale factor errors, and from several other, more subtle, mechanisms. If the computer could somehow estimate the unwanted output axis torques, then their effect on platform orientation could be removed by applying appropriate angular rate compensating signals to the gyro output axis torquers.

As indicated in Figure 1, the accelerometer outputs are the measured specific force which is sent to the digital computer in the form of pulse trains. The specific force vector, $\vec{f}$, is given as

$$\vec{f} = -\vec{a} + \vec{G},$$

where $\vec{a}$ is the inertial acceleration vector and $\vec{G}$ is the gravitational attraction per unit mass. If the position is known, $\vec{G}$ can be calculated.

Hence, it can be seen that the navigation equations perform the functions of accepting specific force and performing the other computations necessary so that position and velocity are continuously available.

To initialize operation, the system must be aligned so that the coordinate frame of the stable platform and computer are coincident. The initial position and velocity must also be known.

The inertial navigation equations used in the C-5 are referred to as a wander azimuth geodetic vertical mechanization. This means that: (a) the coordinate frame has two axes in the locally-level tangent plane, and (b) no attempt is made to maintain the level axes in a preferred azimuth direction (such as north). The platform is initially aligned by driving it with the torquers until the outputs of the two level accelerometers are zero. Once it has thus been leveled, its azimuth relative to north is determined by monitoring the level gyro command rates, which are necessary to keep it level despite the rotation of the earth. The arc tangent of the ratio of these two rates then provides an estimate of the platform heading relative to north. This operation is called self-gyrocompassing. During subsequent operation, position and velocity data are computed and applied through the gyro torquers to maintain the platform locally-level despite aircraft motion across the surface of the Earth.

2.1.2 *Kalman Filter Theory*

Because of the various error sources in gyros, accelerometers, the initialization process, and navigation equation approximations, the position and velocity indicated by the inertial navigation system is in error. The error also grows with time after initialization.

Data which are functionally related to the true aircraft position and velocity can also be obtained from many other sources. As examples, barometric pressure provides an indication of altitude, Doppler radar provides

---

* Reference 3 is highly recommended for those desiring further information.

† The output axis of a single-degree-of-freedom gyroscope is that axis (normal to the spin axis) about which the spinning rotor is free to rotate with respect to the gyro case. The input axis (normal to the spin and output axes) is therefore the direction about which the gyro is sensitive to rotational case motion with respect to inertial space.

a measure of ground speed, and so forth. Conceptually, it is desired to use this external data to update the inertial navigation system; unfortunately these data, too, are imperfect, because of measurement errors and other effects (e.g., the pressure-altitude relationship is non-constant).

Kalman filter theory provides a means by which algorithms can be derived for digital computer processing of such measurement data observations*. The resultant algorithms are feasible for use with on-board computers in that there is no requirement for storage of large quantities of data; that is, each observation is processed sequentially as it occurs in time, and after it is processed the raw data can be overwritten with new data. As originally presented, the Kalman algorithms were restricted to:

(a) Linear dynamic systems.

(b) Observations linearly related to the state vector.

(c) Observations occurring at equally spaced time intervals.

The practical† application of the theory requires removal of these three restrictions. Removal of the third restriction is trivial and is readily accomplished by stating the original algorithm given by Kalman in the two parts:

(1) at an observation;

(2) between successive observations.

Theoretical proofs which remove the first two restrictions for nonlinear systems are not available. Practically, however, one should recognize that almost all real systems obey approximately linear relationships in the vicinity of a given solution. As a result, a practical filter for nonlinear systems can be derived by appropriate use of variational equations about a solution which is "close to" the desired solution. The desired solution in the case at hand would be the true value of the navigation state vector (aircraft position, velocity, etc.) which the observations are used to estimate. If this solution were known there would be no need for the filter. Hence, since it is not, the most obvious second choice is linearization about the best estimate of the navigation state vector, which is indeed available. If the solution converges as more observations are taken, the linearization approximation becomes more and more valid.

The algorithms are presented here in one of the many forms which are applicable to a general nonlinear problem**. Following this, the overall implementation of the filter in the C-5 navigation system is discussed.

Given a system which obeys the differential equations

$$\dot{Z} = H(Z;U,p,t) , \tag{2.1}$$

where

Z = normally defined state vector;

U = vector of forcing functions; U may be composed of known functions of time and/or random functions of time;

p = vector of constants whose values are required to give a unique set of Equations (2.1).

The vector observations of this system obey

$$Y(t) = S(Z,V,t) + q(t) , \tag{2.2}$$

where

V = a vector of deterministic quantities (e.g., constants) necessary to uniquely define the equations;

$q(t)$ = random, time-uncorrelated errors of observation††.

An augmented state vector X can now be constructed, viz.:

$$X = \begin{pmatrix} Z \\ U \\ p \\ V \end{pmatrix}$$

such that***

$$\dot{X} = F(X,t) + u(t) . \tag{2.3}$$

---

* The word observation is chosen here to indicate a measurement used in the sense of Reference 2. This distinction is made because not all measurements are processed in this sense, as will be seen later.

† The first known study results of a practical application of the filter are given in Reference 4.

** Several forms are presented in Reference 5.

†† Any time-correlated measurement errors can be defined as components of V. This also requires definition of differential equations for these components with white noise type driving functions.

*** These techniques are presented in more detail in Reference 5, although the notation there is somewhat different.

where

$u(t)$ = white noise (random forcing functions).

Also,

$$Y(t) = G(X,t) + q(t) . \tag{2.4}$$

2.1.3 Modified Kalman Filter Equations

The modified Kalman filter for such a system and its observations is summarized by Equations (2.5) – (2.11), for which the following definitions and assumptions apply:

(a) $\nabla_X()\big|_{X=\hat{X}}$ denotes the gradient of the bracketed function evaluated at $X = \hat{X}$.

(b) t_m denotes the observation time, t_0 denotes the time of the last change in the estimated state (which coincides with the time of the last observation).

(c) The subscript a denotes "after the observation is included" and the subscript b denotes "before the observation is included". Hence, from Equation (2.7), at an observation time, $P_b = P(t_m)$, and P_a in Equation (2.9) becomes $P(t_0)$ in Equation (2.7) when updating to the next observation time. Further comments (regarding multiple observations occurring at the same time) are included in the subsequent discussion.

Between observations (time update)

$$\hat{X}(t) = \hat{X}(t_0) + \int_{t_0}^{t} F(\hat{X},\tau)\, d\tau \tag{2.5}$$

$$\frac{d\Phi}{dt} = \nabla_X[F(X,t)]\Big|_{X(t)=\hat{X}(t)} \Phi(t;t_0)\Phi(t_0;t_0) = I \tag{2.6}$$

At observation time, t_m (observation update)

$$P(t_m) = \Phi(t_m;t_0)P(t_0)\Phi^T(t_m;t_0) + R(t_m;t_0) \tag{2.7}$$

$$\hat{x}(t_m) = (K_1 + K_2)[Y(t_m) - \hat{Y}(t_m)] \tag{2.8}$$

$$P_a(t_m) = P_b - K_1MP_b + K_2(MP_bM^T + C)K_2^T \tag{2.9}$$

$$\hat{x}(t_m+\tau) = \Phi(t_m+\tau;t_m)\hat{x}(t_m) \tag{2.10}$$

$$\hat{X}_a(t_m+\tau) = \hat{X}_b(t_m+\tau) + \hat{x}(t_m+\tau) . \tag{2.11}$$

Equation (2.5) simply represents an integration of the differential equations,

$$\dot{\hat{X}} = F(\hat{X},t) ,$$

using $\hat{X}(t_0)$ as initial conditions. $\hat{X}(t)$ is the best estimate of $X(t)$, because the random function $u(t)$ of Equation (2.3) is a zero mean white noise variable.

Equation (2.6) defines the differential equations for the transition matrix, which relates small deviations in $\hat{X}(t_0)$ to small deviations in $\hat{X}(t)$, that is,

$$\Phi(t;t_0) = \frac{\partial\hat{X}(t)}{\partial\hat{X}(t_0)} .$$

By defining

$$\tilde{X} = X - \hat{X} ,$$

then, approximately,

$$\tilde{x}(t) = \Phi(t;t_0)\tilde{x}(t_0) + \int_{t_0}^{t} \Phi(t;\tau)u(\tau)\, d\tau .$$

Hence, defining the covariance matrix $P(t_0)$ by

$$P(t_0) = E[\tilde{x}(t_0)\tilde{x}^T(t_0)] ,$$

where $E[]$ means the expected value of the bracketed quantity, and using the white noise definition of the random variables, u, it follows that:

$$P(t_m) = E[\tilde{x}(t_m)\tilde{x}^T(t_m)] = \text{Equation (2.7)} .$$

$R(t_m;t_0)$ in Equation (2.7) is the random uncertainty in the state caused by the random variable u.

Equation (2.8) defines the change in the state estimate resulting from the observation, and Equation (2.9) updates the covariance matrix P to include the effects of the observation.

In these equations:

$$Y(t_m) = \text{actual observation}$$

$$\hat{Y}(t_m) = G(\hat{X}, t_m) = \text{computed (estimated) observation}$$

$$Y(t_m) - \hat{Y}(t_m) = \text{observation residual}$$

$$\left.\begin{aligned} K_1 &= P_b M^T/(MP_b M^T + C) \\ K_2 &= M^T\epsilon/(MP_b M^T + C) \end{aligned}\right\} \text{See footnote}$$

$$M = \nabla_X[G(X, t_m)]\Big|_{X=\hat{X}}$$

$$\epsilon = \text{control gain of epsilon technique}$$

$$C = E[q(t_m)q^T(t_m)] .$$

The equations are for a single observation, that is, each observation type is sequentially processed. K_1 is the conventional gain for the Kalman linear filter. K_2 represents one type of fix [see References 7 and 8] for preventing divergence of the filter resulting from unmodeled and computational errors. In this chapter, the use of K_2 is called the "epsilon" technique.

K_2 causes overweighting of the most recent observation in Equation (2.8). This, in turn, causes an additive growth of the covariance matrix in Equation (2.9).

Because of the computation time required to carry out the calculations, $\hat{X}(t_m)$ is only available after a small delay. Equation (2.10) therefore predicts $\hat{X}(t_m)$ ahead to time $t_m + \tau$, where τ is larger than the computation time. When real-time reaches $t_m + \tau$, Equation (2.11) is used to modify the total system state estimate.

2.1.4 C-5 Navigation System Considerations

From the preceding discussion on fundamentals of inertial navigation systems, a method of implementing the Kalman filter in a navigation system can be recognized. This is to use the basic inertial navigation system in the role of solving Equation (2.5). To facilitate this, the components of the state vector $X(t)$ were chosen to include quantities (such as position and velocity), estimates of which are carried in the digital computer registers. Also included are the attitude of the inertial platform of the IMU, gyro drift rates, accelerometer biases, measurement errors, and so on. This augmented state vector then contains all the quantities necessary for describing the dynamic response of the real system for given initial conditions.

The $u(t)$ of Equation (2.7) can be interpreted as random forcing functions in the inertial system or in the measurement devices which prevent perfect time updating. Physical evidence has shown that gyros and many other real devices have errors which vary with time and/or other environmental factors. Hence, this random forcing function concept provides a means of including such effects in the filter algorithms.

Figure 2 is a block diagram depicting the conceptual real-time implementation of a Kalman filter, in accordance with the above standpoint, in an on-board navigation computer. In particular, the sequence of operations involved in a single cycle of the on-board filter is shown. Although the figure is largely self-explanatory, it should be noted that careful attention to sequencing these operations, so as to minimize or eliminate possible harmful control lags, is inherent in the diagram.

Details of the Kalman filter which is mechanized in the C-5 navigation system are discussed in later sections.

2.2 Navigation System Description

The C-5 navigation system consists of the Inertial-Doppler Navigation Equipment (IDNE) and a number of auxiliary navigation aids, manufactured by several sub-contractors. Northrop Electronics Division is the system integrator as well as the supplier of major equipment items, including the IMU and the on-board digital computers. Figure 3 is a simplified block diagram of the system*. As shown in the block diagram, a high degree of redundancy was incorporated in the design to enhance overall reliability. The two on-board digital computers are the heart of the basic system. The navigator's control and indicator panels enable him to decide what mode to select, what navigation information to use or reject, and so forth.

Footnote: The measurement dimension is assumed here to be 1×1. This allows use of scalar division, rather than the matrix inversion required in the more general case of an $m \times 1$ measurement.

* Reference 1 gives some additional details.

The Digital Computer

The computers shown in Figure 3 are Northrop Electronics Division general-purpose, parallel operation digital computers (NDC-1051A). The word length is 28 bits. The memory sizes are 12,288 and 8,192 words, respectively, for the primary and auxiliary computers. Other characteristics are as follows:

Memory type:	Random access core, non-volatile DRO
Instructions:	51, with two instructions per word and half-word arithmetic capabilities
Add time:	Half-word 6 microseconds Full-word 8 microseconds
Multiply time:	Half-word 28 microseconds Full-word 74 microseconds
Other features:	Indexing Indirect addressing Roll table (push-down stack) Interrupt capability.

The IMU

The IMU (Inertial Measurement Unit), shown in Figure 3, is the Northrop floated ball, gimballess platform. Three Northrop single-degree-of-freedom floated, rate-integrating gyroscopes are orthogonally mounted inside the floated ball. Each gyro error signal is detected and amplified. The amplified signals are applied to ball-mounted torquers which cause the ball to rotate. This high-gain, three-axis, feedback attitude control system rotationally isolates the floated ball from the case, maintaining a stable platform to a high degree of accuracy. Each gyro has an input torquer whose command signals are controlled by the digital computer. Hence, the platform attitude relative to a given reference frame can be changed by quantities computed by the digital computer.

Three Kearfott linear pendulous accelerometers are also orthogonally mounted inside the floated inner ball. The accelerometer-sensed signals are sent to the digital computer in the form of a train of pulses. The number of pulses in a given unit of time divided by the time interval is a measure of the average specific force exerted on the IMU case over the interval. The frequency of the pulses for the C-5 system is high enough for this average specific force to be considered as a continuous quantity, for all practical purposes.

The IMU also contains an electrostatic, three-axis, spherical angle measuring system. This "ball readout" provides a measure of the orientation of the platform case (outer ball) relative to the stable member (inner ball).

Doppler Radar

The Doppler radar, manufactured by the GPL Division of General Precision Systems, Inc., supplies the computer with measurements of ground speed and drift angle*. By utilizing information from the "ball readout" mentioned previously, it is possible to obtain the ground speed relative to platform axes.

LORAN

The LORAN receiver is manufactured by Collins Radio Company. When available, the LORAN data is converted by the digital computer to position information and displayed to the navigator. Subject to his judgment and control, the LORAN data can also be used to update the best estimate of the state vector which includes position, velocity, and other quantities.

TACAN

The TACAN receivers, manufactured by Hoffman Electronics, provide slant range and bearing information to ground stations at known locations. The navigator can use this data, when it is available, in a similar manner to that described for the LORAN.

Multi-mode Radar (MMR)

The multi-mode radar, manufactured by Norden Division of United Aircraft, provides many modes of operation for the C-5. Two modes are used to obtain navigation information. In the terminal phase of flight, the navigator can, through a cathode ray tube display, sight on landmark patterns whose position relative to desired supply and/or personnel drop location is known. Slant range and bearing data is provided by the MMR during the terminal phase. The navigator can also use the MMR for sighting on landmarks whose absolute position (latitude and longitude) are known. In this mode the measurements of range, bearing, and landmark location go directly to the Kalman filter.

---

* The drift angle is the angle between the ground speed vector and the aircraft longitudinal axis.

Attitude-Heading Reference Unit (AHRU)

The two redundant attitude and heading reference units, shown in Figure 3, are manufactured by Lear Siegler. Data from these units is used to provide a back-up mode of operation in case of an IMU malfunction.

Central Air Data Computer (CADC)

The two air data computers are manufactured by Elliot Brothers. One of the functions of the CADC is to provide barometric pressure altitude data to the navigation system. (See Reference 1 for other functions.)

Inertial-Doppler Navigation Equipment

Figure 4 is a functional block diagram of the C-5 Inertial-Doppler Navigation Equipment, which performs the navigation function in the primary mode of operation. This major subsystem of the navigation system is supplied by Northrop Electronics Division. It contains elements organized and located as shown in Figure 4. The Doppler antenna and IMU are in the nose visor of the aircraft. The electronics rack contains the A-D/D-A (analog-to-digital and digital-to-analog) converters, the digital computers, and the Doppler signal data converter. The control and indicator panels are in the navigator's area.

Before installation in the aircraft, a tape reader, which is part of the auxiliary ground equipment, is used to load the real-time computer program and associated constants into the memory of the primary and/or auxiliary computer. The navigator also has access to some parts of the computer memory through a keyboard input.

2.3 Navigation System Functional Capabilities and Accuracy Goals

2.3.1 System Operating Modes

As can be seen from the discussion relative to Figure 3, the C-5 navigation system has numerous equipment redundancies to enhance overall system reliability. In addition, the navigator can select any one of four basic navigation modes, depending upon equipment availability. A brief description of these primary modes is now given.

Free Inertial This mode uses the IMU, digital computer, and barometric pressure measurements of altitude.

Doppler-Inertial This mode adds the Doppler measurements to those used in the free-inertial mode.

Doppler Dead-Reckoning In case of IMU failure, if the Doppler and AHRU equipments are available, this mode can be selected.

True Airspeed Dead-Reckoning If both the IMU and the Doppler radar fail, AHRU and airspeed data from the CADC can be used to continue dead reckoning.

In any one of these basic modes, position data from LORAN, TACAN, multi-mode radar, or visual information can also be used to update the best estimate of the state vector.

In addition to these navigation modes, there are also the platform alignment (ground or air) modes. Here, as in the navigation modes, the Kalman filter is used for weighting the various available data to continually optimize navigation and platform alignment.

Finally, a special terminal navigation mode uses multi-mode radar data to attain the high relative-position accuracy required for determining the computed air release point (CARP) in cargo or personnel drops.

2.3.2 IDNS Performance Goals

The design performance goals for the C-5 navigation system are those defined for system initialization and navigation in the Doppler-inertial and free-inertial modes, as follows.

(a) Ground alignment (-75° to +75° latitude):

(1) Warm-up from -65°F and alignment after warm-up, both in 25 minutes.

(2) Initial position uncertainty: 1.0 nautical mile (3σ).

(b) Airborne alignment (-75° to +75° latitude):

(1) Alignment after warm-up in 20 minutes.

(2) Initial position uncertainty: 5.0 nautical miles (3σ).

(c) Navigation (free-inertial and Doppler-inertial):

(1) Ground alignment – first 5 hours: 0.75 nautical mile per hour CEP,
– second 5 hours: 1.25 nautical miles per hour CEP.

(2) Airborne alignment – first 5 hours: 1.0 nautical mile per hour CEP,
– second 5 hours : 2.0 nautical miles per hour CEP.

These design goals are explained as follows. The system should be able to be aligned within 25 minutes from cold start (-65°F) at latitudes -75° to +75°, even when the initial location is not known to better than a 1 nautical mile (3σ) tolerance.

Following ground alignment, the ensemble position error growth rate should be no larger than 0.75 nautical mile per hour CEP* for the first five hours and 1.25 nautical miles per hour CEP for the second five hours, for either the Doppler-inertial or the free-inertial mode. Following air alignments, the corresponding goals are 1.0 and 2.0 nautical miles per hour, even with a 5 nautical mile (3σ) initial position uncertainty.

The free-inertial mode performance goals are, of course, the more stringent. That is, Doppler data provides added measurements which, if properly used, must improve the system performance over that obtained with the free-inertial mode.

Obviously, the operational environment for ground alignment is far more severe than a laboratory environment. The C-5 is a heavy logistics transport and the IMU and other equipment are located in the nose visor. The nose visor is open while the aircraft is loading cargo. The ground alignment process must be carried out during visor motions created by winds and cargo-loading transients. Definitions of approximate, but realistic, statistical models for such factors were required during the Kalman filter development effort.

2.3.3 *Scheduling Requirements*

An unusually stringent schedule was maintained on the entire C-5 program by Lockheed, Northrop, and the other equipment suppliers. For that reason, when considering various system refinements in which a number of alternatives presented themselves, decision deadlines had to be met, sometimes requiring a less thorough than desired evaluation of each alternative.

2.4 Problem Summary and Approach

The C-5 Kalman filter development problem can be summarized as follows.

Given: (a) Performance specifications and physical characteristics of the C-5 subsystem hardware.

(b) Required overall C-5 system functional capabilities and performance goals.

Define: The algorithms for the Kalman filter portion of the real-time program, and all interfaces with other sub-programs of the computer.

A logical first question was: "does a solution exist?". At the start of the problem, relevant information from many sources (such as practical experience and data from conventional inertial navigation systems) was available. Theoretical considerations, preliminary analysis, and practical experience all contributed to give a high degree of confidence that a solution must exist. In fact, as will be seen later, probably many alternative solutions exist. Hence, as the design gradually evolved, engineering judgment played a large role in the selection of the specific solution from the available alternatives.

In solving a problem of this type, it is logical to first ask four primary questions:

(1) What is known about the problem?

(2) What additional information must be obtained?

(3) How is this additional information to be obtained?

(4) Can the generation of this information be organized so that it is available when needed?

The answers to question (1) are summarized in the paragraph above, and the other questions are discussed below. Note that answers to question (3) are the solutions to the problems posed implicitly by question (2). The answer to question (4) is really the approach to solving the overall problem. That is, the overall problem is separated into many small problems, the solutions of which are organized to provide a logical time sequence of overall system development. Typical problems and solution approaches involved in the C-5 Kalman filter design development are illustrated in the following paragraphs.

Example 1

Preliminary statistical information was needed on system performance in its principle modes of operation (e.g., Doppler-inertial navigation following a ground alignment). Adequate statistical error analysis techniques for linear systems were already available. Further, past applications had shown that such techniques were applicable to nonlinear problems, if the idea of a nominal trajectory representing the mean value of the ensemble were used. Use of linear variational equations for small deviations from the nominal trajectory therefore provided the desired linear equations for propagating and relating subsystem and system errors for C-5.

---

* The CEP (circular error probability) usage for ensemble position error growth is interpreted as meaning that, for an arbitrary system selection and test evaluation, the probability is 0.5 that the position error is less than the specified number.

To use this theory several steps were required:

(1) Derivation of an approximate but realistic mathematical model of the inertial system.

(2) Derivation of variational equations for the inertial system error model.

(3) Development of subsystem error models.

(4) Development of general-purpose digital computer programs for calculating ensemble error analysis data.

These error analysis programs, though preliminary in some respects, must be sufficiently flexible to enable investigation of performance sensitivity to measurement cycling rates, non-included error sources, and the like. Early answers to questions in such areas were mandatory to allow reasonably accurate, first-cut assessment of the demands which the final Kalman filter program would make on the on-board computer.

Example 2

Information was needed relating the number of state variables of the Kalman filter to real-time computer speed and memory requirements.

Since, given the algorithms, an experienced real-time programmer can estimate time and memory factors, this problem was solved by cut-and-try procedures. Initially, "ball park" answers were found. These act as guidelines until new information requires refinements. As described in Section 3, such data, together with the results from the ensemble error analysis programs, indicated a need for development of a special data smoothing algorithm for the raw data, before Kalman filter processing.

Example 3

Information was needed about such nonlinear phenomena as the build-up of computational error in the on-board Kalman solutions. No known theory could provide answers to such questions. Simulation of the system on a general-purpose digital computer, however, promised a very practical solution. Hence, development of a simulator and study of such effects under simulated conditions was the solution chosen. As the design evolved prior to laboratory and flight tests, the simulator was expected to become the primary tool for validation and/or checking of algorithm refinements. If the simulator design were made sufficiently versatile, it would also help in understanding the causes of, and determining fixes for, anomalistic behavior of the real hardware in later stages of the system development.

These are a few examples of the considerations which influenced the approach to the overall problem solution. From the three examples given, it should be reasonably obvious that the problems could not all be solved simultaneously. An orderly, time-sequenced approach had to be formulated with overlaps in obvious areas, and with contingencies for redirection based on new information. Such an approach adapted itself naturally to the overall system development time schedule. The sequence of presentation in the next two sections follows, to a large degree, the overall approach adopted by Northrop management at the start of the development effort.

3. DESIGN DEVELOPMENT AND DESCRIPTION

The previous section has briefly discussed some of the problems which had to be solved in the development of a practical C-5 Kalman filter. The most important initial steps were the derivations of:

(a) the system state vector,

(b) dynamic equations for the state vector,

(c) variational equations for propagating errors in the state vector,

(d) statistical models for the error sources.

Furthermore, a general-purpose computer program was needed early for evaluating effects of various subsystem design parameters on system performance.

As data was gathered from the above tasks, the constraints imposed by the limitations of the real-time computer were investigated. These were:

(a) real-time availability,

(b) computational precision,

(c) memory availability.

This section follows the overall development in the sequence given, and summarizes the design algorithms of the real-time computer that resulted from this phase of the development.

3.1 Mathematical Modeling

The general definition of an appropriate mathematical model for a real system is one of the most difficult tasks in the application of Kalman filter theory. The decision as to what degree of approximation is appropriate is primary, since unnecessary complexity will result in unacceptable real-time computer storage and memory requirements.

To illustrate, consider the problem of defining a C-5 system state vector. The primary objective of the navigation system is to provide continuous, accurate indications of the C-5 aircraft position, velocity, and heading*.

The differential equations for aircraft acceleration can be written in the form

$$\ddot{\vec{z}} = \vec{G} + \vec{A} + \vec{T}, \tag{3.1}$$

where

$\ddot{\vec{z}}$ = aircraft acceleration

$\vec{G}$ = gravitational force/unit mass,

$\vec{A}$ = aerodynamic force/unit mass,

$\vec{T}$ = thrust force/unit mass.

The velocity $\dot{\vec{z}}$ and position $\vec{z}$ would comprise six vector elements each. To describe $\vec{A}$ and $\vec{T}$, a great many more state variables would be needed. This definition would permit treatment of the IMU accelerometer outputs and many other sensed quantities (e.g., engine settings, control deflections, etc.) as observations. If, however, the differential equations were written in the form

$$\ddot{\vec{z}} = \vec{G} + \vec{f}, \tag{3.2}$$

where $\vec{f}$ represents IMU accelerometer outputs, then it would no longer be necessary to model the aircraft attitude motions, engine dynamics, etc. A large number of state variables would thereby be replaced by the outputs of a measuring instrument. In the use of Equation (3.2), however, the IMU accelerometer outputs are *not* observations in the Kalman filter theory sense. Instead, they are forcing functions. Nevertheless, as will be seen, errors in these "measured" forcing functions can be accounted for with this approach.

For either of the two sets of Equations (3.1) or (3.2), Doppler data, barometric pressure, LORAN, TACAN, and the like, all of which are only functions of position and/or velocity, can be treated as observations. Suppose now that the IMU fails. Assuming that Doppler data is still available, the equation

$$\dot{\vec{z}} = \vec{g}(t) \tag{3.3}$$

can now be written, where $\vec{g}(t)$ is the Doppler velocity measurements. The Doppler data is no longer an observation in the filter theory sense; however, the LORAN and TACAN data, which are functions of the position state variables, $\vec{z}$, are still observations.

The definition implied by Equation (3.2) was the one selected for C-5. As mentioned earlier, a wander azimuth, geodetic vertical frame had already been selected for the navigation equations. Hence, the differential Equations (3.2) were changed to define the vehicle acceleration and velocity in this rotating reference frame.

Care in state vector definitions also allowed the equations for the dead-reckoning modes to be sub-sets of the equations for the Doppler-inertial modes. This permitted a much simpler overall formulation for the multi-mode system.

3.1.1 Free-Inertial Mode Model

As already discussed, it was planned to use the free-inertial navigator (in the primary modes) for updating the total estimate of state in the time intervals between measurement cycling of the Kalman filter. Therefore, the need for an integration of nonlinear differential equations as part of the Kalman filter was obviated. However, a derivation of an approximate model of the system was needed as a basis for an approximate transition matrix for the Kalman filter. As described earlier, the differential equations for the system state were

$$\dot{X} = F(X,t) + u. \tag{3.4}$$

Using these, the variational equations:

$$\dot{\tilde{x}} = A(t)\tilde{x} \tag{3.5}$$

were required, where

$$A(t) = \nabla_X[(F(X,t))]\Big|_{X(t)=\hat{X}(t)}. \tag{3.6}$$

$\hat{X}(t)$ is the estimate of system state. Several of the components of $\hat{X}(t)$ are available from the inertial navigator.

The inertial navigator consists of the IMU and the navigation equations solved in the NDC-1051A computer (see Figure 1). A reasonable model for Equation (3.4) can be defined by writing approximate differential equations for the IMU computer system. To illustrate how this can be accomplished, consider the modeling of the way by which the specific force vector is obtained for input to the navigation equations (see Figure 5).

* An inertial, non-rotating coordinate frame is assumed here for simplicity of description.

The true, vector specific force $\vec{f}_I$ is transformed to the platform reference frame by the transformation $T_{PI2P}(t)$. Since the accelerometer-sensitive axes are not precisely orthogonal, the non-orthogonal transformation, T_{P2A}, transforms the specific force to accelerometer axes. In addition, accelerometers have scale factors which are not precisely known; the diagonal matrix SFA represents these factors. Finally, the accelerometer outputs are also corrupted by bias offsets and noise, represented by the two vectors $\vec{a}_b$ and $\vec{a}_n$ respectively. The vector $\vec{f}_o$ represents the simplified, continuous model for the outputs of the actual accelerometers.

Standard C-5 system procedures include provision for estimation of all the above accelerometer-related errors and for a set of gyro-related errors discussed below as well, during periodic, laboratory calibration operations on each IMU. These numbers are then incorporated into the computer to appropriately compensate the accelerometer inputs to, and the gyro command outputs from, the computer during actual system operation. In this connection, the vector $\vec{a}_{bc}$ in Figure 5 denotes the calibration for accelerometer null bias $\vec{a}_b$, and the 3 x 3 matrix $[a_{ij}]$ denotes the calibration for the accelerometer scale factor and input axis orthogonality errors, that is

$$[a_{ij}] \approx [T_{P2A}SFA]^{-1} .$$

The estimated specific force vector $\vec{A}_s$ can therefore be written in terms of all these quantities. The error in $\vec{A}_s$ due to each one of the contributing accelerometric error sources can, in turn, be approximated by first-order variations of these relationships.

The error in $\vec{A}_s$ caused by the errors in the transformation $T_{PI2P}(t)$ must also be derived. These latter errors are called tilts because ideally, $T_{PI2P} = I$, if $\vec{f}_I$ is defined in the true local tangent plane reference system.

Continuing with Figure 5, the vector $\vec{A}_s$ feeds the navigation equations, which are shown schematically in Figure 6. Although the NDC-1051A calculations are discrete rather than continuous, their repetition rate is high (12 Hz). Hence the error differential equations for these calculations can be approximated by taking gradients of the nonlinear equations with respect to the various variables.

Three of the nine fundamental state variables can be identified by referring to Figure 6. These are the integrator outputs:

$$V_l, \vec{V}_L = \text{vertical and level velocities}$$

$$h = \text{altitude} .$$

Two others, latitude and longitude, are obtained from the transformation T. The final fundamental state variables are the three rotations necessary for defining the transformation $T_{PI2P}(t)$ discussed above.

The closure of the loop (i.e., the gyro signal processing) is illustrated in Figure 7. The gyro torquing signals, represented by the vector ω_o, are transformed by the calibration matrix $[g_{ij}]$. This matrix is a calibration for the gyro scale factor errors (diagonal matrix SFG) and the gyro input axis non-orthogonalities (matrix T_{GI2P}). Note that T_{GI2P} transforms rates from the gyro input axis frame to the platform reference frame. Also,

$$[g_{ij}] \approx [(SFG)(T_{GI2P})]^{-1} .$$

Shown also are the $[u_{ij}]$ matrix, which compensates for gyro drift rates caused by static mass unbalance, and the vector $\vec{u}_{bc}$, which is a calibration of the fixed gyro drift rate. In Figure 7, $\vec{\omega}_d$ is the actual vector drift rate of the gyros, referred to the input axis. Reference 3 gives many details on the modeling of such factors.

The final output is the total inertial rate of the platform, $\vec{\omega}_p$. Again, first-order variations of the various equations are used to define errors in $\vec{\omega}_p$ created by the various error sources.

The process just outlined led to the definition of approximate variational equations for the free-inertial navigator errors in the state-space form:

$$\underset{(9 \times 1)}{\dot{e}} = \underset{(9 \times 9)}{\alpha(t)} \underset{(9 \times 1)}{e} + \underset{(9 \times 41)}{\beta(t)} \underset{(41 \times 1)}{e_f} .$$

There are nine error state variables and 41 error forcing functions. Some of the error forcing functions were assumed to be constants (e.g., accelerometer and gyro misalignments), while others were prescribed functions of the flight path (e.g., static mass unbalance drifts); however, known fluctuations in gyro drift rate and accelerometer noise needed a random characterization. This was accomplished by assuming that such errors were representable by white noise driving a first-order filter in accordance with Figure 8.

Before the final, augmented state vector formulation is defined, a brief consideration of the observation modeling is in order.

3.1.2 *Observation Models*

The process of appropriately describing the observations and their error models follows a pattern similar to that described for the free-inertial model.

In certain instances (e.g., a LORAN time-difference pair and inertial navigator latitude and longitude), a nonlinear equation relates the inertial and subsystem measurements; in others, a direct linear relation exists (e.g., barometric pressure altimeter and inertial altitude indications). In either case, however, linear relations (M matrices) between inertial and subsystem errors at a measurement time were developed by simple first-order variations.

Modeling of the time propagation of subsystem errors necessitated the use of time-correlated random variables. These were all approximated by white noise into a first-order filter, as already mentioned. This process led to the definition of 17 more error state variables (constant biases and/or correlated noise) of the type described by the vector V of Equation (2.2).

3.1.3 *Final Linear Dynamic Error Model*

Augmentation of the fundamental state vector to include all the deterministic type of error sources led to a final dynamic error model of the form

$$\left.\begin{array}{c} (67\times 1)\ (67\times 67)\ (67\times 1)\ (67\times 1) \\ \dot{\tilde{x}}(t) \;=\; A(t)\tilde{x}(t) + n(t)\ . \end{array}\right\} \tag{3.7}$$

The vector $\tilde{x}(t)$ represents the total augmented error state vector. The vector $n(t)$ represents the white noise inputs for some of the augmented state variables. Most of the components of $n(t)$ are identically zero. It should be noted that this model assumes a smooth time propagation of system errors. In reality, because of the quantized, iterative nature of digital computer signal processing, the actual errors really exhibit staircase behavior. However, because of the fineness of the quantization (typically one part in 2^{27}), and the high iteration and sampling rates employed (e.g., 12 Hz), the continuous approximation is an adequate one and, of course, highly convenient for simplifying any necessary mathematical manipulations of the model.

3.2 Ensemble Error Analysis Program

At this stage of the overall development effort it was possible to start the construction of a 67-variable ensemble error analysis program, hereafter called Program I for convenience of reference. The purpose of this program was to establish the system performance possible if the C-5 Kalman filter incorporated all 67 state variables. This error model, although still approximate, contained all the error sources considered as having some significance; it is therefore hereafter referred to as a "complete" error model.

Data from this program was intended to serve as a reference standard, against which the performance of subsequent, less sophisticated alternatives* could be measured. Construction of the overall program required:

(1) A means of generating desired flight profiles.

(2) Input means for controlling all the assumed initial error sources, the flight profile, and the sequence of observations.

(3) The algorithm for the change in the error covariance matrix resulting from use of the observations.

(4) The algorithm for updating the error covariance matrix between the observations.

Of these, items 1 and 2 are program details which will not be pursued here. For item 3, an optimal filter was assumed. Item 4, however, merits some discussion, since it incorporated some approximations which were subsequently used in the actual, on-board Kalman filter program.

3.2.1 *Prediction Between Observations*

The algorithm for prediction of the covariance matrix across the intervals between successive operations was developed as follows.

If $\tilde{x}(t)$ denotes the error in the state estimate at time t, then

$$\tilde{x}(t_m) \;=\; \Phi(t_m;t_0)x(t_0) + \int_{t_0}^{t_m} \Phi(t_m;\tau)n(\tau)\,d\tau \tag{3.8}$$

describes the error propagation across the prediction time interval ($\Delta t = t_m - t_0$) between successive observations†. Defining the covariance matrix of the error by

---

* It is apparent that the on-board computer could not solve such a voluminous problem in real-time and, as will be seen, it is not practical to do so.

† Later this time interval will be seen to be the actual cycling time of the Kalman filter.

$$P(t_m) = E\{\tilde{X}(t_m)\tilde{X}^T(t_m)\} , \tag{3.9}$$

it follows that

$$P(t_m) = \Phi(t_0+\Delta t;\ t_0)P(t_0)\Phi^T(t_0+\Delta t;\ t_0) + R(t_0+\Delta t;\ t_0) , \tag{3.10}$$

where

$$R(t_0+\Delta t;\ t_0) = \int_{t_0}^{t_0+\Delta t}\int_{t_0}^{t_0+\Delta t} \Phi(t_0+\Delta t;\ \tau)E\{n(\tau)n^T(\xi)\}\Phi^T(t+\Delta t;\ \xi)\, d\tau d\xi . \tag{3.11}$$

Equation (3.10) is the desired algorithm for updating the error covariance matrix between observations. However, suitable approximations for the transition matrix, Φ, and for the system noise matrix, R, needed to be devised. As an illustration, the development of the transition matrix proceeded as follows.

The transition matrix obeys the differential equations:

$$\dot{\Phi}(t) = A(t)\Phi(t)$$

$$\Phi(0) = I ,$$

When A is a constant matrix, a closed-form solution for Φ can be obtained:

$$\Phi(\Delta t) = \mathcal{L}^{-1}[(sI - A)^{-1}] , \tag{3.12}$$

where s is the Laplace variable, and I is an identity matrix of the same dimension as A. When the dimension of A is large, however, it is generally hopelessly intractable to algebraically carry out the indicated matrix inversion. Other, approximate methods must therefore be used. One of the simplest of these involves the use of the power series expansion:

$$\Phi(\Delta t) = \sum_{i=0}^{\infty}\left(\frac{1}{i!}\right)(A\Delta t)^i = I + (A\Delta t) + \frac{1}{2!}(A\Delta t)^2 + \ldots . \tag{3.13}$$

The series is truncated at a point sufficient to provide the desired accuracy. However, if A is time-dependent, which is the case for C-5 airborne modes, the corresponding series is:

$$\Phi(\Delta t) = I + (A\Delta t) + \frac{1}{2!}(\dot{A} + A^2)(\Delta t)^2 + \ldots . \tag{3.14}$$

The higher-order terms not shown contain higher-order derivatives of A. There are some difficulties associated with attempting to use a truncated form of this expansion. These stem from the impracticality of obtaining the first- and higher-order derivatives of A, if second- or higher-order accuracy is required. This impracticality arises because A naturally contains only the basic inertial navigation signals, up to the acceleration level, as elements. Use of terms involving $\dot{A}$ or higher derivatives would therefore require the artificial, approximate generation of otherwise unavailable acceleration derivatives.

On the other hand, the accuracy involved in the use of only the first two terms in the expansion depends on the size of Δt. It was evident early in the program that, because of the multiple non-Kalman functions required of the C-5 computer, only a small percentage of its real-time use could be devoted to cycling the Kalman computations. In addition, even with the large anticipated reduction in the size of the state vector (from 67), these computations would still be lengthy and therefore time consuming. All of this suggested that the design, even at this stage, should anticipate the need for a large Δt, i.e., of the order of many seconds to a few minutes. To accommodate this need, the following approach was adopted.

One of the important properties of the transition matrix is its exponential character; that is, the transition matrix for state evolution between times t_1 and t_3 is the product of the transition matrices associated with state evolution from t_1 to t_2 and from t_2 to t_3, respectively. A simple extension of this property gives:

$$\Phi(t_m; t_0) = \Phi(t_m; t_{n-1})\Phi(t_{n-1}; t_{n-2}) \ldots \Phi(t_2; t_1)\Phi(t_1; t_0) , \tag{3.15}$$

where $t_n = t_m$.

Denoting $t_i - t_{i-1} = \Delta t_i$, and setting $\Delta t_i = \Delta t/n$, where $\Delta t = t_m - t_0$, the Δt_i could then be visualized as the n equal time intervals between $n+1$ successive updates of all navigation signals by either the navigation computer (C-5 filter mechanization) or the navigation sub-routine (ensemble error analysis program), and Δt could be visualized as the Kalman cycle time itself. Since Δt_i could now be assumed quite small (down to 1/12 second in the navigation computer, and arbitrarily small in the ensemble program), the truncated form,

$$\Phi(\Delta t_i) = I + [A_{i-1}\Delta t_i] + \tfrac{1}{2}[A_{i-1}\Delta t_i]^2 , \tag{3.16}$$

was tentatively assumed sufficiently accurate for extrapolation over a Δt_i interval. Substitution of Equation (3.16) into (3.15) gives

$$\Phi(\Delta t) = \prod_{i=1}^{n}\left[I + (A_{i-1}\Delta t_i) + \tfrac{1}{2}(A_{i-1}\Delta t_i)^2\right] \tag{3.17}$$

or, for $n > 1$ (see footnote),

$$\Phi(\Delta t) = I + \Delta t^* \sum_{i=1}^{n} A_{i-1} + \tfrac{1}{2}(\Delta t^*)^2 \sum_{i=1}^{n} A_{i-1}^2 + (\Delta t^*)^2 \sum_{i=2}^{n} A_{i-1} \sum_{k=0}^{i-2} A_k , \tag{3.18}$$

where $\Delta t^* = \Delta t_i = \Delta t/n$, and terms of order higher than $(\Delta t^*)^2$ have been dropped.

Equation (3.18) still presented difficulties, primarily because mechanization of the last (double summation) term suggested considerable complexity. This term was therefore tentatively dropped, as were all the remaining second-order terms, except for certain principal ones associated with constant elements of the matrix A_{i-1}.

The residual expansion finally adopted was thus

$$\Phi(\Delta t) = I + A_0\Delta t + \tfrac{1}{2}A_0^2\Delta t^2 + \sum_{i=1}^{n}\left[A_{NC_{i-1}}\Delta t^*\right], \tag{3.19}$$

where $A_0 + A_{NC_{i-1}} = A_{i-1}$, and A_0 and $A_{NC_{i-1}}$ are respectively the constant and non-constant portions of the matrix A_{i-1}. A_0^2 represents a matrix of selected, principal constant elements of A_{i-1}^2.

For preliminary evaluation, a small program was constructed to check the validity of Equation (3.19) if used recursively to approximate stationary linear dynamics (e.g., the C-5 ground alignment situation) over long periods of time. Although this program used only a limited-dimension model, the model incorporated the principal dynamical characteristics associated with the stationary elements of the A matrix. The results indicated extremely small percentage drifts relative to an exact solution (obtained by use of Equation (3.12) on the limited-dimension model) over 10-hour periods, even with Δt as large as several minutes.

The airborne modes require additional non-stationary A matrix terms that depend on the aircraft flight profile. Evaluation of the adequacy of Equation (3.19) in airborne modes was therefore deferred until after the development of the system simulation program, described later in this chapter.

3.2.2 *Observations*

Three types of information external to the inertial navigation subsystem would be available for comparison with its outputs in the ground alignment mode and the inertial and Doppler-inertial navigation modes:

(a) barometric altitude from the CADC;

(b) groundspeed and drift angle from the Doppler;

(c) the knowledge that the nose visor is essentially stationary* during ground alignment.

Observation of the information described by (c) is discussed here to illustrate the general approach.

Since the aircraft is parked during ground alignment, either the zero velocity or the fixed position of the aircraft can be used as the information reference against which to compare, respectively, the inertial navigator velocity or position†. However, since the use of velocity-level measurements was already anticipated for Doppler-inertial navigation, velocity comparison was selected for ground alignment. This provided a greater uniformity between the ground alignment and Doppler-inertial mode mechanizations.

If either one of the inertial navigator level velocity readings is denoted by v the corresponding Kalman observation Y, used in ground alignment, could therefore be

$$Y = v .$$

The signal v can be decomposed into the form:

$$v = v_A + \delta v , \tag{3.20}$$

where v_A is the actual visor velocity component, due to wind and loading along the level direction associated with v, and δv is the error in the signal v arising primarily from initial tilts of the platform with respect to the vertical, and from uncalibrated platform drift rates.

Footnote: The notation from here on has been purposely simplified for compactness; thus, $\Phi(\Delta t)$ is used for $\Phi(t_n; t_0)$, and the product notation $\prod_{i=1}^{n}$ implies the *ordered* product of matrix factors ($i = 1$ to n from right to left).

* Except for the motions induced by wind and aircraft loading.

† Either reference, if properly used, is entirely equivalent to the other.

From the standpoint of Equation (3.20), it would appear that δv is the signal containing information, and v_A is the noise from which it must be separated. However, this assumption about v_A is justified only if v_A consists mostly of noise which is of high frequency with respect to both the δv signal content and the contemplated observation iteration rate. Because of the frequency response characteristics of the massive C-5 aircraft to random wind and loading disturbances, this proved to be true for ground alignment. If it were not true, it might have been necessary to introduce error state variables to account for systematic low-frequency behavior of v_A to meet alignment performance goals. Such variables, which generalize the system error state vector concept, were indeed introduced elsewhere to account for low-frequency components in other mode observations.

Since δv was an early, natural choice as one of the essential all-mode state variables, and since the notation relating measurements to state is

$$Y - \hat{Y} = M(\tilde{x} - \hat{x}) + q$$

by association, M was therefore defined as:

$$M = [0 \;\; 0 \;\; 0 \ldots \; 1 \ldots \; 0] \quad \begin{matrix} 1 & 2 & 3 \ldots & \delta v & \ldots & n \end{matrix}$$

The choice of two scalar measurements instead of one (two-dimensional) vector measurement for ground alignment deserves comment. This choice does not require matrix inversion. Instead, two successive passes are made through the observation update portion of the Kalman equations. Each pass only involves a division for the inversion of the quantity $(MPM^T + C)$ in Equation (2.9). In fact, all observations were defined as scalars to make this processing simplification uniform.

Although this approach is theoretically exact only if the noise associated with each separate observation is uncorrelated with that associated with every other observation, ground alignment proved to be the only C-5 case where this condition was not substantially met. The resulting performance of ground alignment, as will be seen, nevertheless met design goals, despite this theoretical infringement.

Having defined the observation matrices associated with all C-5 modes and information sources, it was then possible to complete the 67-variable ensemble error analysis program (Program I) by incorporating the following matrix equations:

$$\left.\begin{aligned} Q &= MPM^T + C \\ b &= PM^TQ^{-1} \\ P &= P - bMP \end{aligned}\right\} \qquad (3.21)$$

These, together with Equation (3.10), completed the set of recursion equations necessary to obtain the time history of the covariance matrix P of navigator errors through a typical C-5 flight operation.

To simulate such flights, the program was equipped with a general, flexible flight profile sub-routine capable of providing great circle cruise, climb, descent and turn phases in any commanded sequence. The navigation signals thus generated were used to drive the dynamic elements of the transition matrix.

Program I was coded in FORTRAN IV single precision for use on the IBM 7090 or 360 computers. On completion of checkout, the immediate, extensive use of Program I at this stage established the ability of a 67-variable, full precision 7090 or 360 mechanization to meet or exceed all C-5 navigation system performance goals.

3.3 Design Review and Discussion

At this stage of the development the overall final filter design was beginning to take shape. Approximately half the development time had elapsed, however, and many of the difficult problems imposed by the real-time computer constraints had yet to be faced.

3.3.1 Overall Design

The overall design mechanization for the multi-mode system was reasonably well defined. Figure 9 is a functional block diagram of the overall mechanization, showing the role of the filter in the overall navigation information processing. Although this is essentially an analog diagram, definition of the overall sequencing of the digital calculations was also well advanced, as will be seen below.

In this connection, it was recognized that the actual Kalman filter would require an appreciable time to process a single set of observations and to produce revised error estimates for application as system corrections. While this was going on, data on current position and velocity would continue to be required for inputs to other, non-navigational functions (e.g., aircraft steering) and for display. Since such signals might be available at several different points (e.g., groundspeed from corrected Doppler groundspeed or from the horizontal navigation equations in Figure 9), one source had to be uniformly selected as "best" between successive Kalman corrections. Because of their relatively high update frequency and their relative smoothness, the (corrected) inertial navigation signals were chosen as the best outputs for this purpose.

3.3.2 *Kalman Filter Design*

Table I summarizes the C-5 Kalman filter equations. The computations shown, together with the formation of the various matrices involved, comprise one full Kalman processing cycle in the on-board computer, and are given in the order of actual processing.

As discussed earlier in this chapter, the Kalman equations are modified to include the "epsilon" weighting technique and to provide for system and estimator control. At one time it was contemplated not to mechanize control at all, but rather to perform only open-loop error estimation. Navigation signals would then be corrected with these estimates only for output to other equipments. This approach was discarded in favor of the closed-loop control technique shown, because it was recognized that otherwise the linearization required would be done about the uncontrolled system state. A large error build-up in this uncontrolled state would cause a correspondingly larger departure from the desired linear properties.

The essentials of the control technique used, which are summarized symbolically in Table I by the notations f() and g(), are as follows:

(a) *Impulsive Control* Wherever possible, the contents of a computer register whose error is modeled as a state vector element are corrected by adding the estimate of error (with the proper sign) to the computer register. Simultaneously, the register containing the error estimate is set to zero. This is done to each such register pair one in every Kalman computational cycle.

(b) *Finite Duration Control* The above impulsive type of control is used uniformly for all registers controlled by the C-5 Kalman mechanization, with one important exception. This exception pertains to those two state vector elements which model the two components of platform tilt; i.e., the two components of the small angular deviation between the actual and desired platform verticals. To remove the actual tilts, the gyros are commanded to torque the platform vertical at a uniform rate over the next Kalman interval. This rate is exactly that required to take out the estimated tilt in the platform in one Kalman cycle. At the same time that these leveling rates are first applied, the platform drift rate estimates are set to these computed leveling rates. The current tilt estimates are left untouched.

Note that a basic, built-in invariant implicit in these control techniques is the difference between the actual and the estimated system behavior. This invariance allows retention of the simple, open-loop covariance matrix recursion relations unchanged, if it is assumed that the additional errors created by the application of control are negligible. This assumption was made at this stage of development for later validation.

The off-optimal "epsilon" factor was introduced in the design algorithms as a precautionary measure to combat unmodeled and numerical errors (see References 7 and 8).

3.3.3 *Other Relevant Information*

Concurrently with the effort described thus far, some additional data was being accumulated as a result of related work on another system. For example, an NDC-1051 computer (24-bit word) had been programmed for Kalman estimation in the fine ground align mode, and laboratory tests were conducted with an INU of different design from another program. The results were not overly encouraging.

One of the problems was the difficulty of maintaining a positive definite covariance matrix. The other appeared to be connected with the use of fixed-point arithmetic. Only poor estimates of gyro drift rates were obtained. This seemed to be tied to the loss of numerical significance in the covariance matrix due to the fixed-point calculations. These results were one of the reasons for the introduction of the "epsilon" technique in the basic mechanization equations for the C-5 system.

3.4 Constraint-Imposed Problems and Solutions

One of the principal problems which had to be dealt with at this point was the identification of the smallest system error model which could be mechanized for the on-board Kalman filter, and still meet system performance goals. Use of the complete, 67-variable model of Program I, for example, would require far too much memory and real-time of the on-board computer. However, Program I results did indicate that the complete model, if mechanized, would substantially surpass system performance goals. This, therefore, provided a reasonable performance margin to work with, in terms of the loss of performance inevitably expected from on-board use of smaller models.

Actual quantitative evaluation of this performance loss is a relatively subtle problem which cannot be solved, for example, by comparison of 67-variable model Program I results with results from an analogous program incorporating the smaller model. Rather, a more general type of computer error analysis program is required. As will be seen, the stringent overall C-5 development schedule had a direct effect at this time on the selection of the type of program to be constructed to fill this role.

In addition, definition of the airborne computer tape sub-routines had been proceeding concurrently in all other non-Kalman areas. Since design in these more conventional areas naturally proceeded faster than it did in the more novel Kalman area, the programs for these sub-routines, together with their attendant word count and time allocations, began to crystallize somewhat ahead of those for the Kalman sub-routine. This process led to the early establishment of word count and real-time allocations for the Kalman sub-routine. These allocations constituted the two primary constraints on the filter design.

These contraints, together with the accuracy goals and all-mode operation requirements, comprised the framework within which the Kalman design had to be completed. The problems created by these contraints, the methods of attack used, and the final design solutions adopted are now discussed.

3.4.1 *Computer Storage Considerations*

The relative fractions of the total computer storage allocated to the various functions of a real system like this one is always a matter of compromise, requiring continuous trade-off evaluations. A Kalman filter having sufficient dimension to satisfy the C-5 requirement was new, although experience did exist on more limited implementations. However, these implementations were inadequate as a basis for estimating the C-5 Kalman storage requirement. Nevertheless, the storage available was defined quite closely, because the 12,288-word total (primary computer) was fixed and the requirement for most of the other routines could be closely estimated on the basis of experience.

Depending on how much program compaction effort was to be applied to the other routines, the storage available for the Kalman routine ranged from 2000 to 3000 words. It was immediately evident that the model size must be reduced (from 67) as far as performance goals would allow.

In conjunction with this effort, it became evident that the minimal word length to be used in carrying out all arithmetic operations would have to be determined. If double, rather than single, precision operations were used, these would obviously double the non-instruction storage requirements. This consideration of word length also naturally raised a question as to the type of arithmetic to be used (fixed- or floating-point).

Finally, further consideration strongly suggested that the requirement for multi-mode filter operation might demand a sizeable portion of the overall Kalman word allocation for intermode switching.

3.4.2 *Error Model Dimension*

It was recognized immediately that a natural, two-step iterative process, together with the techniques necessary to carry out these steps, would be required to determine the minimal error model for the Kalman filter mechanization. These steps were (1) a method for systematically discarding error variables from the complete (67-variable) model, and (2) a method for evaluation of the reduced models thus obtained.

This latter need - for a reduced model evaluation tool - led to the development of the second major ensemble error analysis program (hereafter referred to as Program II). Two types of formulation, of differing capabilities, were considered. The first and more general of these was based on an extended set of recursive covariance matrix equations (see Reference 9). This type of formulation allows complete freedom in discarding (a) arbitrarily selected error variables, and (b) arbitrarily selected elements of the residual transition and system noise matrices.

A second and simpler formulation was also available[10]. This one allowed calculation of the increase in error produced by discarding *constant* state vector elements only. Despite this theoretical limitation, it had the overriding advantage that it allowed rigorous handling of all bias errors (which were predominant) and permitted much faster program construction by only requiring the addition of a few simple operations to Program I. For practical reasons, then, and in the light of the tight development schedule requirements, the second and simpler formulation was adopted.

An efficient method for discarding variables had to be devised, since the turn-around time for Program II results was one or two days. Hence, the obvious brute-force approach involving removal of one variable at a time could not be used. Instead it was necessary to try to identify at once some extensive sets of variables whose removal would, in a few steps, closely approximate the desired minimal error model. These discarded sets, which finally comprised no less than 51 of the original 67 variables, were arrived at as follows.

Between the time that Program I (the original 67-variable model) was formulated and Program II was completed and available, it had been determined for a variety of practical reasons, not to provide outputs from certain on-board navigation equipments as Kalman filter inputs. The state variables originally included in the 67-variable model to account for the errors in these measurements were thus automatically discarded as a first step. For example, although two radar altimeters would be aboard the C-5 aircraft, these would only furnish altitude above terrain, and up to a maximum of only 2500 feet. Since only the estimation of altitude above sea level was contemplated for the Kalman filter, and since the C-5 aircraft would typically spend only a negligible percentage of mission time below 2500 feet over the ocean, use of the radar altimeter data by the filter was excluded. The radar altimeter state variables in the original error model could therefore be discarded.

In addition, where several elements had been used in the complete model to describe the high, intermediate, and low-frequency components of the error in a subsystem navigation reference signal (e.g., Doppler groundspeed), these were compressed into one, where plausible, by eliminating all but the largest amplitude component.

Finally, the largest of the trial sets of discarded elements was composed of the velocity and/or acceleration-sensitive errors associated with the IMU gyros and accelerometers. These errors, although important in a tactical aircraft design application, were expected to produce only minor performance degradations in such a huge transport aircraft as the C-5.

Tables II and III summarize the 67-to-16 model size reduction finally attained by the techniques just described. Figure 10 illustrates the small ensemble performance degradation of the final 16-variable model, relative to the complete model, for a typical C-5 great circle flight, as obtained by Program II.

3.4.3 Word Length/Arithmetic Type

In parallel with the effort to reduce error model size, an investigation was conducted to determine the most suitable word length/arithmetic type for the filter.

One of the problems identified almost at the start of this effort was the large dynamic range requirements of certain covariance matrix elements during the transient phases of filter operation. For example, it was estimated that platform tilts of, typically, one degree would have to be reduced to about one second of arc in a short period at the beginning of ground alignment. This reduction was necessary to achieve sufficiently accurate platform drift rate calibration during the remaining ground alignment time. The ratio of starting tilts to residual tilts is about 3600 to 1. The tilt variances carried in the covariance matrix must, however, traverse the square of this range (i.e., about 10^7 to 1). A single word of the airborne computer is only 28 bits (27 bits plus sign) in length. Hence, if single precision, fixed-point arithmetic were used, the accuracy of settled tilt variance computations would only be about 3 binary bits. The effect of round-off and truncation errors on a word of only 3-bit significance during this critical alignment period was believed to be intolerable.

At this juncture, a wide variety of alternative solutions seemed available, comprising all remaining conceivable combinations of (i) word length, and (ii) fixed- or floating-point arithmetic. However, in the light of further scrutiny, these options were not all practical possibilities.

Before the final reduction to 16 variables, a model of dimension 24 had been attained in the parallel error model size reduction investigation. With this size as a basis, a preliminary sizing of the overall Kalman subroutine storage requirements had produced an estimate of 3300 words, predicated on the use of double precision, fixed-point arithmetic.

Although this figure was still in the ball park, it was evident at this point that 3300 words could be made available only if a considerable programming effort on the other routines was undertaken. (Because of experience gained on previous programs, most of the other routines already made efficient use of storage and further compaction would have been difficult.) Therefore, the necessity for implementing the double precision fixed-point arithmetic approach (or any method that required more than a single 28-bit word) was questioned. Attention was then focused on the use of floating-point arithmetic. Fixed-point arithmetic requires careful scaling of all variables so that adequate significance is retained, while avoiding any danger of overflow. The scaling consideration is unusually complicated in the Kalman filter routine because of the extensive matrix operations. A significant advantage of the floating-point approach, therefore, is that the scaling task is avoided entirely.

The hardware peculiarities of the machine itself now completed the rapidly narrowing restriction of choice. The NDC-1051A computer had a built-in (i.e., hardwired) capability for performing either full-word or half-word fixed-point arithmetic. It was not practical to contemplate a split of the full 28-bit word into a mantissa comprising most of these 28 bits and a small exponent to provide just enough dynamic range to satisfy covariance matrix requirements. (Although such a split could be used if accompanied by attendant sub-routines to provide arithmetic operations, such sub-routines were expected to be somewhat slow. This would have resulted in an inefficient use of computer time.)

The design choice was thus narrowed to the use of floating-point arithmetic, utilizing a single 28-bit word shared equally by mantissa and exponent. In particular, the restriction to a short 14-bit (13 bits plus sign) mantissa created a need for evaluation of filter performance with respect to round-off and truncation errors. Since this evaluation was clearly beyond the existing or easily augmented capabilities of either Program I or Program II, a requirement for a new type of evaluation program was established.

Finally, it is pointed out here, for reference later in this chapter, that the adoption of floating-point arithmetic was the first and major contribution toward increasing the single Kalman cycle execution time. Some preliminary estimates, made at about the time this technique was adopted, indicated uninterrupted execution times of about 10 seconds. The preliminary real-time budget allocated about 20% of computer real-time during ground alignment and 5% in flight to the Kalman routine. This led to estimates of one to three minutes for on-board execution. As will be seen, the unavoidable emergence of these slow Kalman cycles led to a challenging problem later in the development.

3.4.4 Multi-mode Requirements

The preceding paragraphs have centered on the effort to optimize design of the Kalman filter algorithms with respect to storage requirements. These efforts were characterized by use of the error model associated with the most complicated mode of C-5 navigation system operation. This is the Doppler-Barometric-Inertial mode, augmented with position fixes from any or all of the on-board position-measuring sub-system equipments.

However, as has been pointed out, many different modes of system operation were contractually required. The error model associated with each such mode differed with every other to varying degrees. Hence the problems of how to mechanize all such modes, and how to switch between them, had to be addressed.

Multi-mode operation implies mode switching in somewhat arbitrary sequences, and optimal operation requires continuation or relatively complex transfer of the estimation process. For these reasons a unified, multi-mode Kalman filter design was formulated in solving the problem of minimizing storage requirements with this selected approach, the concept of the standard error model proved particularly useful, as follows.

This standard model was composed of the separate error models for the inertial navigator, the Doppler (ground-speed and drift angle) inputs, and the CADC (barometric altitude) input. As such, it constituted the largest dimension (16) model required for dynamic error propagation in any navigation and alignment mode. Further, the state vector, transition matrix, and system noise matrix required to characterize any other mode could in every case be described either in terms of sub-sets of the standard state vector and transition and system noise matrices, or as slightly modified sub-sets whose dimensions never exceeded 16. To a large extent, these convenient error model sub-sets existed because of the care which had been previously taken, with regard to inter-mode uniformity, in the design of the signal flow in each of the mode mechanizations.

The use of this concept made it possible to view all non-standard modes as involving significant information in only certain well-defined portions of the estimate and control vectors and of the covariance and system noise matrices, during both time and measurement update operations. The overall filter operations could then always be done on a fixed dimension (16) basis, independent of mode, without corruption of the information-carrying portion of these matrices by the portions carrying meaningless information. The continuity of state variables common to all modes (e.g., position errors) also facilitated the design of inter-mode switching and initialization. Rigid conformity to these principles, in conjunction with incorporation of the 16-variable model and single-word floating-point arithmetic, ultimately led to an all-mode design which satisfied the 2200-word allocation.

3.4.5 Computer Real-Time Allocation

Throughout the early and intermediate design definition phases, there was a gradual but nearly inevitable drift toward a design requiring more and more execution time. This drift was recognized, and investigation of its implications on the final design was therefore initiated.

The first problem was the need for development of a transition matrix capable of extrapolation across relatively large time intervals. The second problem was the potential loss of reference information (e.g., Doppler groundspeed and drift angle measurements in Doppler-inertial operation). This latter problem, which would arise in the less frequent sampling imposed by a slower filter cycle, was initially ignored because it was assumed that simple measurement averaging over the slow Kalman computational cycle was a solution. This assumption led to the appearance of an unanticipated problem late in the development.

3.4.6 Transition Matrix

The early effort in this area has already been discussed in some depth. The need for evaluation of the preliminary formulation developed at that time had already created the first need for a new tool with capabilities beyond those of ensemble Programs I and II. The system simulation program (SSP) developed to fill this and many other roles in the overall Kalman filter development is described later in this chapter.

3.4.7 Measurement Averaging

One of the more stringent design goals of the C-5 navigation system design was the short time allowed for the alignment and calibration of the inertial platform prior to inertial or augmented inertial navigation. This led to the attendant specification of the Kalman filter as the only approach believed capable of achieving the desired calibration levels in such short times. However, as with all statistical filter methods, success of the Kalman filter depends on the amount of measurement data available in a specified time interval.

As an example, the velocity signal to be used as the basic ground alignment measurement would be corrupted by the velocity errors produced by the effects of wind and cargo loading on the parked aircraft during the alignment period. The time required to complete a single Kalman computational cycle was expected to be about one minute. However, if the standard Kalman formulation, which can accept only a single measurement in each cycle, were used, only about one noisy data sample could be processed each minute. On the other hand, it had already been determined, almost at program inception, that data sampling rates of at least one sample per second were required for adequate noise smoothing*.

To this end, it was proposed simply to pre-average the (once per second) velocity readings over the Kalman interval and use the result for the measurement once per minute. This seemed such an obviously satisfactory solution that it was accepted on the spot, and attention was directed to other areas. Much later in the program this decision was re-examined. A small ensemble error analysis program was constructed to quantitatively evaluate the performance of this averaging process with respect to platform drift rate recovery. The results were unsatisfactory: rather than converging, the solution diverged – to a point far beyond that required to meet minimal system navigation performance goals.

Analysis of this phenomenon led to a simple conclusion. Not only was the signal noise being smoothed (as desired) by the averaging process, but so was the signal itself! This created serious driving error (e.g., platform drift rate) identification problems for the standard [illegible]an formulation, which is based on use of signal samples rather than signal averages.

* This information had been obtained from a small, special-purpose alignment evaluation program, incorporating a simplified, single-axis leveling loop model.

A modification of the Kalman measurement processing formulation, designed to accommodate averages rather than samples, was therefore evidently necessary. This was undertaken, and led to a simple modification of the standard formulation. Finally, a simple ensemble error analysis program was constructed to evaluate the effect of this method. Figure 11* shows a sample of the successful results of compensated averaging, as well as the divergent characteristics obtained with uncompensated averaging.

The modification developed is based on the assumption that one may refer measurements to a fixed epoch (time point) of the state vector†:

$$y(t) = M(t)\Phi(t;t_0)x(t_0) + q(t) . \tag{3.22}$$

If n measurements $y_i(t_i)$ in the interval $t_0 \rightarrow t_0 + \Delta t$ are averaged, it follows that

$$\frac{1}{n}\Sigma y_i(t_i) = \frac{1}{n}[\Sigma M_i(t_i)\Phi(t_i;t_0)]x(t_0) + \frac{1}{n}\Sigma q(t_i) . \tag{3.23}$$

Since $q(t_i)$ is assumed to be a random variable, an averaged measurement can be defined as

$$\bar{y} = \bar{M}x(t_0) + \bar{q} , \tag{3.24}$$

where

$$\bar{y} = \frac{1}{n}\Sigma y_i(t_i)$$

$$\bar{M} = \frac{1}{n}\Sigma M_i(t_i)\Phi(t_i;t_0)$$

$$\bar{q} = \frac{1}{n}\Sigma q$$

$$n = \text{number of measurements.}$$

The technique just described was incorporated into the design. Figure 12 compares this formulation with the normal Kalman formulation.

4. DESIGN EVALUATION AND REFINEMENT

As already mentioned, evaluation of the effect on system performance of many important, specific areas of the basic Kalman filter design lay beyond the capabilities of any existing closed-formula theory, as well as those of the ensemble error analysis Programs I and II. These areas included the single-word floating-point arithmetic, approximations to the elements of the transition and data average observation matrices, the linear approximation to system error behavior, and the like.

Such evaluation could, of course, be left until it could be accomplished with the real system hardware, as soon as this became available. However, not only would this kind of an approach be highly wasteful of precious development time, but also – however carefully designed and monitored – real system tests can rarely provide the kind of environmental control essential to understanding all-mode system performance sensitivities to individual effects such as those which were in question here. The considerations led to the development of a digital simulation of the C-5 aircraft navigation system. This section describes this simulation, and the way in which it was used to evaluate, validate, and refine the filter design.

4.1 Simulator Description

Simulation, as the name implies, is only an approximation to reality. Since the C-5 simulator was, in fact, a program designed for use in a general-purpose digital computer (IBM 7090 or 360), its approximations were embodied in the mathematical models used to simulate real system hardware operation.

Figure 13 is a functional block diagram illustrating the major elements which are required to simulate the navigation system. As illustrated, a flight profile simulator is necessary for generating the true specific force, $\bar{f}_i(t)$, acting on the IMU, and for supplying to the measurement simulation the true position and velocity, denoted by the vector $Z(t)$. In particular, the flight profile simulator was made sufficiently versatile to simulate take-off, landing, cruise, and in-flight maneuvers of the C-5 aircraft.

The IMU simulator accepts the true specific force from the flight plan simulator, and the commanded platform torquing rates from the 1051A airborne computer simulator. The Kalman filter, which is a part of the 1051A simulation, accepts information from the simulated inertial navigation computations and from the measurement simulator, and feeds back corrections to the navigation simulation.

---

* The single-cycle time of only 10 seconds used for generating Figure 11 results was based on a very early (incorrect) assumption about the Kalman filter computational complexity. For corresponding results for cycle times more appropriate to C-5, see Reference 14.

† The notation here and in Figure 12 has been simplified for brevity by means of the symbols

$$x = \bar{x} - \hat{x} , \quad y = Y - \hat{Y} = \Delta y .$$

4.1.1 *IMU Simulation*

The IMU model used in the simulation program is illustrated in Figure 14. The vector $\vec{f}_I$ is the true specific force obtained from the Flight Profile block of Figure 13. The matrix $T_{PI,B}(t)$ is the transformation from the coordinate system in which $\vec{f}_I$ is computed to the platform reference frame. In particular, corruption of the specific force vector by accelerometer-related errors is simulated, as shown across the top of the figure (see also Figure 5), and results in the measured specific force vector, $\vec{f}_0$. The corruption of the commanded torquing rate vector, $\vec{\omega}_c$, by gyro-related errors is simulated as shown in the remainder of the figure (see also Figure 7). In addition to the gyro errors discussed earlier and shown in Figure 7, this more detailed model includes a breakdown of gyro drift rates, $\vec{\omega}_d$, into bias, programmed time-function, and mass unbalance drift rates. Generation of these for the i^{th} gyro ($i = 1,2,3$) is represented in this figure by the box G_i. Generation of the mass unbalance drift rate, in particular, requires resolution of the specific force vector into the misaligned input, output and spin axes of each gyro. This is accomplished for the i^{th} ($i = 1,2,3$) gyro by the matrix T_{PSG_i}, which produces the resolved specific force vector $\vec{f}_{G_i}$. The overall resultant rate vector $\vec{\omega}_p$ is the actual inertial rotation rate of the platform in platform reference coordinates[3].

Updating of the matrix $T_{PI,P}$ is done in a separate portion of the program in accordance with a set of matrix differential equations[3]. In addition to $\vec{\omega}_p$, these equations also require as inputs the components of the inertial rate of rotation of the true locally tangent, ideal wander azimuth angle frame in this frame itself. These components are, in turn, generated by the Flight Profile computations, using $\vec{f}_I$ resolved into this ideal reference frame.

Some of the simplifying assumptions built into this simulation were:

(a) The measured specific force and the gyro torquing rates are continuous.

(b) The inherent, dynamic instrument lags of the gyro and the accelerometers are negligible.

(c) The dynamic lags of the platform servo loops are negligible.

These simplifying assumptions, which were mandatory if the simulation were to run at efficient speeds – i.e., much faster than real time, were motivated by the fact that the principal natural frequencies of inertial system error propagation are extremely low (e.g., Schuler frequency 1 cycle per 84.4 minutes). With respect to these, the lags and sampling rates of the above neglected effects lay in ranges which were, respectively, too short and too high, to effect output errors of interest.

4.1.2 *NDC-1051A Navigation Equation Simulation*

Figure 15 shows some of the details of the computations performed in the 1051A computer. As already discussed, laboratory calibrations are performed on each IMU to establish the elements of the calibration vectors and matrices shown in the block diagram. For example, the null bias calibration will cancel the effects of vector $\vec{a}_b$ of Figure 14, except for errors in the calibration and variations in $\vec{a}_b$ caused by environmental factors.

Under idealized conditions of perfect calibration, with reference to Figures 14 and 15,

$$[a_{ij}] = [(T_{PSA})(SFA)]^{-1} \tag{4.1}$$

and

$$[g_{ij}] = [(SFG)T_{GISP}]^{-1}. \tag{4.2}$$

Also the $[u_{ij}]$ matrix and the constant drift rate calibration vector produce a commanded drift rate. This will exactly cancel the fixed and static mass unbalance drifts of the gyros under the idealized conditions. This type of calibration "marries" the specific IMU and its 1051A computer so that the remaining errors are small. These same matrices are used in the simulation of the C-5 aircraft navigation system. Inputs to the simulator control the assumed calibration errors for any specific case.

The wander azimuth equations are solved by numerical integration. No attempt is made to simulate any part of the actual 1051A arithmetic for the operations shown in Figure 15. Parallel studies in the development of the real-time program had shown that the errors introduced by this assumption are insignificant.

4.1.3 *Kalman Filter Simulation*

Rather than attempting to design an all-mode simulator, different versions of the simulated Kalman filter sub-routine were programmed for each mode under study. This approach was used to reduce complexity and run time. Also, special versions were developed to isolate causes of certain problem areas. For example, versions with and without the 14-bit arithmetic mentioned below, and with and without data averaging, were constructed.

For final validation, the sequence of calculations, the 14-bit floating-point arithmetic, and all other factors were simulated to a high degree of realism. To this end, a special machine language sub-routine was constructed for masking the 7090 word to simulate the 14-bit floating-point arithmetic. This simple sub-routine is used in the appropriate computations to introduce the additional errors caused by the 14-bit floating-point arithmetic. Relative care was used in the order of calculation to preserve as much accuracy as possible. These program details were later used in the real-time program for the same purpose.

All valpes of all the controllable quantities (for example, the constants of the R matrix, the value of the ϵ gain factor, the variance of the random error in observation, etc.) are controlled by input and/or overlay of block data. Observation sequences, Kalman cycle time, etc., are also input in a similar manner.

4.1.4 *Observation Simulation*

The observation simulation models necessary for the various modes of study were made a part of the Kalman filter sub-routine. Models were simulated as realistically as possible to include all the relevant error sources. Those errors, which could only be described by a random process, were included by using a selected sample or samples. This was justified on the basis that the primary purpose of the simulation was to detect and correct major problems. Once problem areas are corrected, the real system should behave in a reasonably linear manner as far as the error propagation is concerned.

4.2 Example Uses of the Simulation Program

It should be reasonably clear from the preceding discussion that the digital computer simulation program is not only quite versatile but can also be readily modified. By means of appropriate input data, values can be selected for quantities controlling

(a) the flight plan,

(b) the initial errors,

(c) the constant and time-varying IMU errors,

(d) the observation sequence and the Kalman cycle time,

as well as many others.

To a novice, such versatility poses a problem of where to begin. This problem can be best addressed by simply reviewing what the program was designed to test. The possible problem areas in the Kalman filter could be expected to fall in the following categories:

(a) Nonlinear phenomena.

(b) Numerical inaccuracies and problems.

(c) Effects of inherent computational delays.

(d) Unmodeled or improperly modeled errors.

(e) Approximations in the equations.

It was therefore planned to set up computer runs designed to exercise the simulated filter in a worst-case sense with respect to each of these categories. Analysis of the resulting data would then uncover problems, if present. Although selection of a worst case is by no means a simple problem, theory can be used to provide sensible guidelines. For example, the basic design is based on linear theory. Hence, one of the first questions is "how linear is this system?". To obtain the answer, the simulator was used to generate time histories of actual navigation system errors for various sets of initial errors. These were compared to errors predicted by the transition matrix in accordance with

$$\hat{x}(t) = \Phi(t;t_0)\hat{x}(t_0) .$$

These runs validated that, for all practical purposes (i.e., within reasonable values for $\tilde{x}(t_0)$), the system errors propagate linearly. Similar checks were made on the measurement functions. These also indicated that the residuals were linearly related to estimate errors for all practical purposes. Hence the linearity assumptions had been proved sound.

In checking the remaining areas, a two-step process was generally adopted:

(a) Demonstration that the filter behaved normally when the only errors in the system were those included in the filter.

(b) Adding unmodeled errors to determine degradation in performance or the existence of otherwise hidden problems.

4.2.1 *Fine Ground Align Mode Description*

The purpose of the fine ground alignment mode is to:

(a) Precisely level the platform. Platform deviations from level are called tilts.

(b) Obtain accurate coincidence between platform and computer azimuths. The difference between the azimuths is called a heading error.

(c) Precisely estimate drift rates of the level channel gyros (gyros 2 and 3).

The fine ground align mode starts after 10 minutes of warm-up, during which time coarse alignment of the platform is accomplished. At the start of fine ground alignment the platform is thus reasonably close to its desired attitude. Tilts are fractions of a degree, and coarse gyro-compassing has reduced the heading error to a few degrees.

Figure 16 is a simplified schematic of the operation of a single level channel during fine ground alignment. As illustrated, the accelerometer on the platform senses a tilt proportional to gravitational force times the sine of the tilt angle ϕ. The velocity register output, v, will ramp as a result of the tilt angle. Because of platform drift rate, an approximate quadratic effect (in time) will also exist, as shown in Figure 17. The total output shown is the sum of the two effects.

The data averaging described earlier samples the velocity register at a high rate to accumulate an average reading. At 45 seconds*, this average is input to the Kalman filter, which computes estimates of the state errors. However, because of the time required for their computation, these estimates are not applied until the end of the second cycle. At this time, the velocity register is reset to a small value by the estimated velocity error, and leveling rate and platform drift rate estimates are also applied.

During the third cycle, the platform is almost perfectly leveled by the commanded leveling rate†. In subsequent cycles, the residual quadratic sawtoothing of the velocity register is steadily reduced as the platform drift rate is more and more accurately estimated and compensated.

During the operation just described, the platform is held in a single, fixed-azimuth orientation which, to facilitate the discussion, is here taken as cardinal (i.e., the platform axes are aligned approximately north and east). In this orientation, the filter after a short time forces the platform into a steady-state, level condition in which, although the input axis north gyro drift rate has been accurately calibrated, the input axis east gyro has not. Instead, the platform (and therefore the input axis of this gyro) has a heading error just sufficient to couple enough of the Earth's rotation rate** into this gyro to cancel its drift rate. Thus, after single-position gyrocompassing, as the above operation is called, both a significant heading error and an east gyro drift rate remain, although the total *platform* drift rate about the east axis is compensated.

To remove these important residual errors, two-position gyrocompassing is employed for C-5 fine ground alignment. That is, after a 6- or 7-minute period in the first position, the platform is slewed 90 degrees in azimuth, and the first-position process is repeated. Because of the reorientation of the level gyro input axes, however, the heading error is removed, and the input axis east (now north) gyro is also calibrated during this second operation.

4.2.2 *Nominal Results of Fine Ground Align Mode Simulation*

The Kalman fine ground alignment technique just described is most easily evaluated when no unmodeled errors are included in the simulation run. This is because unmodeled errors produce extra error effects at the modeled variable levels which unfortunately do not behave in accordance with the filter model. For example, a scale factor error in the input axis north gyro in the first position produces an apparent fixed drift rate about the north axis (in the horizon plane), due to the torquing rate applied to this gyro to counter the Earth's rotation rate in this direction. This is duly compensated by the filter in the first position by means of a fixed re-bias rate. After repositioning, however, this scale factor effect is absent, since no Earth rate compensation torquing is applied to this gyro when its input axis is west. However, the compensating re-bias rate remains to produce an effective drift rate, and therefore a balancing heading error, in the second position.

Since such effects would unnecessarily complicate the evaluation of performance with respect to other specific filter design areas, unmodeled errors were not included in the simulation runs used to evaluate the fine ground alignment filter design until the design had been essentially completed and validated.

In addition, in order to establish a fixed frame of reference for comparisons between runs, a single set of nominal values for modeled errors was used, as well as a nominal alignment configuration. Specifically, all runs simulated two-position fine ground alignment at 45 degrees north latitude, with the input axes of gyros 2 and 3 respectively east and north in the first position, and north and west in the final position. The time spent in each position was about 7 minutes, which, together with the 1-minute slew, accounted for the overall maximum goal of 15 minutes††. As to modeled errors, initial tilts about axes 2 and 3 were assumed to be respectively -0.32 and +0.40 degree, gyros 2 and 3 were each assigned a +0.02 degree per hour drift rate, and the platform heading error was initiated at +2.0 degrees.

Figures 18, 19, and 20 summarize the results of a simulation run with this nominal configuration, after most of the major problem areas had been ironed out. In particular, Figure 18 is the simulated time history of the north velocity register for this run. The ramps due to the large initial platform tilts in each position are evident. The velocity error due to this tilt is largely removed at the beginning of the third Kalman cycle in each position, based on the average of the velocity data collected in the first Kalman cycle. Based on the same data, the initial tilt is also accurately estimated and largely removed by the application of an appropriate platform leveling rate during the third cycle. In subsequent cycles, the drift rate and its tilt and velocity error effects are gradually estimated and removed.

---

* This was the expected duration of the final Kalman cycle time midway in the development. It was reduced to 22.5 seconds later in the program.

† This leveling is reflected in the velocity register readings, whose instantaneous slope is proportional to the instantaneous tilt of the platform.

** Specifically the northward component, in the horizon plane.

†† For investigation purposes, operation in the second position was often extended beyond 15 minutes overall, and alignment performance in inter-cardinal orientations of the platform were also examined.

Although velocity and platform tilt errors are thus quickly removed, calibration of the level gyro drift rates and removal of the heading error is a much slower process. This is illustrated in Figures 19 and 20, from which it is evident, in particular, that significant calibration of the number 2 level gyro drift rate is still proceeding at the end of the allotted 15 minutes of alignment time*.

4.2.3 *Effects of Large Unmodeled Errors*

Early laboratory results for the ground align mode were obtained before the calibration constants of Figure 16 were available. Figure 21 shows a typical time history of the north velocity register obtained from the actual system at that time. Comparison with Figure 18 shows that the behavior of the real system was far from nominal. It was recognized, however, that large unmodeled errors could easily produce such performance. Since the final system would be precalibrated, there was little need for concern if the effects seen in Figure 21 were indeed caused by these errors.

The simulator was used to confirm this hypothesis: Figure 22 shows the north velocity time history obtained from a simulated fine ground alignment, for which large uncalibrated scale factor errors on the gyros were assumed. The unmistakable similarity between the simulated and actual velocity responses in Figures 22 and 21 gave some confidence that the irregularities observed in the real system results were in fact caused by its lack of calibration, rather than any filter design deficiencies. In particular, it is evident that, in each case, the platform was *not* adequately leveled by the command rate applied during the third cycle. In the simulation, this was known to have been caused by the large (10%) gyro scale factor error, which produced a net leveling displacement of the platform during this cycle which was 10% larger than that commanded. The resulting irregular behavior of the simulated system provided a unique signature which was used in this case, as well as throughout later phases of system test, to identify problems due to lack of, or to faulty, system calibration.

4.2.4 *Problems Caused by the 14-Bit Floating-Point Arithmetic*

As mentioned earlier, several versions of the sub-routine which simulated the on-board filter computations were available. This made it possible to address the overall problem in a reasonably orderly manner. For example, it was first established that satisfactory results were obtained with the single precision floating-point 7090 arithmetic (28 bits and sign), before simulating the 14-bit arithmetic.

It was also suspected, on the basis of prior experience, that there would likely be problems in maintaining a positive definite covariance matrix. As described earlier, however, two techniques for combating potential numerical problems and unmodeled errors had been provided[6]. To recapitulate, these two techniques were the addition of the R matrix in the time update operation:

$$P(t) = \Phi P(t_0)\Phi^T + R \,, \tag{4.3}$$

and the "epsilon" technique in the measurement update operations:

$$P_+ = P - PM^TMP/(MPM^T + C) + M^T[\epsilon^2/(MPM^T + C)]M \,. \tag{4.4}$$

Neither of these techniques, however, could provide a satisfactory solution to the numerical problem that occurred as a result of the use of the proposed 14-bit, floating-point arithmetic. This was the loss of positive definiteness (i.e., the occurrence of negative diagonal terms) in the covariance matrix on the first or second filter cycles in the simulated fine ground alignment mode. The problem did not occur initially, when the standard, Kalman discrete-measurement formulation was simulated, but only later, the first time the new, averaged-measurement formulation was tried. This tended to cloud the fundamental source of difficulty. Previous experience had definitely indicated that the most likely problem area would be in the execution of Equation (4.4). The principal changes in this equation introduced by the averaged-measurement formulation were:

(a) the M vector had more non-zero elements;

(b) the C scalar (variance of the random error in measurement) was significantly smaller.

Once these factors were recognized, the value of MPM^T was compared with the value of C. The problem was then immediately recognized. C was so small that early in the sequence (when P is large) it was effectively zero† with respect to addition to MPM^T. Also, the third term of Equation (4.4), which contains the "epsilon" gain, was effectively zero† in comparison to the second term.

Increasing either ϵ or C to prevent the difficulty was not reasonable. The results of such increases would prevent good estimation later in the measurement sequence. Perhaps, in retrospect, an appropriate choice would have been to:

(a) Test if $C < kMPM^T$.

(b) If true, assign $(MPM^T + C) = (1+k)MPM^T$, where the quantity k could be selected somewhere in the range $0.001 < k < 0.1$.

---

* However, the level of calibration achieved in this time was more than adequate. This was established by the excellent inertial navigation performance (not shown) achieved in a simulation of a flight following this alignment.

† i.e., less than the last bit in the 14-bit register.

The actual choice used was to assign

$$(MPM^T+C) = (1.01)(MPM^T) + C ,$$

The 1.01 factor prevented loss of positive definiteness early in the sequence. Later measurements were weighted slightly less than optimally, but the performance loss was small in comparison to that produced by other factors (e.g., unmodeled error sources).

4.2.5 The Approximate Transition Matrix

The transition matrix for the time-varying linear variational equation

$$\dot{x}(t) = A(t)x(t)$$

was to be approximated by the formulation given in Equation (3.19). Two potential problems were evident in this expansion:

(1) Were the second-order terms retained sufficient for adequate prediction across the Kalman interval?

(2) Could the non-constant portion, A_{NC}, of the A matrix be treated as a constant across the interval?

In the ground align mode the matrix A is constant. Hence, one of the early uses of the simulator was to verify the adequacy of Equation (3.19) for use in this mode. This investigation, quickly completed, resulted in the addition of a few more selected second-order (constant) terms to the elements of the original formulation.

In the case of airborne mode filter operation, the question as to whether or not the non-constant elements of A could be treated as constants across each Kalman interval, required a much more extensive investigation with the simulator. This was because results in this area were (not surprisingly) found to depend strongly on (a) the nature and extent of the particular maneuvers assumed, (b) the Kalman cycle time, and (c) the order in which filter operations were carried out (see Table I).

At the time of writing, these investigations were essentially complete, and the constancy assumption had been proved valid, subject to certain conditions on the above factors which were realizable in the C-5 application. However, any detailed discussion of this area, which, it is to be noted, has general importance far beyond the C-5 application, is outside the scope of this chapter.

4.3 Simulated Ground Align/Barometric-Inertial Performance*

Figure 23 shows the gyro calibration history for a simulated fine ground alignment, in which all errors believed to be significant (modeled and unmodeled) were included. In this case, unlike that of Figure 19 where only modeled system errors were included, the final, steady-state, level gyro compensation values attained represent lumped compensations for the combined effects of gyro bias drift rates and the effective drift rates produced by gyro scale factor and input axis alignment calibration matrix errors.

It is evident from Figure 23 that essentially no calibration of the number 1 (azimuth) gyro drift rate is attained in the overall 15-minute fine alignment mode period. This is because the sensitivity of the level velocity register averages, which are the Kalman filter inputs in this mode, to azimuth drift rates is considerably less than it is to level drift rates. This, however, is a limitation, not of the Kalman filter approach, but rather of the basic, two (level) position gyrocompassing technique; no other filter type could (statistically) produce a significant azimuth drift rate calibration within the 15-minute alignment time using the same positioning sequence and geometry.

Fine alignment performance in the presence of the unmodeled gyro errors can only be honestly evaluated by simulation of follow-on barometric inertial flight. The goal (see Section 2) is a barometric-inertial navigation CEP growth rate during ten-hour flights of no larger than 0.75 nautical mile per hour during the first five hours and 1.25 nautical miles per hour in the second five hours.

Figure 24 gives the navigator performance in terms of the actual east and west position errors that accumulated for this particular set of unmodeled gyro errors. Shown also are the "ensemble" results taken from the Kalman filter covariance matrix. This was a simulation of a barometric-inertial flight, starting at 45 degrees latitude, accelerating east to 600 ft/sec horizontal velocity, climbing to 33,000 feet altitude, and thereafter cruising at constant altitude and speed on a great circle path. It is seen in the figure that the actual east position error exceeds the standard deviation obtained from the Kalman filter covariance matrix. This was not unexpected because:

(a) The 16-by-16 Kalman filter covariance matrix does not include all the error sources.

(b) The sample errors chosen were larger than would, on average, be obtained from a statistical sample.

The overall trends of the two curves (actual and standard deviation) have the same behavior. This gives confidence that the simplifications used in the filter are appropriate for long-time propagation.

---

* The vertical channel of a three-degrees-of-freedom free-inertial navigator is unstable. For the C-5 navigator, stability is maintained by using the generalized all-mode Kalman filter with an altitude observation based on the barometric altimeter. Therefore, this mode of operation is referred to as "barometric-inertial" instead of free-inertial.

Previous ensemble analysis had shown that straight flights produce slightly worse barometric-inertial performance than flights which have turns. It was therefore decided to establish a flight that included a turn, as a further test for the filter in the barometric-inertial mode.

The path chosen was a level acceleration in the north direction to 600 ft/sec, simultaneous with a climb to 33,000 feet. This was followed by cruise for 20 minutes, then a 90-degree right turn taking 60 seconds, and finally a long terminal great circle cruise on the new leg at 600 ft/sec. Figure 25 gives the navigation results for these conditions. Alignment and error conditions were identical to those used in obtaining the results of Figure 24.

As can be seen in comparing Figures 24 and 25, the actual errors are slightly smaller for the latter. This is a (one-case) confirmation of the data obtained from the ensemble analysis programs.

An interesting property of the results of Figures 24 and 25 is that sustained oscillations at both earth rate frequency and Schuler frequency are contained in the errors. This is particularly evident in the north position error time histories. The approximate error model used in the Kalman filter reflects these same oscillations and retains a fairly good phase relationship. This gives confidence that the entire Kalman filter error covariance matrix does remain a good model of the actual sets of errors. This is not too important for fine alignment followed by otherwise unaugmented barometric-inertial operation in the subsequent flight. If, however, additional data such as a LORAN position fix becomes available later in the flight, correct weighting of the observation to improve navigation and to calibrate the IMU depends on the validity of the covariance matrix.

4.4 Laboratory Test of Ground Align/Barometric-Inertial Performance

Reliable confirmation of the design of the C-5 Kalman filter in the fine ground align mode of operation became possible for the first time only when the IMU precalibration program checkout was completed, and this program became available for use in determining the gyro and accelerometer error compensation matrix elements (see Section 4.1).

This allowed laboratory system fine align tests to be conducted for the first time with calibrated IMU's. Since filter design had been predicated from the start on the use of such precalibrated IMU's and quite a few simulation results incorporating this assumption were already available, extensive, direct, and meaningful comparisons between actual and simulated performance could now be carried out.

Figures 26 and 27 show, respectively, the east velocity register and the drift rate compensation histories during typical laboratory system fine align tests with calibrated IMU's at this stage of the program.

Comparison of Figures 26, 21, and 18 reveals that IMU calibration had removed the mis-modeling difficulties, the effects of which are evident in Figure 21, so that velocity register behavior of the calibrated system was very similar to that shown in Figure 18 for a simulated, *perfectly* precalibrated system.

Comparison of Figures 27 and 23 further substantiates that the (16-variable) filter-incorporated system error model was adequate, as indicated by the highly similar simulated and laboratory system test results shown for the gyro drift rate compensation histories during ground align.

In addition, the laboratory test depicted in Figure 27 was extended (as shown) to establish azimuth drift rate (gyro 1) compensation behavior if more alignment time were available in a real C-5 operational situation*. It was evident that, despite the weak sensitivity involved in extracting azimuth drift rate information from level velocity averages, complete calibration could nevertheless be accomplished in about half an hour†.

Finally, Figure 28 shows the inertial navigation performance of a typical laboratory system following nominal (i.e., 15-minute, two-position) fine ground align. The excellent performance (relative to the performance goal of 0.75 nautical mile per hour for the first five hours (see Section 2.3)) shown here was a strong overall confirmation, beyond that obtainable from operation in the fine align mode only, of filter design adequacy.

4.5 Flight Test of Ground Align/Barometric-Inertial Performance

Although results (such as that shown in Figure 28) clearly demonstrated adequate performance of the C-5 filter design for ground align in the laboratory, and proved a high degree of confidence for its expected performance in this mode in the field, *demonstrated* field capability could, of course, only be obtained in actual flight test.

At the time of writing, performance results obtained from the first series of flight tests have confirmed not only the adequacy of filter design for fine ground alignment, but for barometric-inertial navigation as well.

In addition, these results have given a high degree of confidence that the design will also prove adequate in the remaining system modes of filter operation, which will be evaluated in the test program phases now being entered.

---

* Calibration of the azimuth, as well as the level gyro drift rates, provides even better subsequent inertial navigation.

† This had been predicted by earlier simulation results (not shown).

5. SUMMARY

The C-5 Kalman filter software development effort, which has been outlined in this chapter, resulted in a final design which was optimum in the practical sense that it

(a) met the software performance specifications;

(b) fitted into the real-time, on-board computer;

(c) met the pertinent, overall software development schedules.

To the extent possible, modern sample-data estimation and control theory was not only used in developing the design, but in some areas actually extended (e.g., the "epsilon" and averaged-measurement techniques). However, as is typical of most practical problems, in a few areas, because theory was inadequate or development time limited, approximate solutions had to be constructed, based on engineering judgment and experience. Much of the development effort was devoted to the construction and use of the three major error analysis and system simulation programs (Programs I, II, and III). These programs were used to validate or reject alternatives at almost every major filter design decision point throughout the program. Costly redesign and tape reprogramming efforts late in the system development were thereby avoided. Program III, in particular, also proved an invaluable tool for isolation and correction of both hardware and software "bugs" during system test.

6. FUTURE TRENDS

It is only in the past few years that digital computers have become available which are sufficiently fast, capacious, light and small to render feasible real-time on-board solutions of problems the size of that mechanized in the C-5 computer. The rapid advance in technology which created these computers continues: next-generation models, with vastly improved capabilities, will soon be available. In particular, the use in such systems of advanced, optimal, information processing techniques like Kalman filtering, which were hitherto ignored because of their attendant memory and real-time cost, have now become feasible. Further, the superior performance available in practice from such techniques, with respect to prior methods, ensures their future as the eventual *standard* techniques in such applications. This future will be all the more rapidly realized, however, if better solutions to some of the significant current problems attending the actual mechanization of such techniques are reached. With respect to Kalman filter mechanizations, two principal current problems are:

(a) *The Need for an Extremely Accurate System Description.* A Kalman filter must incorporate both a model of the dynamic behavior of the system to which it is applied and the statistics of the noise inputs forcing the system. Mechanized performance depends on the fidelity of the mechanized model to the real system, and on the accuracy with which the assumed statistics characterize the actual random input processes. Such information cannot in general be easily or cheaply obtained in the course of a real system development, where time and funds are limited. Further work on appropriate modifications of the Kalman gain equations, which would provide a high degree of insensitivity to system and input mis-modeling or under-modeling, is necessary*.

(b) *Round-off Error Sensitivity.* The standard recursive Kalman formulation for determining optimal measurement weighting involves the use of essentially squared, covariance matrix information. Round-off errors associated with the fixed computer word length, which are introduced at each computational step, produce an unduly rapid rate of weighting error build-up, when compared with what could be obtained with a formulation based on *unsquared* information. Although several such formulations have recently been advanced, more work is needed to establish whether these do not cost more in attendant speed and memory requirements than they gain in accuracy relative to the standard formulation.

---

* The "epsilon" technique, as applied to the C-5 filter, is one of the most attractive such modifications yet developed. However, even this technique requires much more work to establish a firm general basis for choosing parameter values.

REFERENCES

1. Stein, Kenneth J. — *Self-Contained Avionics Broaden Scope of C-5 Missions.* Aviation Week and Space Technology, Vol.87, No.21, November 20 1967, pp.192-202.

2. Kalman, R.E. — *A New Approach to Linear Filtering and Prediction Problems.* American Society of Mechanical Engineers Transactions, Journal of Basic Engineering, Vol.82, Series D, No.1, 1960, pp.35-45.

3. Broxmeyer, C. — *Inertial Navigation Systems.* McGraw-Hill, New York, 1964.

4. Smith, G.L. et al. — *Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle.* NASA Technical Report R-135, Washington, DC, 1962.

5. Schmidt, S.F. — *State Space Techniques Applied to the Design of a Space Navigation System.* Proceedings of the 1962 Joint Automatic Control Conference, Session 11, Paper 3.

6. Leondes, C. (Editor) — *Advances in Control Systems.* pp.1-68, 219-339 in Vol.3, Academic Press, New York, 1966.

7. Schmidt, S.F. — *Estimation of State with Acceptable Accuracy Constraints.* Analytical Mechanics Associates, Inc., Palo Alto, California, Technical Report 67-16 (Contract NAS 5-11048), 1967.

8. Schmidt, S.F. — *Compensation for Modeling Errors in Orbit Determination Problems.* Analytical Mechanics Associates, Inc., Palo Alto, California, Technical Report 67-168 (Contract NAS 5-11048), 1967.

9. Fagin, S.L. — *Recursive Linear Regression Theory, Optimal Filter Theory, and Error Analyses of Optimal Systems.* Institute of Electrical and Electronic Engineers, International Convention Record, Part I, Session 30, 1964, pp.216-240.

*10. Lukesh, J.S. — *Simulation of the Nortronics Kalman Filter-Augmented C-5 Inertial-Doppler Navigator.* Proceedings of the National Aerospace Electronics Conference (NAECON), Dayton, Ohio, May 1968, pp.1.1-180.

*11. Sekiguchi, J. Berg, R. — *The C-5 Navigation Computing System.* Proceedings of the National Aerospace Electronics Conference (NAECON), Dayton, Ohio, 1968, pp.189-194.

*12. Miller, B.J. — *Information Processing in the C-5 Navigation Computer.* Proceedings of the ION National Space Meeting on Simplified Manned Guidance, Navigation and Control, Cocoa Beach, Florida, 19-21 February 1968, pp.255-280.

AUTHORS AND REFERENCE AUTHORS

Berg, Richard G.	Northrop Corporation, Electronics Division, Hawthorne, California.
Broxmeyer, Charles	Raytheon Co., Massachusetts.
Fagin, Samuel L.	Sperry Gyroscope Co., Division of the Sperry-Rand Corporation, Great Neck, NY.
Gunckel, Thomas L. II	Autonetics Division, North American Rockwell Corporation, Anaheim, California.
Kalman, R.E.	Research Institute for Advanced Study, Baltimore, Maryland.
Leondes, C.T.	Department of Engineering, University of California, Los Angeles, California.
Lukesh, John S.	Northrop Electronics Division, Hawthorne, California.
McGee, L.A.	NASA Ames Research Center, Moffatt Field, California.
Miller, Bruce J.	Northrop Electronics Division, Hawthorne, California.
Samuelson, H.R.	Northrop Electronics Division, Hawthorne, California.

* References 10 – 12 are not directly referred to in the text; however, they are included for general information.

Schmidt, Stanley F.	Analytical Mechanics Associates, Palo Alto, California.
Sekiguchi, Joh	Northrop Electronics Division, Hawthorne, California.
Smith, G.L.	NASA Ames Research Center, Moffatt Field, California.
Sorenson, H.W.	AC Electronics Division, General Motors Corporation, El Segundo, California.
Stein, Kenneth J.	Aviation Week and Space Technology, McGraw-Hill Co., New York, NY
Steybe, D.E.	Northrop Electronics Division, Hawthorne, California.
Weinberg, John D.	Northrop Electronics Division, Hawthorne, California.

TABLE I

Kalman Filter Equations

Operation	*Matrix formulation*
Observation Residual	$\Delta y = Y - \hat{Y}$
Estimator Control	$\hat{x} = f(\hat{x})$
Observation Update	$Q = MPM^T + C$ $b = (P + \epsilon I)M^T Q^{-1}$ $\hat{x} = \hat{x} + b(\Delta y)$ $P = P - PM^T Q^{-1} MP + M^T \epsilon^2 Q^{-1} M$
Transition Matrix Calculation	$\Phi(\Delta t) =$ Equation (3.19)
Time Update	$\hat{x} = \Phi(\Delta t)\hat{x}$ $P = \Phi(\Delta t) P \Phi^T(\Delta t) + R(\Delta t)$
System Control	$u_{sc} = g(\hat{x})$

Quantity and definition	*Dimensions*
$\hat{x}$ = system error estimate vector	n × 1
P = estimate error covariance matrix	n × n
Φ = system error transition matrix	n × n
R = system noise covariance matrix	n × n
Δy = observation residual	1 × 1
Y = observation	1 × 1
$\hat{Y}$ = observation estimate	1 × 1
M = measurement matrix	1 × n
C = observation noise	1 × 1
Q = observation residual covariance matrix	1 × 1
b = weighting vector	n × 1
I = identity matrix	n × n
ϵ = "epsilon" factor	1 × 1
u_{sc} = system control vector	n × 1
$f()$ = estimator control algorithm; applies estimate to system	
$g()$ = system control algorithm; computes corrections to be applied to system	
Δt = period of Kalman cycle	
n = state vector dimension	

TABLE II

Overall C-5 Error Model

Error source	*Model*	
	Complete	*Mechanized*
Inertial Measurement Unit	50	13
Doppler		
Groundspeed	3	1
Drift Angle	3	1
Vertical Velocity	3	0
Altitude Reference Error		
Barometric	2	1
Radar Altimeter	3	0
Position Reference Error		
Horizontal Position	2	0
Heading Reference Error		
AHRU	1	0
Totals	87	16

TABLE III

C-5 Inertial Navigator Error Model

Error source	*Model*	
	Complete	*Mechanized*
Velocity Error	3	3
Position Error	3	3
Platform Attitude Error	3	3
Gyro Drift Rate Error		
Correlated	3	3
Bias	3	0
Scale Factor	3	0
Ramp	3	0
Mass Unbalance	6	0
Input Axis Alignment	6	0
Accelerometer Error		
Correlated	3	1
Bias	3	0
Scale Factor	3	0
Input Axis Alignment	6	0
Ball Readout	2	*
Totals	50	13

* A single state element is used to model IMU readout/Doppler drift angle azimuth error.

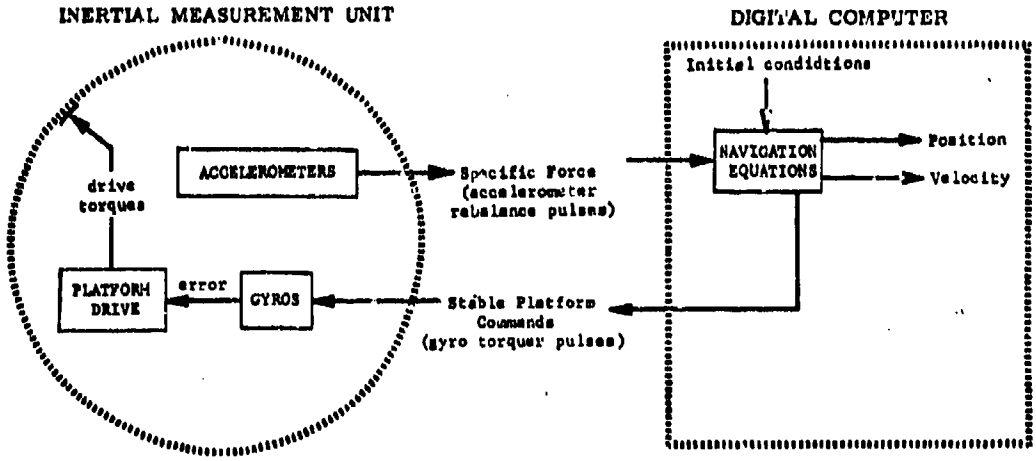


Fig. 1 Simplified block diagram of an inertial navigation system

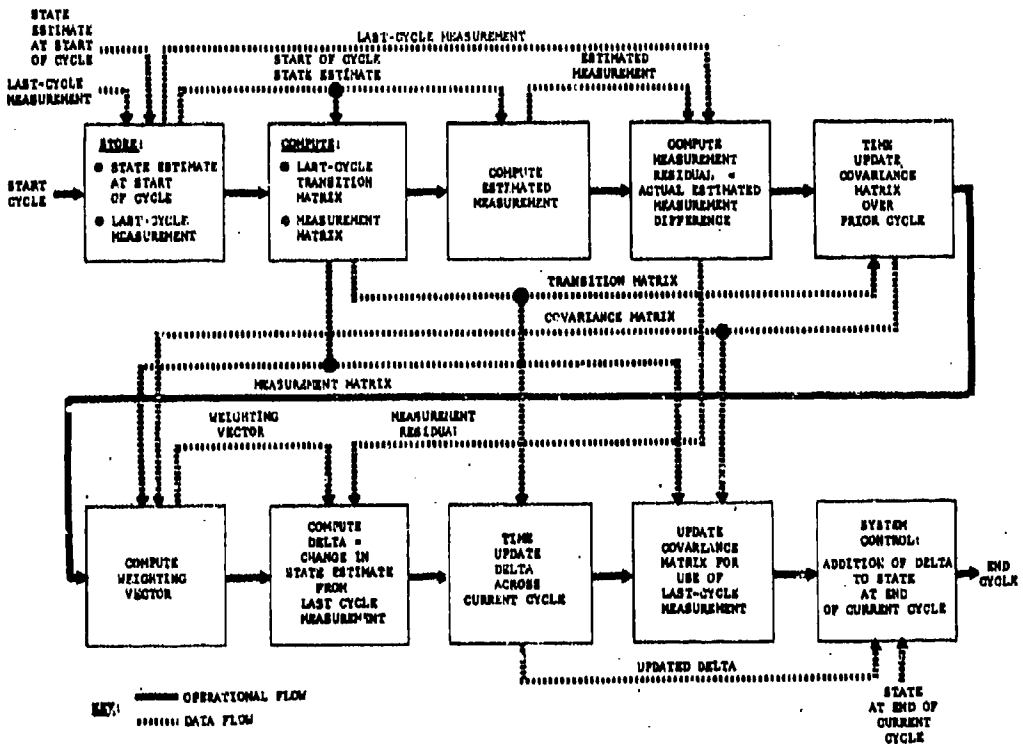


Fig. 2 Block diagram of Kalman filter operation sequence

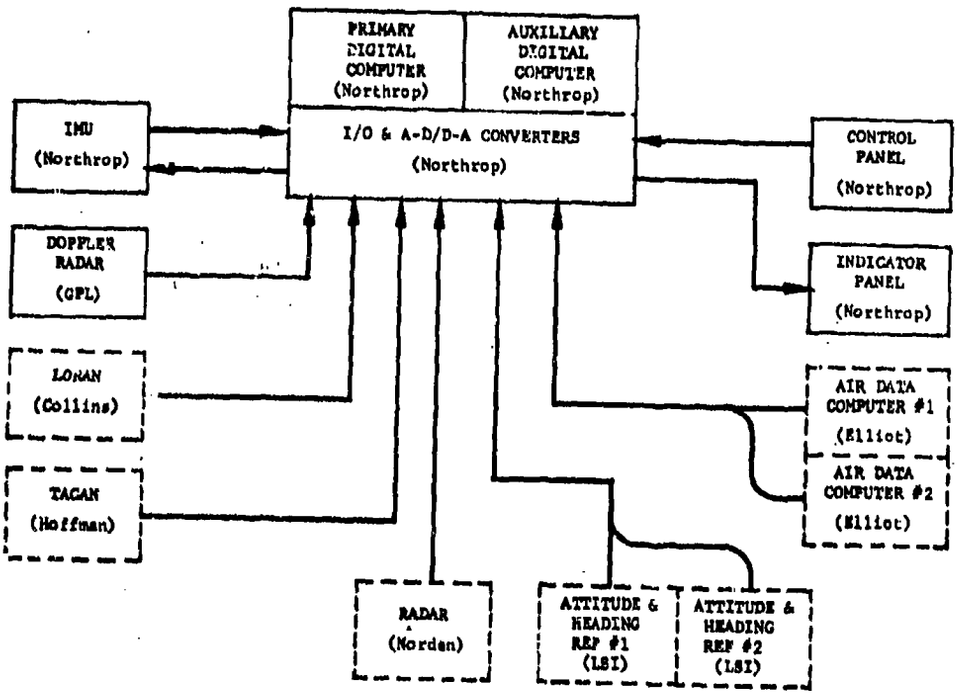


Fig.3 Block diagram of C-5 navigation system equipments

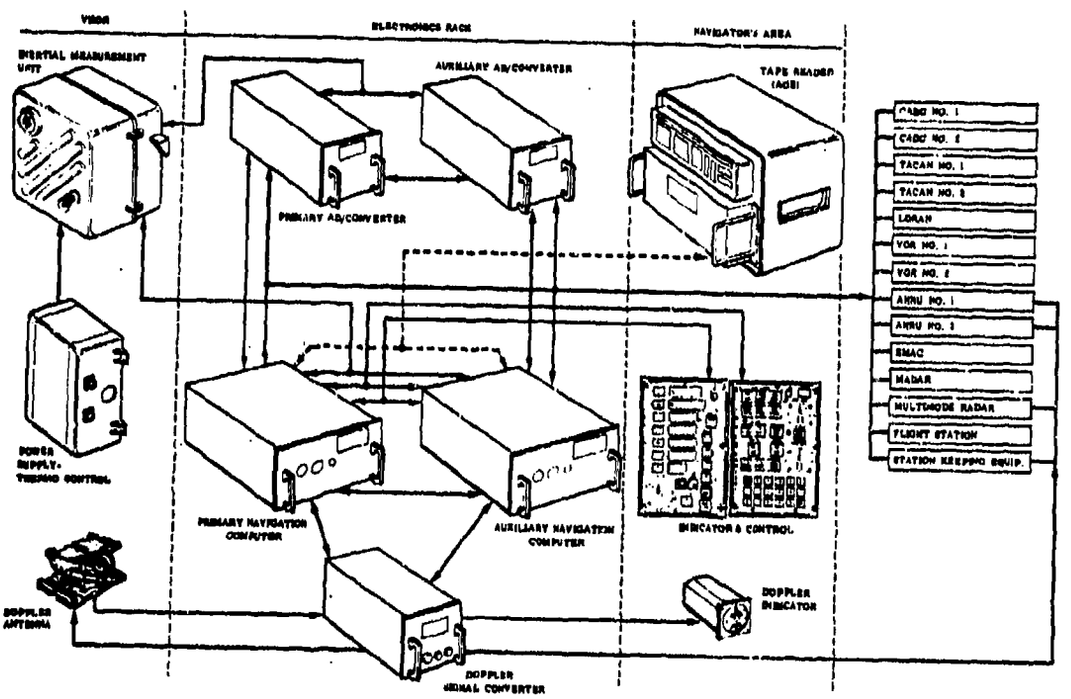


Fig.4 C-5 inertial-Doppler navigation equipments

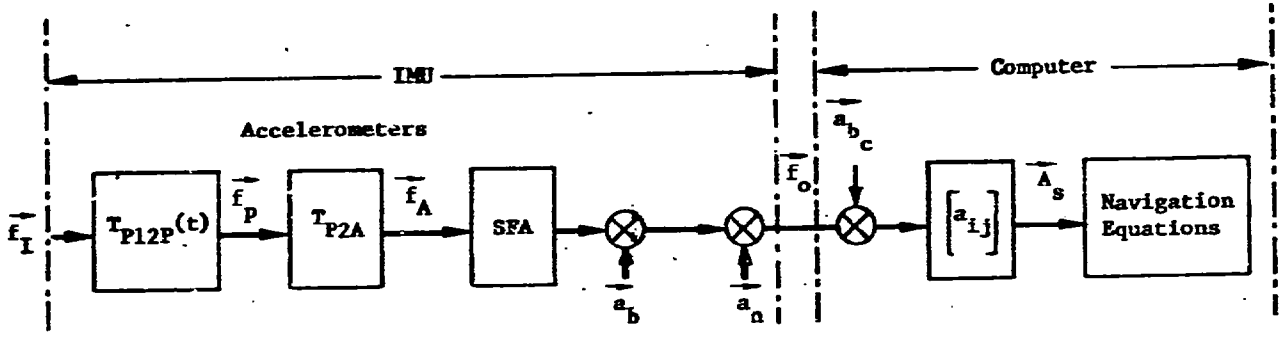


Fig. 5 Accelerometer signal processing

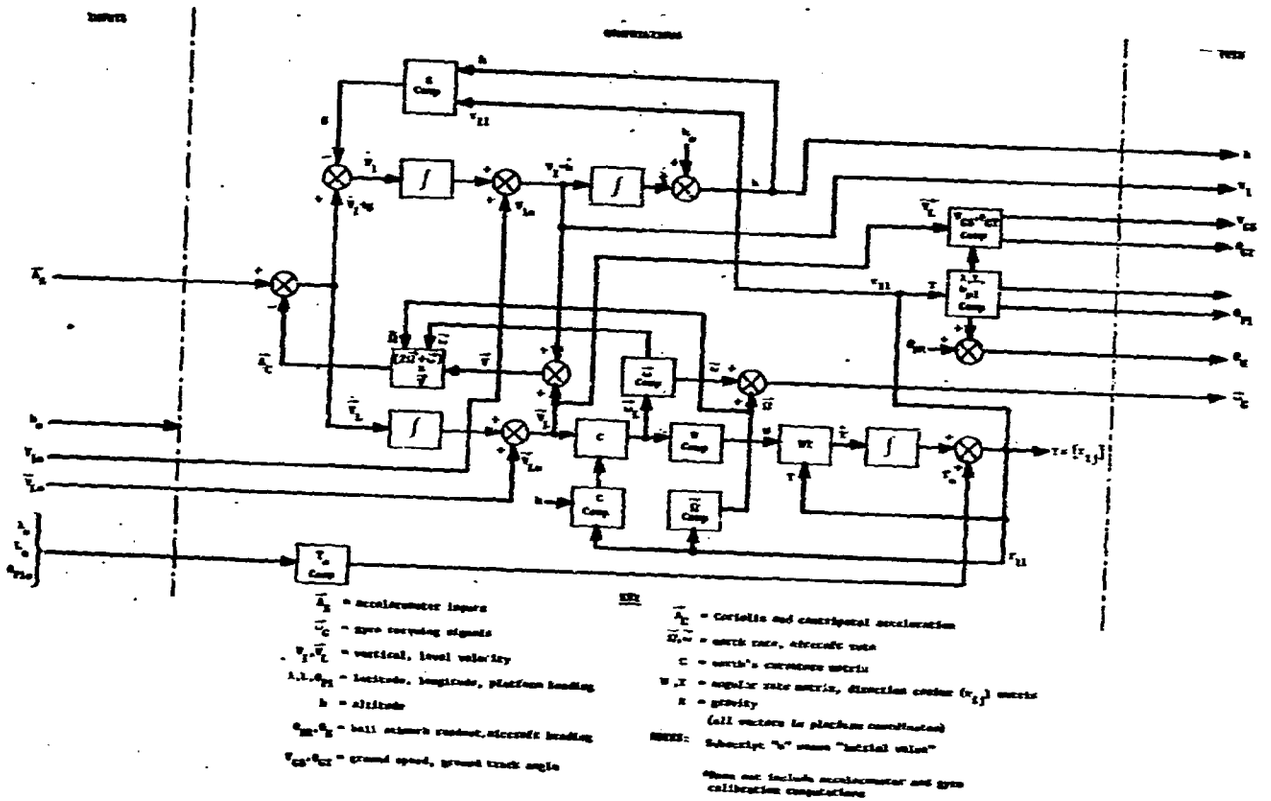


Fig.6 Simplified block diagram of free-inertial mechanization (excluding accelerometer and gyro calibration computations)

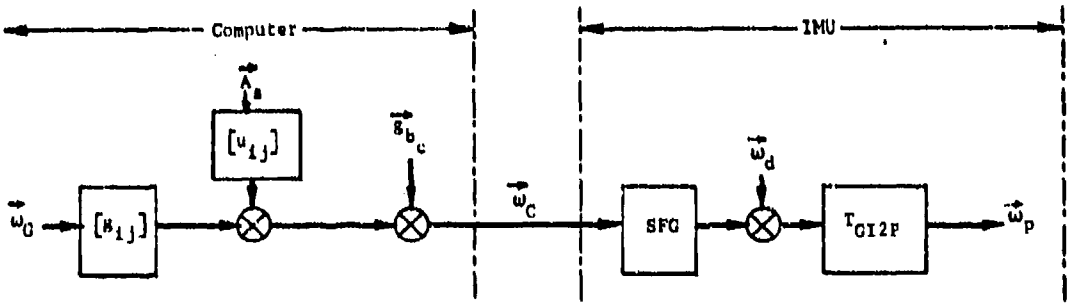
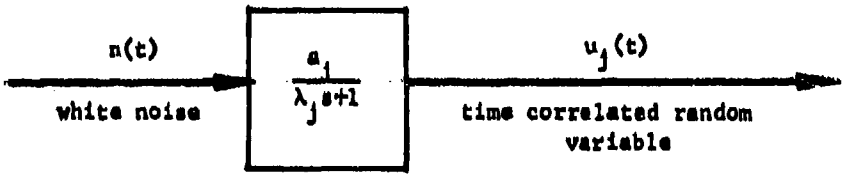


Fig.7 Gyro signal processing



where:

a_j = gain constant assumed for particular variable

λ_j = time constant assumed for particular variable

s = Laplace variable

Fig.8 Random noise model

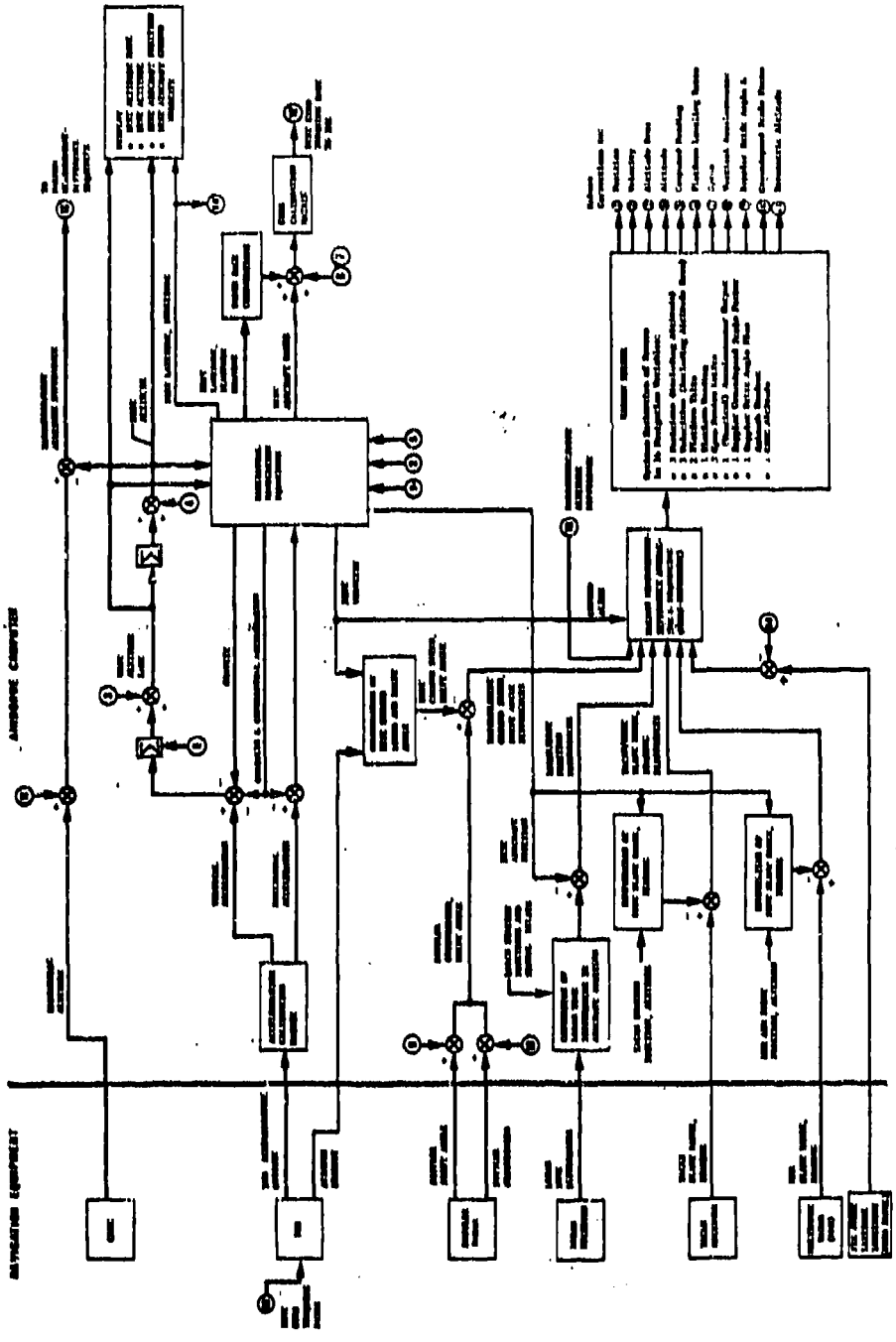
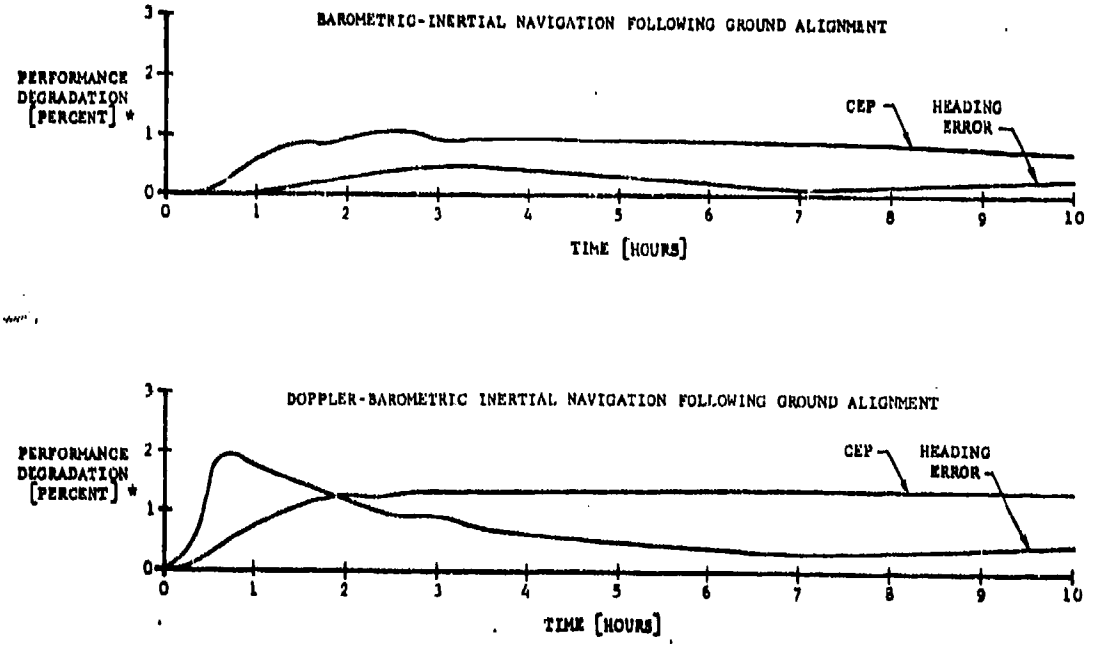


Fig.9 Functional block diagram of C-5 Kahan navigation/alignment mechanism



*Percent Degradation of 16-Variable Relative to 67-Variable (Complete) Mechanization

Fig. 10 C-5 ensemble performance simulation: 67- against 16-variable mechanizations

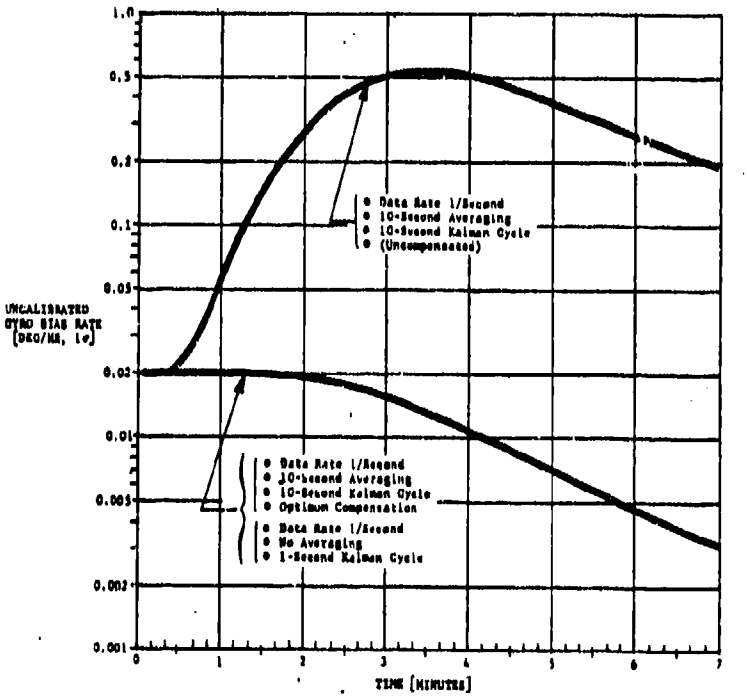
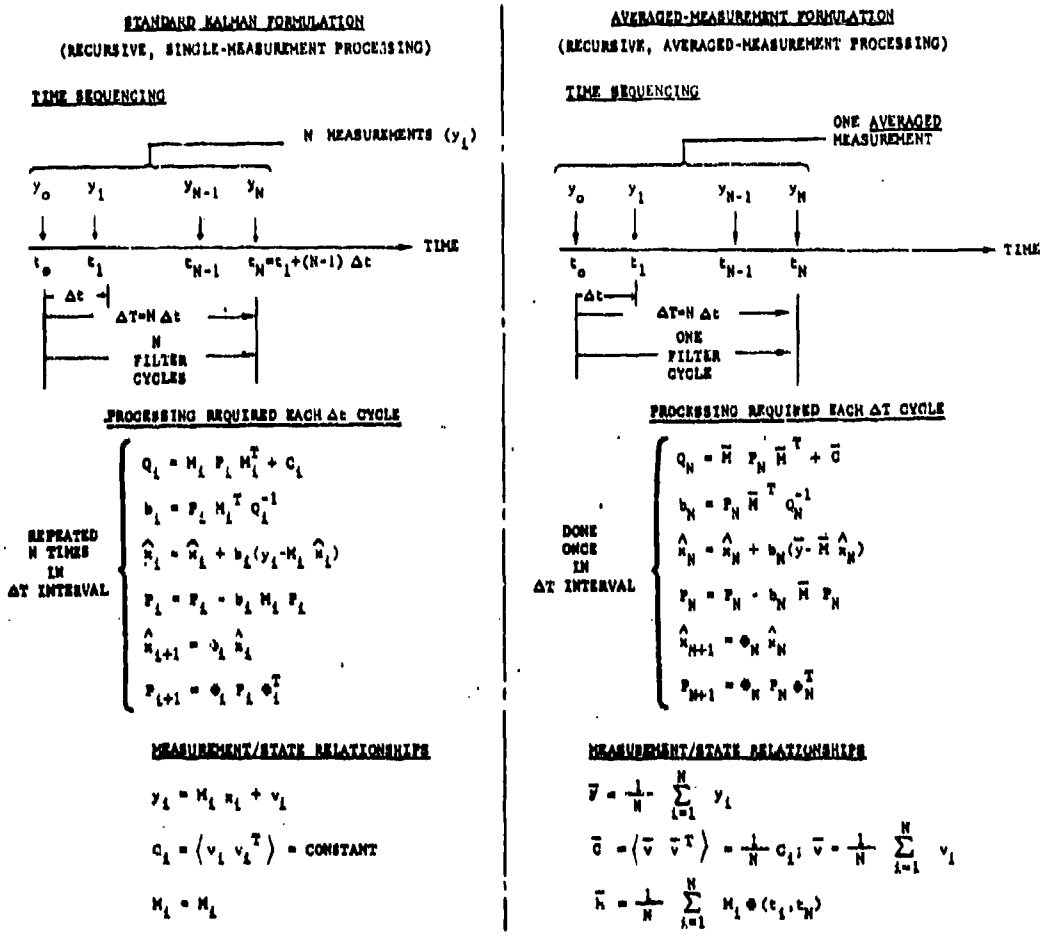
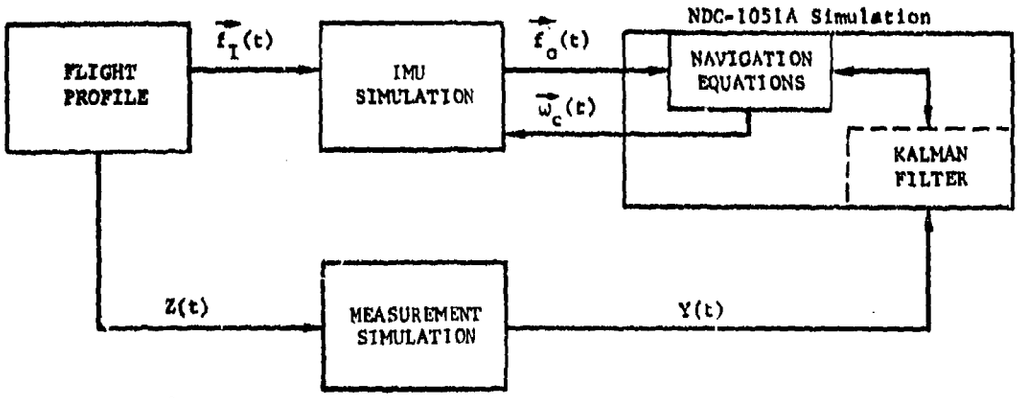


Fig. 11 Effect of optimum data averaging compensation during ground alignment



*EXCLUDES SYSTEM FORCING NOISE TERMS

Fig. 12 Simplified comparison of averaged-measurement and standard Kalman filter formulations (excluding system forcing noise terms)



$\vec{f}_I(t)$ = Specific force vector acting on IMU

$\vec{f}_o(t)$ = Measured specific force vector (by accelerometers)

$\vec{\omega}_c(t)$ = Vector of commanded gyro torque rates

Z(t) = Vector of true position - velocity, etc.

$\vec{Y}(t)$ = Vector of simulated measurements

Fig. 13 Functional block diagram of C-5 navigation system simulation

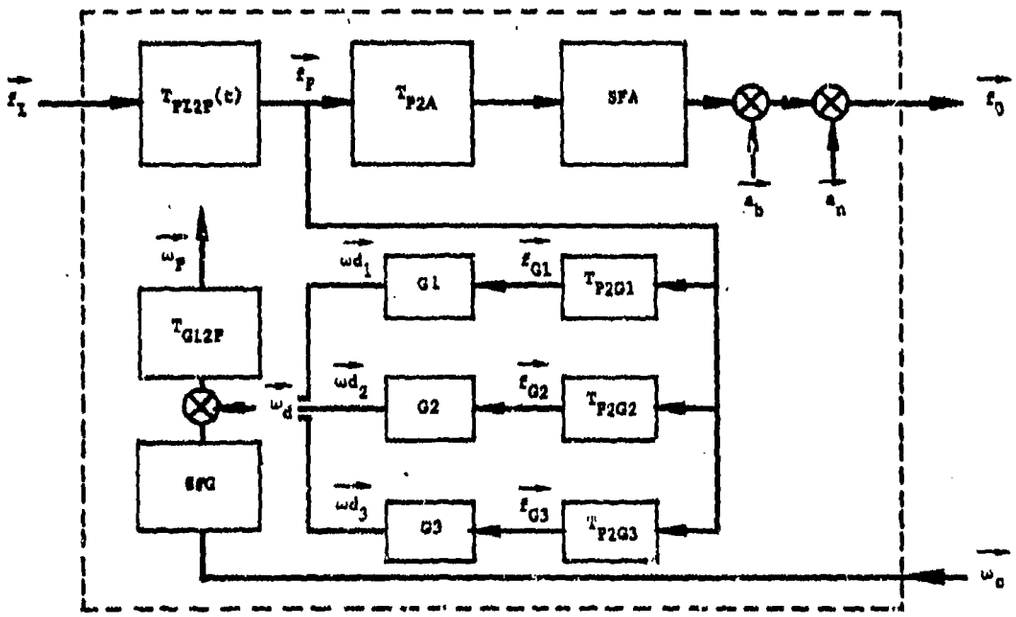


Fig. 14 Functional block diagram of the IMU simulation model

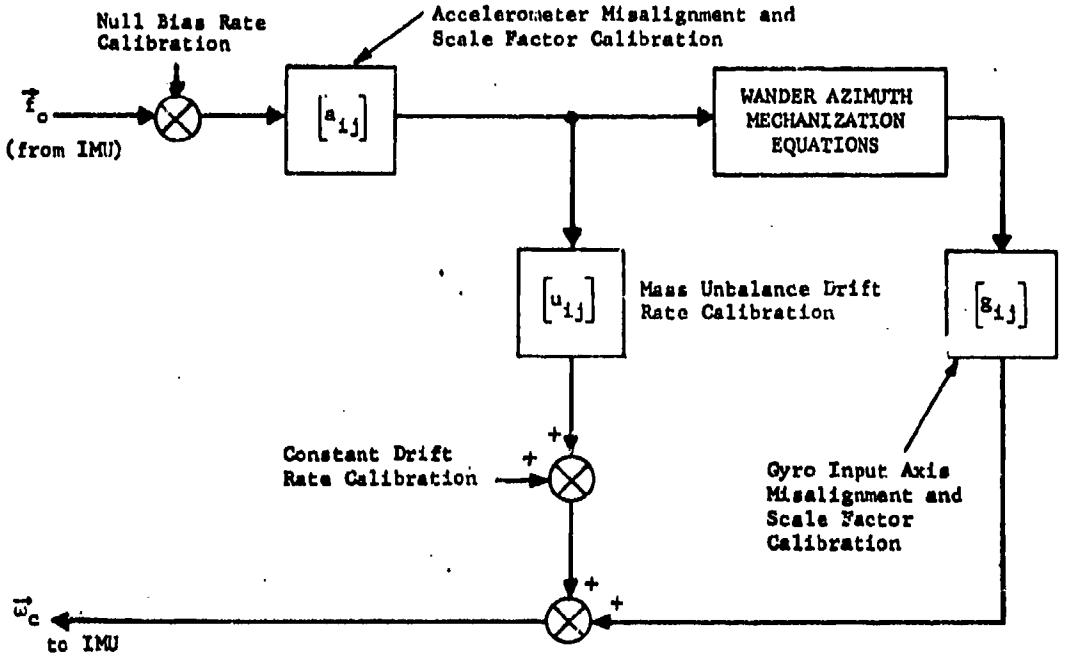


Fig. 15 Functional block diagram of NDC-1051A free-inertial mechanization

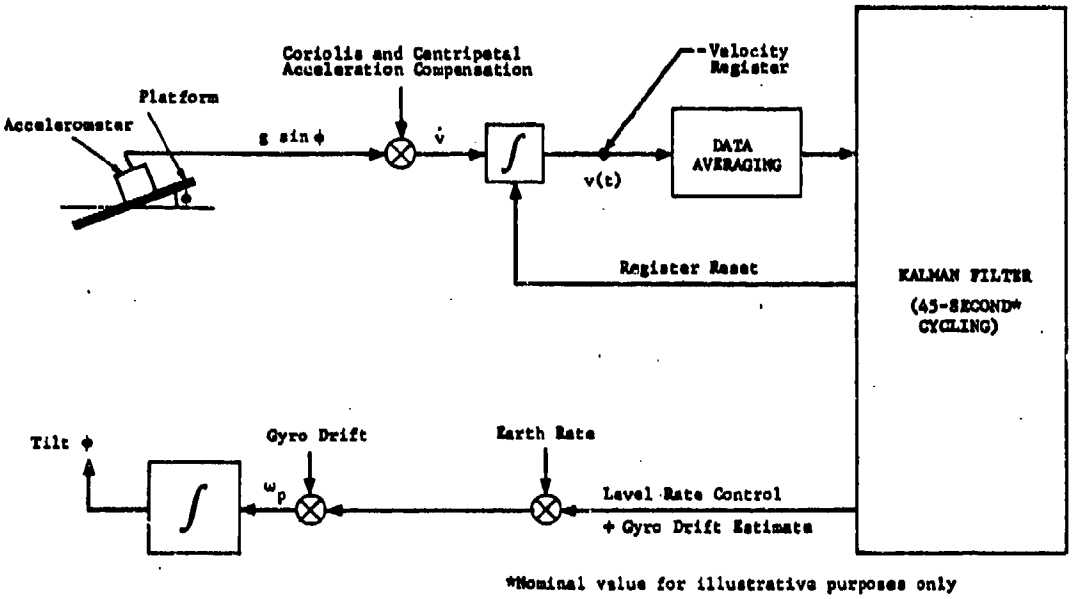


Fig. 16 Single-channel schematic of fine ground align mode

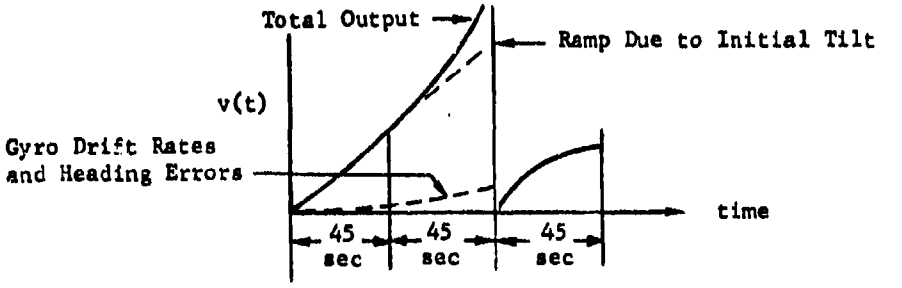


Fig. 17 Typical single-channel velocity history in fine ground align

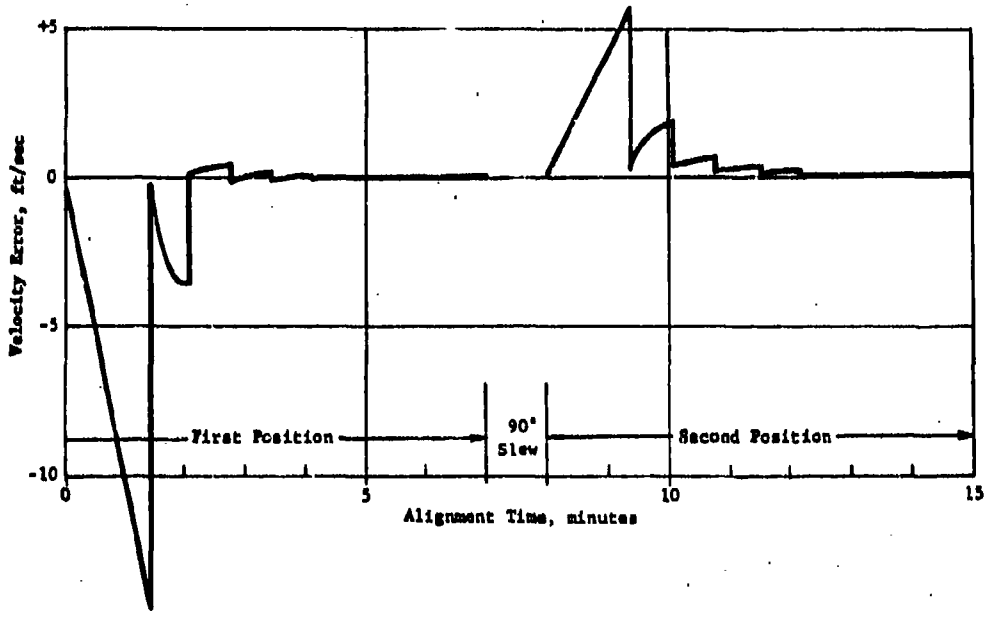


Fig. 18 North velocity history for simulated two-position fine ground alignment (modeled errors only)

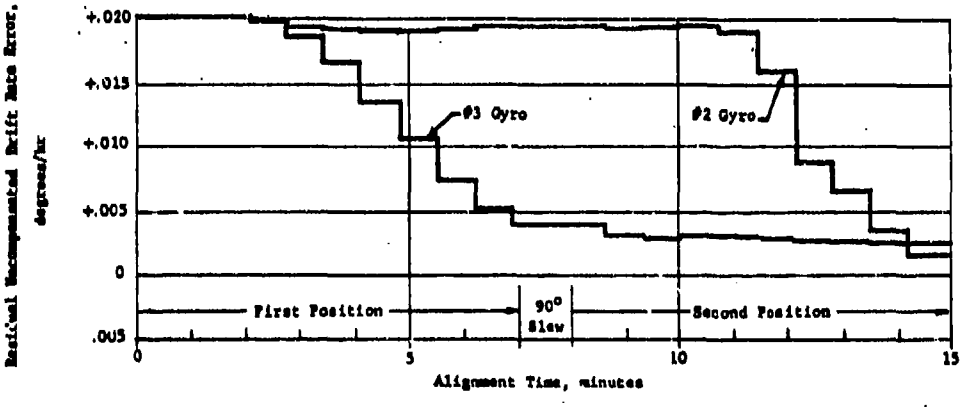


Fig. 19 Level gyro drift rate error histories for simulated two-position fine ground alignment (modeled errors only)

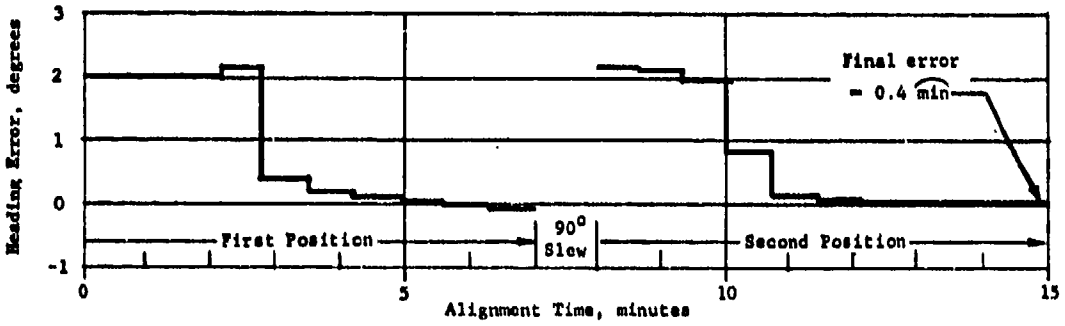


Fig. 20 Heading error history for simulated two-position fine ground alignment (modeled errors only)

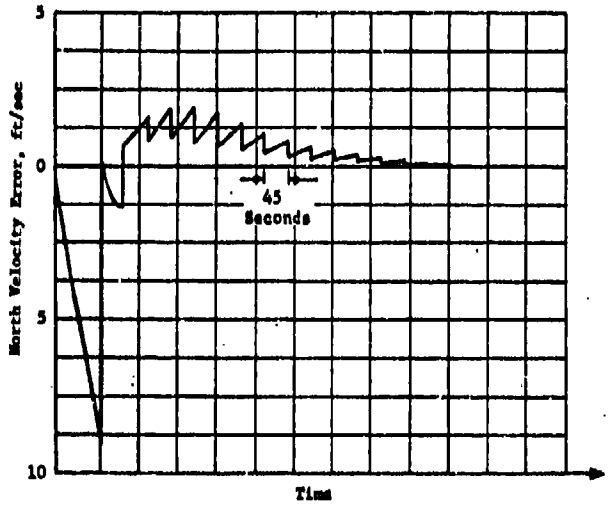


Fig. 21 Actual north velocity history of an uncalibrated system (early, single-position, laboratory fine ground alignment test)

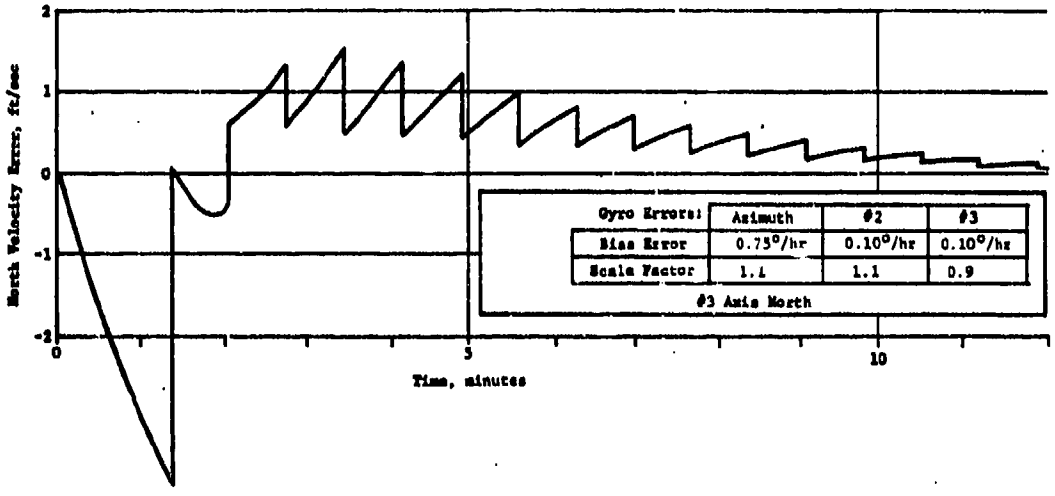


Fig. 22 Simulated north velocity history of an uncalibrated system

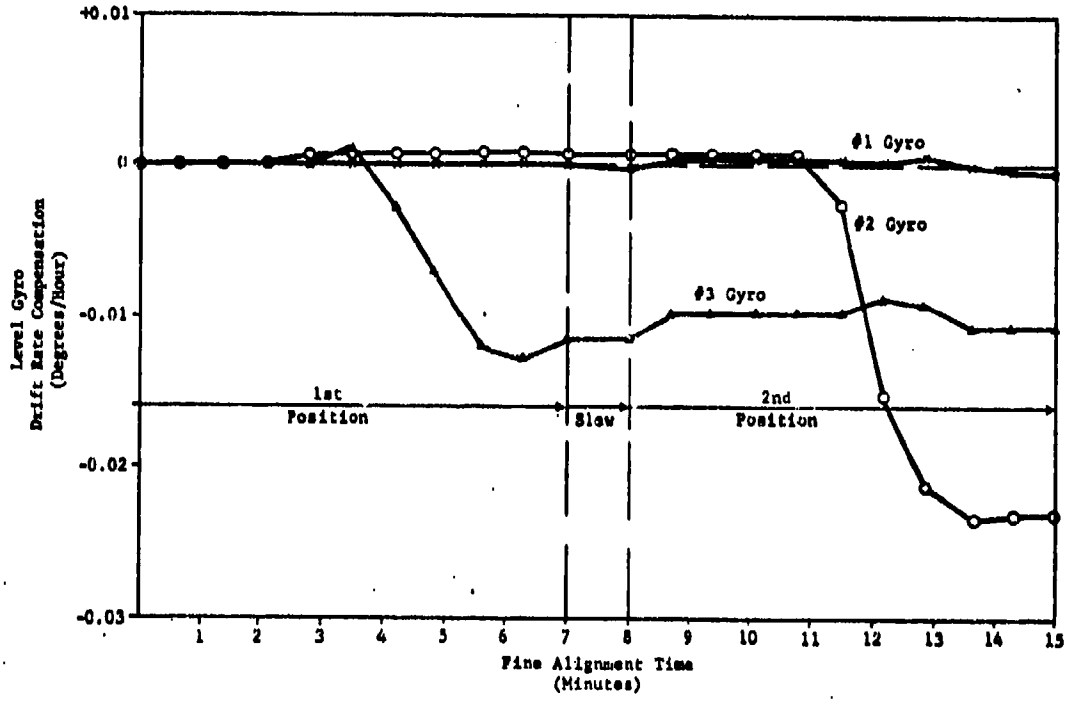


Fig. 23 Simulated ground align gyro calibration history (unmodeled system errors included)

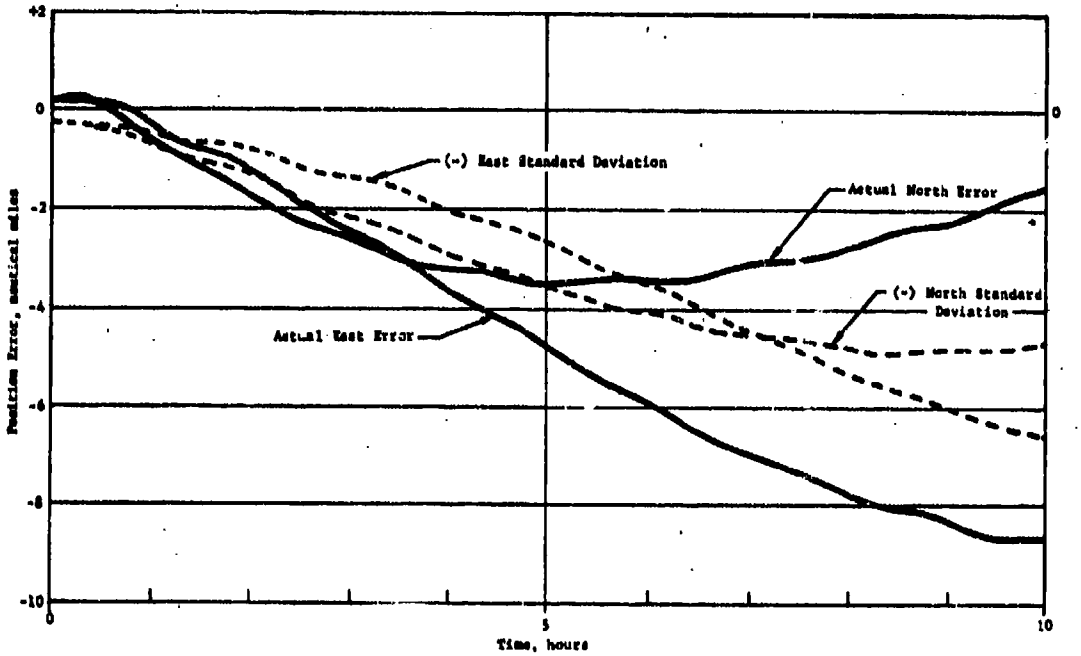


Fig. 24 Simulated barometric-inertial position error time histories (great circle flight)

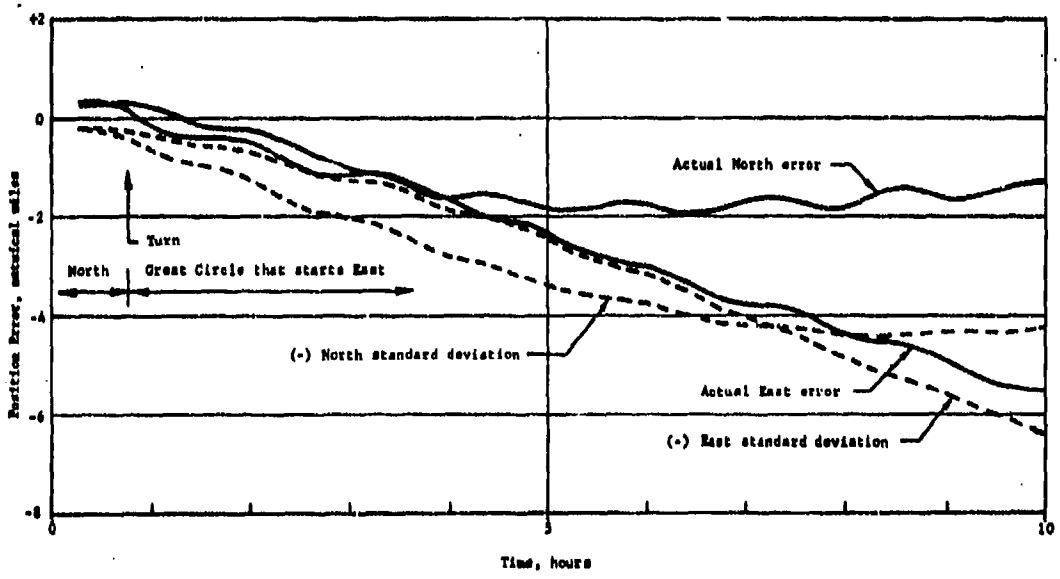


Fig. 25 Simulated barometric-inertial position error time histories (flight with two great circle legs)

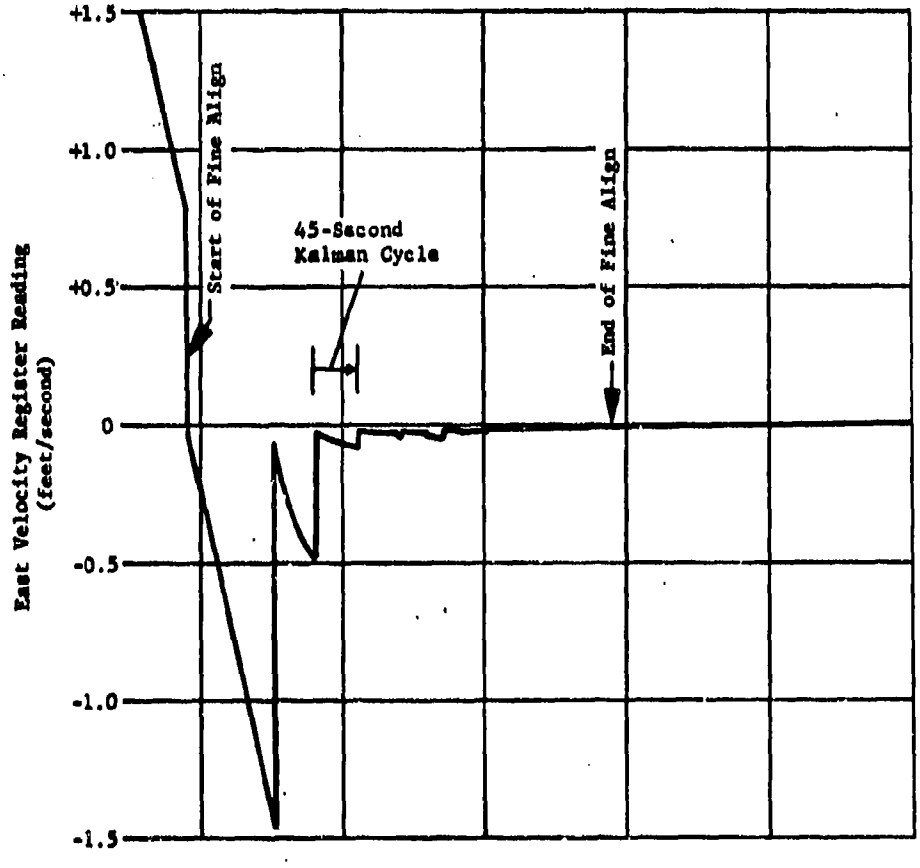


Fig. 26 East velocity history for single-position fine ground alignment in laboratory system test

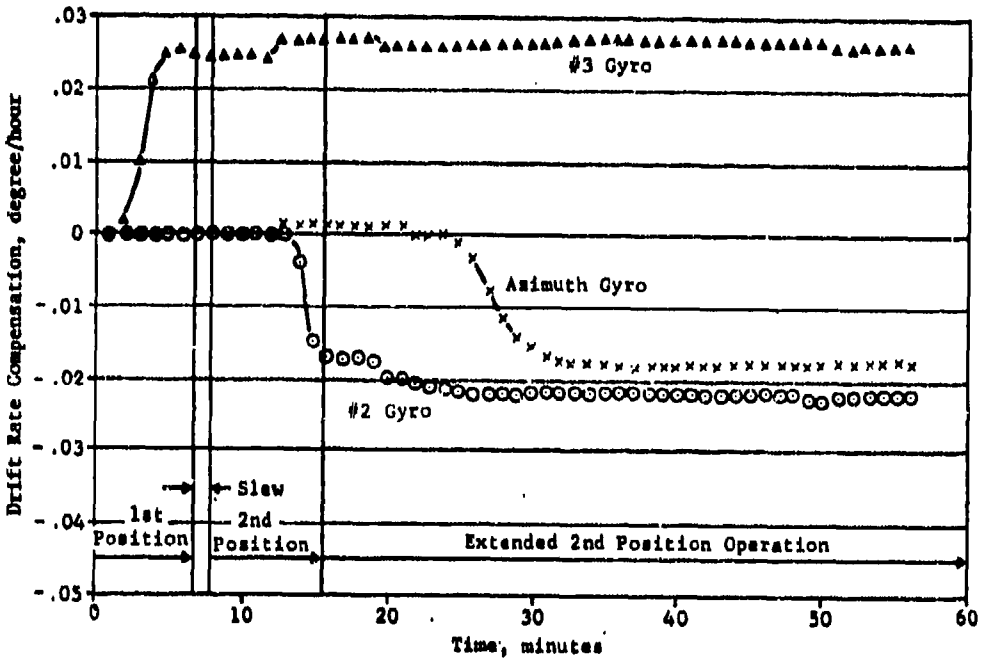


Fig. 27 Laboratory system ground alignment gyro calibration history

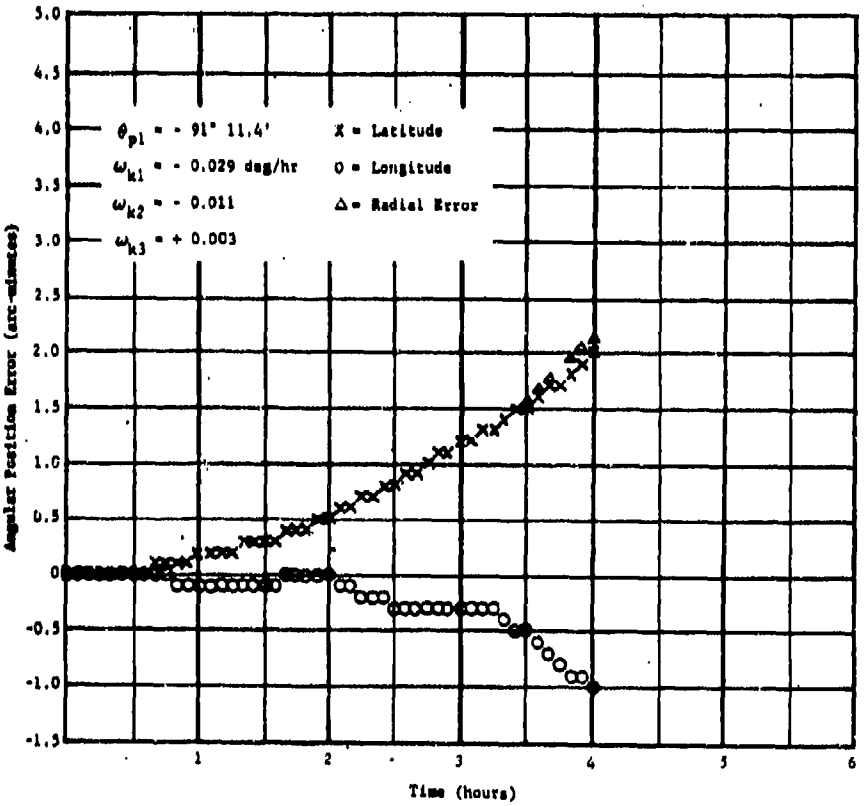


Fig. 28 Laboratory system free-inertial navigation performance following fine ground alignment

CHAPTER 14 - APPLICATION OF KALMAN FILTERING TECHNIQUES TO THE APOLLO PROGRAM

by

Richard H.Battin and Gerald M.Levine

Instrumentation Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, USA

NOTATION

A	measured angle
A_L	angle with respect to vertical in unknown landmark calculation
a	denominator of weighting vector
$\underline{a}_d$	disturbing acceleration
a_H	semi-major axis of horizon ellipse
$\underline{b}$	measurement geometry vector
$\underline{b}_0$, $\underline{b}_1$, $\underline{b}_2$	partitions of geometry vector
b_H	semi-minor axis of horizon ellipse
$C(x)$	special transcendental function
c	speed of light
d	factor in horizon calculation
$\underline{\underline{E}}$	covariance matrix
$\underline{\underline{E}}_{66}$	6×6 partition of covariance matrix
f	general derivative function
$f(q)$	special function in Encke's method
$\underline{\underline{G}}$	gravity gradient matrix
$\underline{g}$	special gravity vector
$\underline{\underline{I}}$	identity matrix
$\underline{i}$	unit vector
$\underline{i}_{AB}$	unit vector B relative to A
J_{kM}, J_{kM}, J_{ksM}	coefficients of gravity potential functions
k_1, k_2, k_3	derivative functions
$\underline{\underline{M}}$	transformation matrix
$\underline{\underline{0}}$	zero matrix
P_k'	derivatives of Legendre polynomials
Q	measured quantity
q	special variable in Encke's method
R	measured range
$\underline{r}_{AB}$	position vector B relative to A
r_E	earth equatorial radius
$\underline{r}_L$	position vector of landmark
r_M	mean lunar radius
r_p	radius of primary body
$\underline{r}_{PV(c)}$	conic position vector
$\underline{r}_V$	position vector of vehicle

r_{xm}	projection of r_{CL}
$S(x)$	special transcendental function
T	transformation matrix
t, t_0	time
$\hat{i}_0, \hat{i}_1$	horizon vectors
U	projection operator
$\text{Unit}(q)$	unit vector in direction of q
v_{AB}	velocity vector B relative to A
$v_{cv(c)}$	conic velocity vector
W	error transition matrix
W_{66}	6×6 partition of W matrix
$W_0, W_1, \ldots, W_8$	3×3 sub-matrices of W matrix
$W_0, W_1, \ldots, W_{24}$	vector partitions of W matrix
x	generalized anomaly in Kepler's equation
x_H	component of position vector in horizon coordinates
x_{MV}	component of position vector in moon coordinates
y_H	component of position vector in horizon coordinates
y_{MV}	component of position vector in moon coordinates
Z	special vector in measurement incorporation
Z_0, Z_1, Z_2	three-dimensional partitions of Z vector
$z_0, z_1, \ldots, z_8$	components of Z vector
α	angular error in landmark tracking
$\overline{\alpha^2}$	mean squared measurement error
α_0	reciprocal of semi-major axis
β	RR shaft angle
γ	factor in measurement incorporation
Δt	integration time step
δ	position deviation vector from conic
δ_q	deviation in quantity q
$\delta\beta, \delta\theta$	RR shaft and trunnion angle biases
ε	position error vector
ζ	parameter error vector
η	velocity error vector
θ	RR trunnion angle
μ_A	gravitational constant of body A
χ	velocity deviation vector from conic

σ_A^2	angle measurement error variance
σ_L^2	landmark error variance
σ_α^2	variance of angular error α
ϕ	co-latitude
Φ	special vector in Nyström's method
χ	polar angle component of landmark tracking error
Ψ	special vector in Nyström's method
ω	weighting vector
$\omega_0, \omega_1, \omega_2$	three-dimensional partitions of ω

Subscripts

C	command module
D	dimension of state vector to be estimated
E	earth
H	horizon
L	lunar module or landmark
M	moon
m	measured
n	stage in integration time step
P	primary body
Q	secondary body
S	sun
s	star
V	vehicle
x, y, z	coordinate axes

Superscripts

T	transpose of vector or matrix
*	altered variable

Abbreviations

CM	command module
CMC	command module computer
DSKY	display and keyboard
IMU	inertial measuring unit
LM	lunar module
LMC	lunar module computer
RR	rendezvous radar
SCT	scanning telescope
SXT	sextant
VHF	very high frequency

CHAPTER 14 - APPLICATION OF KALMAN FILTERING TECHNIQUES TO THE APOLLO PROGRAM

Richard H.Battin and Gerald M.Levine

1. COASTING FLIGHT NAVIGATION

Navigating the Apollo spacecraft in coasting flight involves two processes. First, frequent navigation measurements are made to improve the estimate of the spacecraft's position and velocity. Second, a prediction of the orbit is made periodically so that small corrections to the speed and direction of motion can be applied using the propulsion system if the spacecraft is not following the intended course.

Predicting the course of Apollo during prolonged periods of coasting flight is the same as the astronomer's problem of predicting the positions of the moon and planets. There are several considerations which influence the ability to make long-range predictions. First of all are the mathematical techniques used for solving the equations of motion. Unless rather elaborate computational techniques are employed, numerical errors will propagate and rapidly degrade the solution. Second, the accuracy of predicting position and velocity also is subject to our knowledge of the physical properties of the solar system. Finally, and most important of all, is the accuracy of initial conditions - the values of position and velocity at the time from which the prediction is made.

In order to insure accurate initial conditions, it is necessary periodically to correct the estimate of spacecraft position and velocity using data gathered from optical or radar measurements. Use of a space sextant allows the astronaut, for example, to measure the apparent elevation of a star above the earth's horizon or to measure the angle subtended by the directions to a star and to a landmark on the moon. At the time a measurement is made, the best estimate of the spacecraft's position and velocity is contained in the on-board digital computer. Then, since the directions of the stars and the locations of landmarks are known, it is possible to calculate the expected value of the angle to be measured.

When the expected value of this measurement is compared with the value actually measured, the difference can be used to correct the estimate of the spacecraft's position and velocity. A sequence of such measurements separated in time, together with an accurate mathematical description of the solar system, will eventually produce estimates with sufficient precision to permit corrective maneuvers to be made with confidence.

1.1 Navigation Instruments

To accomplish the lunar landing objectives of the Apollo mission, two vehicles are used - the command module (CM) and the lunar module (LM). Each vehicle is equipped with sensors for acquiring navigation measurement data, together with a digital computer for information processing and orbit prediction (see Reference 1). By appropriate utilization of these instruments, the Apollo astronauts can solve the three major coasting flight navigation problems of the lunar mission: (1) cislunar-midcourse navigation of the CM to and from the moon; (2) navigation of the CM in orbit about the moon; and (3) navigating both the CM and LM during the rendezvous phase in lunar orbit.

Reference orientation is maintained in both vehicles by means of a device called an inertial measuring unit (IMU). This instrument is basically a small platform supported and pivoted so that the spacecraft is free to rotate about it. On this platform are mounted three gyroscopes that sense and prevent any rotation of the platform from occurring.

In the command module a rigid structure mounted to the spacecraft, called the navigation base, provides a common mounting structure for a telescope, a sextant and the base of the IMU gimbal system. Precision angle transducers on each of the axes of the optical instruments and on each of the axes of the IMU gimbals permit the indicated angles to be processed in the command module computer (CMC) to generate the components of the optical targets in inertial system stable-member coordinates.

The sextant (SXT), shown schematically in Figure 1, is a 28-power, narrow-field-of-view instrument having two lines of sight. One of these lines of sight is fixed to the spacecraft and is aimed at a landmark or the horizon by turning the vehicle in space using orientation commands to the attitude control system. The second line of sight can be pointed at a star through the use of a two-axis hand controller. One axis of this motion, the shaft axis, is parallel to the landmark line of sight. The shaft drive changes the plane in which the navigation angle is measured by rotating the head of the instrument as a whole. The trunnion drive sets the navigation angle by tilting the trunnion axis mirror. By superimposing the two images in the field of view and signaling this event to the computer by depressing a mark button, the navigator can measure the angle between the lines of sight to a star and either a landmark or the horizon.

The scanning telescope (SCT), shown in Figure 2, is a unity-power, wide-field-of-view instrument having a single line of sight. Both shaft and trunnion control of this line of sight is possible. The shaft angle is made always to follow the sextant shaft angle, while the trunnion can be selected so that the astronaut may look either along the star or landmark line of sight of the sextant. The wide field of view aids considerably in star and landmark recognition.

A very high frequency (VHF) link between the two vehicles exists which is normally used for inter-vehicle voice communication. However, a signal path through this VHF link from the CM to the LM and back to the CM makes possible a measurement of the range between the vehicles. Data from this automatic VHF range-link is used in the CMC during rendezvous navigation to complement the manually acquired optical measurements.

Optical navigation sightings are not required in the lunar module. However, to aid in a successful rendezvous with the command module, a rendezvous radar (RR), shown in Figure 3, is provided and mounted near the LM inertial measuring unit so that direction data can be related between the two. With this instrument the range and range rate of the CM with respect to the LM as well as the direction, in terms of shaft and trunnion angles of the radar antenna, are made available to the lunar module computer (LMC) for state vector updating.

The command and lunar module computers are designed to handle a relatively large and diverse set of on-board data processing and control functions. Some of the special requirements for this computer include (a) real-time solution of several problems simultaneously on a priority basis, (b) efficient two-way communication with the navigator, (c) capability of ground control radio links, and (d) multiple signal interfaces of both a discrete and continuously variable type. The memory section has a cycle time of 12 μsec and consists of a fixed (read only) portion of 36,864 words, together with an erasable portion of 2048 words. Each word in memory is 16 bits long (15 data bits and an odd parity bit). Data words are stored as signed 14-bit words.

Most of the computer programs relevant to guidance and navigation are written in a pseudocode notation for economy of storage. This notation is encoded and stored as a list of data words. An "interpreter" program translates this list into a sequence of sub-routine linkages. Thus, the small basic instruction set is effectively expanded into a comprehensive mathematical language, which includes matrix and vector operations, using numbers of 28 bits and sign.

The display and keyboard (DSKY), illustrated in Figure 4, serves as the communication medium between the computer and the navigator. The principal part of the display is a set of three registers, each containing five decimal digits, so that a word of 15 bits can be displayed in one register by five octal digits. Three registers are used because of the frequent need to display the three components of a vector. Data are entered in the computer by the astronaut through the keyboard. When the computer requires a response from the astronaut, certain lights are caused to flash on and off in order to attract his attention.

1.2 Navigation Data Processing

The recursive formulation of the optimum linear estimator, as originally devised by R.E.Kalman, is ideally suited to the space navigation problem - especially when a vehicle-borne computer is utilized to process data obtained from on-board instrumentation. With the Kalman estimator, measurement data may be incorporated sequentially, as they are obtained, without recourse to the batch processing techniques required by other methods. Furthermore, within the framework of a single computational algorithm, estimates of quantities in addition to position and velocity, such as radar biases, may be included by the simple expedient of increasing the dimension of the state vector. Finally, matrix inversion, with all of its numerical pitfalls, may also be avoided by regarding all measurement data as single-dimensional or scalar information.

Each computer in the two vehicles maintains an estimate of the position and velocity vectors of both its own and the other spacecraft. These two state vectors are normally six-dimensional but at times are augmented by certain parameters which must also be estimated as part of the navigation process.

In cislunar-midcourse navigation the command module computer only is involved and the CM state vector is of six dimensions. However, when the command module is navigating in lunar orbit, it is necessary also to estimate the position vector of the particular landmark which is being tracked. This is conveniently accomplished by utilizing a nine-dimensional state, the first six elements of which are the components of the CM position and velocity vectors in moon-centered, non-rotating rectangular coordinates, while the last three elements are the components of the lunar landmark position vector.

Two separate rendezvous navigation programs, one in each computer, are used simultaneously during the rendezvous phase of the mission with each solving the navigation problem independently of the other. In the CMC the six-dimensional state vector of either the CM or the LM can be updated from the measurement data obtained with the CM sensors. Normally, it is the LM state vector which is altered, but the mode is at the option of the astronauts. The selection of the updating mode is based primarily on which vehicle's state is more accurately known initially and which vehicle is active in controlling the rendezvous maneuvers.

Since the rendezvous radar of the LM is not structurally mounted with the IMU, significant unknown biases in the knowledge of the direction of the radar antenna are possible. In order to achieve the required accuracy during rendezvous, it is necessary to include the RR angle biases as components of the state vector to be estimated in the LMC. Although the augmented state vector then has eight components, it is treated as nine-dimensional for computational convenience with the ninth element zero.

2. EXTRAPOLATION OF THE STATE VECTOR AND ERROR TRANSITION MATRIX

The estimates of position and velocity are maintained in the spacecraft computers in non-rotating rectangular coordinates and are referenced to either the earth or the moon. An earth-centered equatorial coordinate system is used when the vehicle is outside of the lunar sphere of influence. Inside of this sphere the center of coordinates coincides with the center of the moon. The extrapolation of position and velocity is made by a direct numerical integration of the equations of motion.

The basic equation may be written in vector form as

$$\frac{d^2}{dt^2} \underline{r}_{PV} + \frac{\mu_P}{r_{PV}^3} \underline{r}_{PV} = \underline{a}_d \,, \tag{2.1}$$

where $\underline{r}_{PV}$ is the vector position of the vehicle with respect to the primary body P, which is either the earth or moon, and μ_P is the gravitational constant of P. The vector $\underline{a}_d$ is the vector acceleration which prevents the motion of the vehicle from being precisely a conic with P at the focus.

If $\underline{a}_d$ is small compared with the central force field, direct use of Equation (2.1) is inefficient. As an alternative, the integration may be accomplished by employing the technique of differential accelerations suggested by Encke (see pages 189-190 of Reference 2).

2.1 Encke's Method

At time t_0, the position and velocity vectors $\underline{r}_{PV}(t_0)$ and $\underline{v}_{PV}(t_0)$ define an osculating conic orbit. The vector difference $\underline{\delta}(t)$ between the actual and conic orbits satisfies the following differential equation:

$$\frac{d^2}{dt^2} \underline{\delta} = \frac{\mu_P}{r_{PV(C)}^3} \left[\left(1 - \frac{r_{PV(C)}^3}{r_{PV}^3} \right) \underline{r}_{PV} - \underline{\delta} \right] + \underline{a}_d \,, \tag{2.2}$$

subject to the initial conditions

$$\underline{\delta}(t_0) = \underline{0} \,, \qquad \frac{d}{dt} \underline{\delta}(t_0) = \underline{v}(t_0) = \underline{0} \,,$$

where $\underline{r}_{PV(C)}$ is the osculating conic position vector. The numerical difficulties which would arise from the evaluation of the coefficient of $\underline{r}_{PV}$ in Equation (2.2) may be avoided. Since

$$\underline{r}_{PV}(t) = \underline{r}_{PV(C)}(t) + \underline{\delta}(t) \,, \tag{2.3}$$

it follows that

$$1 - \frac{r_{PV(C)}^3}{r_{PV}^3} = -f(q_0) = 1 - (1+q_0)^{3/2} \,,$$

where

$$q_0 = \frac{(\underline{\delta} - 2\underline{r}_{PV}) \cdot \underline{\delta}}{r_{PV}^2} \,. \tag{2.4}$$

The function $f(q)$ may be conveniently evaluated from

$$f(q) = q \, \frac{3 + 3q + q^2}{1 + (1+q)^{3/2}} \,. \tag{2.5}$$

Encke's method may now be summarized as follows:

(i) Position in the osculating orbit is calculated from

$$\underline{r}_{PV(C)}(t) = \left[1 - \frac{x^2}{r_{PV}(t_0)} C(\alpha_0 x^2) \right] \underline{r}_{PV}(t_0) + \left[(t - t_0) - \frac{x^3}{\sqrt{\mu_P}} S(\alpha_0 x^2) \right] \underline{v}_{PV}(t_0) \,, \tag{2.6}$$

where

$$\alpha_0 = \frac{2}{r_{PV}(t_0)} - \frac{v_{PV}(t_0)^2}{\mu_P} \tag{2.7}$$

and x is determined as the root of Kepler's equation in the form

$$\sqrt{(\mu_P)}(t - t_0) = \frac{\underline{r}_{PV}(t_0) \cdot \underline{v}_{PV}(t_0)}{\sqrt{\mu_P}} x^2 C(\alpha_0 x^2) + [1 - r_{PV}(t_0)\alpha_0] x^3 S(\alpha_0 x^2) + r_{PV}(t_0) x \,. \tag{2.8}$$

The special transcendental functions S and C are defined by

$$\left.\begin{aligned} S(x) &= \frac{1}{3!} - \frac{x}{5!} + \frac{x^2}{7!} - \cdots \\ C(x) &= \frac{1}{2!} - \frac{x}{4!} + \frac{x^2}{6!} - \cdots \end{aligned}\right\} \tag{2.9}$$

(ii) Deviations from the osculating orbit are obtained by a numerical integration of

$$\frac{d^2}{dt^2}\,\underline{\delta}(t) = -\frac{\mu_P}{r^3_{PV(c)}(t)}\left[f(q_c)\underline{r}_{PV}(t) + \underline{\delta}(t)\right] + \underline{a}_d(t) \,. \tag{2.10}$$

The first term on the right-hand side of the last equation must remain small, i.e. of the same order as $\underline{a}_d(t)$, if the method is to be efficient. As the deviation vector $\underline{\delta}$ grows in magnitude, this term will eventually increase in size. Therefore, in order to maintain the efficiency, a new osculating orbit should be defined by the true position and velocity. The process of selecting a new conic orbit from which to calculate deviations is called rectification. When rectification occurs, the initial conditions of the differential equation for $\underline{\delta}$ are again zero and the right-hand side is simply the perturbation acceleration $\underline{a}_d$ at the time of rectification.

(iii) The position vector $\underline{r}_{PV}(t)$ is computed from Equation (2.3) using Equation (2.6). The velocity vector $\underline{v}_{PV}(t)$ is then computed as

$$\underline{v}_{PV}(t) = \underline{v}_{PV(c)}(t) + \underline{v}(t) \,, \tag{2.11}$$

where

$$\underline{v}_{PV(c)}(t) = \frac{\sqrt{\mu_P}}{r_{PV}(t_0)r_{PV(c)}(t)}\left[\alpha_0 x^3 S(\alpha_0 x^2) - x\right]\underline{r}_{PV}(t_0) + \left[1 - \frac{x^2}{r_{PV(c)}(t)}\,C(\alpha_0 x^2)\right]\underline{v}_{PV}(t_0) \,. \tag{2.12}$$

2.2 Disturbing Acceleration

The form of the disturbing acceleration $\underline{a}_d$ to be used depends on the phase of the mission. In earth or lunar orbit only the gravitational perturbations arising from the non-spherical shape of the primary body P need be considered. During translunar and transearth flight, the gravitational attraction of the sun and the secondary body Q (either earth or moon) are relevant forces. A summary of the various cases appears below.

(i) Earth Orbit*

$$\underline{a}_d = \frac{\mu_E}{r^2_{EV}}\sum_{k=2}^{4} J_{kE}\left(\frac{r_E}{r_{EV}}\right)^k \left[P'_{k+1}(\cos\phi)\underline{i}_{EV} - P'_k(\cos\phi)\,\underline{i}_E\right] \,, \tag{2.13}$$

where

$$\begin{aligned} P'_2(\cos\phi) &= 3\cos\phi \\ P'_3(\cos\phi) &= \tfrac{1}{2}(15\cos^2\phi - 3) \\ P'_4(\cos\phi) &= \tfrac{1}{3}(7\cos\phi P'_3 - 4P'_2) \\ P'_5(\cos\phi) &= \tfrac{1}{4}(9\cos\phi P'_4 - 5P'_3) \end{aligned}$$

are the derivatives of the Legendre polynomials;

$$\cos\phi = \underline{i}_{EV}\cdot\underline{i}_E$$

is the cosine of the angle ϕ between the unit vector $\underline{i}_{EV}$ in the direction of $\underline{r}_{EV}$ and the unit vector $\underline{i}_E$ in the direction of the north pole; r_E is the equatorial radius of the earth; and J_{2E}, J_{3E}, J_{4E} are the coefficients of the second, third and fourth harmonics of the earth's potential function. The subscript E denotes the center of the earth as the origin of coordinates.

(ii) Translunar and Transearth Flight**

$$\underline{a}_d = -\frac{\mu_Q}{r^3_{QV}}\left[f(q_Q)\underline{r}_{PQ} + \underline{r}_{PV}\right] - \frac{\mu_S}{r^3_{SV}}\left[f(q_S)\underline{r}_{PS} + \underline{r}_{PV}\right] \,, \tag{2.14}\dagger$$

where the subscripts Q and S denote the secondary body and the sun, respectively. Thus, for example, $\underline{r}_{PS}$ is the position vector of the sun with respect to the primary body. The arguments $q_{(\,)}$ are calculated from

---

* See page 198 of Reference 2.

** See page 12 of Reference 2.

† During the periods of transearth and translunar flight when the vehicle is near the primary body, the disturbing acceleration which is used is the sum of Equation (2.14) and either (2.13) or (2.17), whichever is appropriate.

$$q_{()} = \frac{(\underline{r}_{PV} - 2\underline{r}_{P()}) \cdot \underline{r}_{PV}}{r^2_{P()}} \tag{2.15}$$

and the function f from Equation (2.5).

In the vicinity of the lunar sphere of influence a change in origin of coordinates is made. Thus

$$\left.\begin{aligned} \underline{r}_{PV}(t) - \underline{r}_{PQ}(t) &= \underline{r}_{QV}(t) - \underline{r}_{PV}(t) \\ \underline{v}_{PV}(t) - \underline{v}_{PQ}(t) &= \underline{v}_{QV}(t) - \underline{v}_{PV}(t) \end{aligned}\right\} . \tag{2.16}$$

(iii) Lunar Orbit*

$$\underline{a}_d = \frac{\mu_M}{r^2_{MV}}\left\{\sum_{k=2}^{N} J_{kM}\left(\frac{r_M}{r_{MV}}\right)^k \left[P'_{k+1}(\cos\phi)\underline{i}_{MV} - P'_k(\cos\phi)\underline{i}_Z\right] + 3J_{22M}\left(\frac{r_M}{r_{MV}}\right)^2\left[4\,\frac{x_{MV}y_{MV}}{x^2_{MV}+y^2_{MV}}\,(\underline{i}_{MV}\times\underline{i}_Z) + \right.\right.$$

$$\left.\left. + \frac{x^2_{MV}-y^2_{MV}}{x^2_{MV}+y^2_{MV}}\,((5\cos^2\phi - 3)\underline{i}_{MV} - 2\cos\phi\,\underline{i}_Z)\right]\right\} , \tag{2.17}$$

where r_M is the mean lunar radius; x_{MV} and y_{MV} are the lunar equatorial plane components of $\underline{r}_{MV}$; and J_{22M} is that coefficient of the moon's potential function associated with the asymmetry of the moon about its polar axis.

2.3 Error Transition Matrix

The position and velocity vectors, as maintained in the computer, are only estimates of the true values. As part of the Kalman estimation technique, it is also necessary to record statistical data for the processing of navigation measurements.

If $\underline{\xi}(t)$ and $\underline{\eta}(t)$ are the errors in the estimates of the position and velocity vectors, respectively, then the six-dimensional covariance matrix $\underline{E}_{66}(t)$ is defined by**

$$\underline{E}_{66}(t) = \begin{pmatrix} \overline{\underline{\xi}(t)\underline{\xi}(t)^T} & \overline{\underline{\xi}(t)\underline{\eta}(t)^T} \\ \overline{\underline{\eta}(t)\underline{\xi}(t)^T} & \overline{\underline{\eta}(t)\underline{\eta}(t)^T} \end{pmatrix} . \tag{2.18}$$

As noted in Section 1, for certain applications it is necessary to expand the state vector and the covariance matrix to more than six dimensions, in order to include the estimation of landmark locations in the CMC during orbital navigation and the rendezvous radar tracking biases in the LMC for rendezvous navigation. For this purpose a nine-dimensional covariance matrix is defined as follows:

$$\underline{E}(t) = \begin{pmatrix} & \underline{E}_{66}(t) & & \overline{\underline{\xi}(t)\underline{\zeta}^T} \\ & & & \overline{\underline{\eta}(t)\underline{\zeta}^T} \\ \overline{\underline{\zeta}\underline{\xi}(t)^T} & & \overline{\underline{\zeta}\underline{\eta}(t)^T} & \overline{\underline{\zeta}\underline{\zeta}^T} \end{pmatrix} , \tag{2.19}$$

where the components of the three-dimensional vector $\underline{\zeta}$ are the estimation errors of the three additional variables to be estimated.

To take full advantage of the operations provided by the interpreter in the spacecraft computers, the covariance matrix is restricted to either six or nine dimensions. In the LMC rendezvous navigation procedure, only two additional quantities are estimated. However, a dummy variable is added to the state vector to make it nine-dimensional.

The appropriate relation to be used in extrapolating the $\underline{E}_{66}(t)$ matrix is obtained as follows. The basic equation of motion of the vehicle with respect to the primary body may be written in the form:

$$\frac{d^2}{dt^2}\underline{r}_{PV} + \frac{\mu_P}{r^3_{PV}}\underline{r}_{PV} + \frac{\mu_Q}{r^3_{QV}}\underline{r}_{QV} + \frac{\mu_Q}{r^3_{PQ}}\underline{r}_{PQ} = \underline{g} .$$

where the vector $\underline{g}$ encompasses the gravitational acceleration of the sun as well as the other disturbances arising from the asymmetrical shapes of the earth and moon. This equation can be linearized about the best estimate of the vehicle's state and, to a first order of approximation, the estimate error vectors will satisfy the resulting linear differential equation. We have

---

* See page 3 of Reference 2.

** The transpose of a vector or a matrix is denoted by a superscript T.

$$\frac{d}{dt}\begin{pmatrix}\xi(t)\\ \eta(t)\end{pmatrix} = \begin{pmatrix}\underline{0} & \underline{I}\\ \underline{G}(t) & \underline{0}\end{pmatrix}\begin{pmatrix}\xi(t)\\ \eta(t)\end{pmatrix},$$

where $\underline{I}$ and $\underline{0}$ are the three-dimensional identity and zero matrices, respectively.

The matrix $\underline{G}(t)$ is the three-dimensional gradient of the gravitational field with respect to the components of the position vector $\underline{r}_{PV}$. If we neglect the gradient of the vector $\underline{g}$, then it is easy to show that*

$$\underline{G}(t) = \frac{\mu_P}{r_{PV}^5(t)}\left[3\underline{r}_{PV}(t)\underline{r}_{PV}(t)^T - r_{PV}^2(t)\underline{I}\right] + \frac{\mu_Q}{r_{QV}^5(t)}\left[3\underline{r}_{QV}(t)\underline{r}_{QV}(t)^T - r_{QV}^2(t)\underline{I}\right], \tag{2.20}$$

which is a sufficiently good approximation for translunar and transearth flight. For orbital navigation about the primary body only the first term expression for $\underline{G}(t)$ need be included.

Since

$$\underline{E}_{66}(t) = \overline{\begin{pmatrix}\xi(t)\\ \eta(t)\end{pmatrix}\quad\begin{pmatrix}\xi(t)^T & \eta(t)^T\end{pmatrix}},$$

it is easy to show that the six-dimensional covariance matrix satisfies the following matrix differential equation:

$$\frac{d}{dt}\underline{E}_{66} = \begin{pmatrix}\underline{0} & \underline{I}\\ \underline{G} & \underline{0}\end{pmatrix}\underline{E}_{66} + \underline{E}_{66}\begin{pmatrix}\underline{0} & \underline{G}\\ \underline{I} & \underline{0}\end{pmatrix}.$$

Because of accumulated numerical inaccuracies, it is possible that the covariance matrix may fail to remain positive definite after a large number of computations, as it theoretically must. An innovation to avoid this problem, which has also the advantage of significantly reducing certain computational requirements, is to replace the covariance matrix by a matrix $\underline{W}(t)$, called the error transition matrix. The $\underline{W}(t)$ matrix has the property:

$$\underline{E}(t) = \underline{W}(t)\,\underline{W}(t)^T \tag{2.21}$$

and thus, in a sense, is the square root of the covariance matrix. If needed, the covariance matrix may be determined as the product of the matrix $\underline{W}(t)$ and its transpose, thereby guaranteeing it to be at least positive semi-definite.

The chief computational advantage of the $\underline{W}$ matrix lies in the simplicity of the differential equation which it satisfies. From the differential equation for $\underline{E}_{66}$, the fact that the components of the vector ζ do not change with time, and Equation (2.21), it is obvious that

$$\frac{d}{dt}\underline{W} = \begin{pmatrix}\underline{0} & \underline{I} & \underline{0}\\ \underline{G}(t) & \underline{0} & \underline{0}\\ \underline{0} & \underline{0} & \underline{0}\end{pmatrix}\underline{W}. \tag{2.22}$$

Now let the nine-dimensional matrix $\underline{W}(t)$ be partitioned as

$$\underline{W} = \begin{pmatrix}\underline{W}_0 & \underline{W}_1 & \cdots & \underline{W}_8\\ \underline{W}_9 & \underline{W}_{10} & \cdots & \underline{W}_{17}\\ \underline{W}_{18} & \underline{W}_{19} & \cdots & \underline{W}_{26}\end{pmatrix}. \tag{2.23}$$

Then, we have

$$\left.\begin{aligned}\frac{d}{dt}\underline{W}_i(t) &= \underline{W}_{i+9}(t)\\ \frac{d}{dt}\underline{W}_{i+9}(t) &= \underline{G}(t)\underline{W}_i(t)\\ \frac{d}{dt}\underline{W}_{i+18}(t) &= \underline{0}\end{aligned}\right\} \quad i = 0, 1, \ldots, 8. \tag{2.24}$$

Thus, the extrapolation of the $\underline{W}$ matrix may be accomplished by successively integrating the vector differential equations

$$\frac{d^2}{dt^2}\underline{W}_i(t) = \underline{G}(t)\underline{W}_i(t) \qquad i = 0, 1, \ldots, 8. \tag{2.25}$$

Finally, then, if D is the dimension of the matrix $\underline{W}(t)$, the differential equations for the $\underline{W}_i(t)$ vectors are simply

$$\frac{d^2}{dt^2}\underline{W}_i = \frac{\mu_P}{r_{PV}^3(t)}\left\{3[\underline{i}_{PV}(t)\cdot\underline{W}_i(t)]\,\underline{i}_{PV}(t) - \underline{W}_i(t)\right\} + \frac{\mu_Q}{r_{QV}^3(t)}\left\{3[\underline{i}_{QV}(t)\cdot\underline{W}_i(t)]\,\underline{i}_{QV}(t) - \underline{W}_i(t)\right\} \tag{2.26}$$

$$i = 0, 1, \ldots, D-1$$

with the second term omitted for orbital navigation about the primary body P.

---

* See page 207 of Reference 7.

2.4 Numerical Integration

The extrapolation of the state vector and the error transition matrix requires the solution of $D + 1$ second-order vector differential equations, specifically Equations (2.10) and (2.26). These are all special cases of the form

$$\frac{d^2}{dt^2}\underline{y} = \underline{f}(\underline{y},t) , \tag{2.27}$$

in which the right-hand side is a function of the independent variable and time only. Nyström's method, described in Reference 3, is particularly well-suited to this form and gives an integration method of fourth-order accuracy, while requiring only three computations of the derivatives per time step. (The usual fourth-order Runge-Kutta integration methods require four derivative computations per time step.) The second-order system is written as

$$\left.\begin{aligned} \frac{d}{dt}\underline{y} &= \underline{z} \\ \frac{d}{dt}\underline{z} &= \underline{f}(\underline{y},t) \end{aligned}\right\} \tag{2.28}$$

and the formulae are summarized as follows:

$$\left.\begin{aligned} \underline{y}_{n+1} &= \underline{y}_n + \underline{\phi}(\underline{y}_n)\Delta t \\ \underline{z}_{n+1} &= \underline{z}_n + \underline{\psi}(\underline{y}_n)\Delta t \\ \underline{\phi}(\underline{y}_n) &= \underline{z}_n + \tfrac{1}{6}(\underline{k}_1 + 2\underline{k}_2)\Delta t \\ \underline{\psi}(\underline{y}_n) &= \tfrac{1}{6}(\underline{k}_1 + 4\underline{k}_2 + \underline{k}_3) \\ \underline{k}_1 &= \underline{f}(\underline{y}_n, t_n) \\ \underline{k}_2 &= \underline{f}(\underline{y}_n + \tfrac{1}{2}\underline{z}_n\Delta t + \tfrac{1}{8}\underline{k}_1(\Delta t)^2 , \; t_n + \tfrac{1}{2}\Delta t) \\ \underline{k}_3 &= \underline{f}(\underline{y}_n + \underline{z}_n\Delta t + \tfrac{1}{2}\underline{k}_2(\Delta t)^2 , \; t_n + \Delta t) \end{aligned}\right\} . \tag{2.29}$$

For efficient use of computer storage as well as computing time, the computations should be performed in the following order:

(i) Equation (2.10) is solved using the Nyström formulae (2.29). It is necessary to preserve the values of the vectors $\underline{r}_{PV}$ at times t_n and $t_n + \Delta t/2$ for use in the solution of Equations (2.26).

(ii) Equations (2.26) are solved one at a time using formulae (2.29), together with the values of $\underline{r}_{PV}$ which resulted from the first step.

It has been found experimentally that the maximum value that the integration time step Δt can have is either $0.3(r_{PV}^3/\mu_\oplus)^{3/2}$ or 4000 seconds, whichever is the smaller.

3. INCORPORATION OF MEASUREMENT DATA

An important feature of the Apollo navigation method is that measurement data from a wide variety of sources may be incorporated within the same framework of computation. Associated with each measurement is a D-dimensional vector $\underline{b}$ representing, to a first order of approximation, the variation in the measured quantity Q which would result from variations in the components of the state vector. Thus, each measurement establishes a component of the spacecraft state vector along the direction of the $\underline{b}$ vector in state space.

By algebraically combining the $\underline{\underline{M}}$ matrix, the $\underline{b}$ vector and a mean-squared *a priori* estimation error α^2 in the measurement, there are produced a weighting vector $\underline{\omega}$ and the step change to be made in the error transition matrix to reflect the changes in the uncertainties in the estimated quantities as a result of the measurement. The weighting vector $\underline{\omega}$ has D components and is determined so that the observation data is utilized in a statistically optimum manner.

According to the Kalman estimation theory for a one-dimensional scalar measurement, the weighting vector is determined from

$$\underline{\omega} = \frac{1}{a}\underline{\underline{M}}\underline{b} , \tag{3.1}$$

where $\underline{\underline{M}}$ is the value of the covariance matrix extrapolated to the time of the measurement and

$$a = \underline{b}^T\underline{\underline{M}}\underline{b} + \alpha^2 . \tag{3.2}$$

For convenience, define the vector $\underline{g}$ as

$$\underline{g} = \underline{W}^T\underline{h} , \tag{3.3}$$

so that Equations (3.1) and (3.2) may be written

$$\underline{\omega} = \frac{1}{a}\underline{W}\underline{g} \tag{3.4}$$

$$a = \underline{g}^T\underline{g} + \overline{\alpha^2} . \tag{3.5}$$

Then, if δQ represents the difference between the quantity actually measured and its expected value based on the extrapolated value of the state vector, the change in the state vector is simply $\underline{\omega}\delta Q$.

As a result of the measurement, the statistics embodied in the covariance matrix $\underline{E}$ or its square root $\underline{W}$ must be altered. Again, the Kalman theory dictates that the new value for $\underline{E}$, denoted by $\underline{E}^*$, is obtained from

$$\underline{E}^* = \underline{E}\left(\underline{I} - \frac{1}{a}\underline{h}\underline{h}^T\underline{E}\right)$$

or, in terms of the $\underline{W}$ matrix and the $\underline{g}$ vector,

$$\underline{W}^* = \underline{W}\left(\underline{I} - \frac{1}{a}\underline{g}\underline{g}^T\right)\underline{W}^T , \tag{3.6}$$

where $\underline{I}$ is the D-dimensional identity matrix.

What is desired, of course, is a formula for updating the $\underline{W}$ matrix rather than the $\underline{E}$ matrix. The objective will be achieved if a square root can be found for the parenthesized factor in Equation (3.6). Indeed, the desired result is obtained by determining the value of the parameter γ such that

$$\left(\underline{I} - \frac{1}{a}\underline{g}\underline{g}^T\right) = \left(\underline{I} - \frac{\gamma}{a}\underline{g}\underline{g}^T\right)\left(\underline{I} - \frac{\gamma}{a}\underline{g}\underline{g}^T\right)^T .$$

By straightforward computation it is seen that γ must be

$$\gamma = \frac{1}{1 + \sqrt{(\overline{\alpha^2}/a)}} , \tag{3.7}$$

so that the new value $\underline{W}^*$ is computed as

$$\underline{W}^* = \underline{W} - \gamma\underline{\omega}\underline{g}^T . \tag{3.8}$$

In order to take full advantage of the three-dimensional vector and matrix operations provided by the interpreter in the computer, the nine-dimensional $\underline{W}$ matrix is stored sequentially as follows:

$$\underline{w}_0^T , \underline{w}_1^T , \ldots , \underline{w}_{26}^T .$$

Then, by defining the three-dimensional sub-matrices

$$\underline{W}_0 = \begin{bmatrix} \underline{w}_0^T \\ \underline{w}_1^T \\ \underline{w}_2^T \end{bmatrix} , \quad \underline{W}_1 = \begin{bmatrix} \underline{w}_3^T \\ \underline{w}_4^T \\ \underline{w}_5^T \end{bmatrix} , \quad \ldots . , \quad \underline{W}_8 = \begin{bmatrix} \underline{w}_{24}^T \\ \underline{w}_{25}^T \\ \underline{w}_{26}^T \end{bmatrix} ,$$

so that

$$\underline{W} = \begin{pmatrix} \underline{W}_0^T & \underline{W}_1^T & \underline{W}_2^T \\ \underline{W}_3^T & \underline{W}_4^T & \underline{W}_5^T \\ \underline{W}_6^T & \underline{W}_7^T & \underline{W}_8^T \end{pmatrix} , \tag{3.9}$$

and partitioning the possibly nine-dimensional vectors $\underline{h}$, $\underline{w}$, and $\underline{g}$ as

$$\underline{h} = \begin{bmatrix} \underline{h}_0 \\ \underline{h}_1 \\ \underline{h}_2 \end{bmatrix} \quad \underline{\omega} = \begin{bmatrix} \underline{\omega}_0 \\ \underline{\omega}_1 \\ \underline{\omega}_2 \end{bmatrix} \quad \underline{g} = \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_8 \end{bmatrix} = \begin{bmatrix} \underline{g}_0 \\ \underline{g}_1 \\ \underline{g}_2 \end{bmatrix} , \tag{3.10}$$

the computations developed in this section are conveniently performed as

$$\left.\begin{aligned} \underline{z}_j &= \sum_{k=0}^{D-1} \underline{W}_{j+3k} b_k \,, & a &= \sum_{k=0}^{D-1} \underline{z}_k \cdot \underline{z}_k + \overline{\alpha^2} \\ \underline{\omega}_j^T &= \frac{1}{a} \sum_{k=0}^{D-1} \underline{z}_k^T \underline{W}_{3j+k} & \underline{W}^*_{i+3j} &= \underline{W}_{i+3j} - \gamma \underline{z}_i \underline{\omega}_j \,. \end{aligned}\right\} \tag{3.11}$$

where the subscripts i and j range, respectively, from 0 to $D-1$ and from 0 to $(D/3)-1$.

It is worth emphasizing that the particular formulation of the Kalman estimator, developed in this and the previous section, has achieved what is felt to be a minimum in practical computational efficiency as well as a most sparing use of erasable memory locations in the small vehicle-borne Apollo computers. The latter is possible primarily because of the introduction of the error transition matrix instead of the more conventional covariance matrix.

It should be remarked, however, that straightforward application of the $\underline{W}$ matrix techniques would not have been possible had the dynamics of the state vector model included what is commonly called "process noise". Since we have elected to neglect process noise in favor of computational compactness, the gradual and inevitable decay of the elements of the $\underline{W}$ matrix must be countered in practice by a periodic re-initialization of those elements. It should be clear that, without either including process noise or re-initializing the $\underline{W}$ matrix, eventually all measurement data would be ignored simply because they would be given a zero weighting factor.

4. COMMAND MODULE CISLUNAR - MIDCOURSE NAVIGATION

During the translunar and transearth phases of the Apollo mission, navigation data can be obtained with the sextant by measuring the angle between the lines of sight to a star and an earth or moon horizon or landmark. When the navigator depresses the mark button, indicating to the computer that the two target optical images are properly superimposed in the SXT field of view, the time of the measurement and the measured angle, i.e. the SXT trunnion angle A, are automatically recorded in the CMC. The navigator must then inform the computer, through the DSKY, of the identity of the star and the particular feature of the earth or moon involved in the sighting. These data are used by the computer to determine: (i) the CM state vector estimate and its associated error transition matrix extrapolated to the measurement time, as described in Section 2; (ii) the estimated CM position vector $\underline{r}_V$ relative to that body used in the measurement (since this may be either the primary body P or the secondary body Q, the single subscript will serve without ambiguity); (iii) the unit vector $\underline{i}_s$ in the direction of the particular star used (there are data for 37 stars in the CMC fixed memory which the navigator can identify by code number); and (iv) the position vector $\underline{r}_L$ of the landmark, assuming a star-landmark measurement.

4.1 Star-Landmark Measurements

Because of the extremely high accuracy required for midcourse navigation measurements, observation data must be corrected for aberration prior to processing. Aberration is the term used to describe the change in the apparent direction of an object due to the velocity of the observer normal to the line of sight to the object. It is only when this perpendicular velocity has a magnitude of tens of thousands of feet per second that the aberration correction is necessary. It should be remarked that the correction is not required for rendezvous or orbit navigation.

The apparent direction of the star is computed from

$$\underline{i}_s^* = \text{Unit}\left(\underline{i}_s + \frac{\underline{v}_{PV} + \underline{v}_{SE}}{c}\right), \tag{4.1}$$

where the notation Unit $(\underline{q})$ is understood to mean a unit vector in the direction of the vector $\underline{q}$. The vector $\underline{v}_{SE}$ is the velocity of the earth relative to the sun, and c is the speed of light. The velocity of the sun relative to the star is already taken into account for the basic star directional data $\underline{i}_s$, and the relative velocity between the earth and moon is negligible for the case in which the moon is the primary body. The expected direction of the landmark is obtained from

$$\underline{i}^*_{VL} = \text{Unit}\left(\underline{i}_{VL} + \frac{\underline{v}_{PV}}{c}\right), \tag{4.2}$$

where

$$\underline{i}_{VL} = \text{Unit}(\underline{r}_L - \underline{r}_V)$$

is a unit vector defining the estimated direction of the landmark relative to the vehicle.

Using the corrected unit vectors corresponding to the directions of the two SXT lines of sight, the CMC computes the measured deviation δQ as

$$\delta Q = A - \cos^{-1}(\underline{i}_s^* \cdot \underline{i}^*_{VL}) \,. \tag{4.3}$$

Although for cislunar-midcourse navigation the CM state vector is six-dimensional, only the first three components of the measurement geometry vector $\underline{h}$, defined in Section 3, will be non-zero, since the measured quantity is independent of velocity except through the small aberration correction. To determine this non-zero partition $\underline{h}_0$ of the $\underline{h}$ vector, we compute the first-order differential of the expression

$$r^*_{VL} \cos A = \underline{i}^*_m \cdot \underline{r}^*_{VL} ,$$

with the result that

$$\cos A\, \delta r^*_{VL} - r^*_{VL} \sin A\, \delta A = \underline{i}^*_m \cdot \delta \underline{r}^*_{VL} .$$

Then, since

$$\delta r^*_{VL} = \frac{\underline{r}^*_{VL} \cdot \delta \underline{r}^*_{VL}}{r^*_{VL}} \quad \text{and} \quad \delta \underline{r}^*_{VL} = -\delta \underline{r}_{PV} ,$$

we have

$$r^*_{VL} \sin A\, \delta A = (\underline{i}^*_m - \cos A\, \underline{i}^*_{VL}) \cdot \delta \underline{r}_{PV} .$$

Hence

$$\underline{h}_0 = \frac{1}{r^*_{VL} \sin A} (\underline{i}^*_m - \cos A\, \underline{i}^*_{VL})$$

or

$$\underline{h}_0 = \frac{1}{r^*_{VL}} \text{Unit} [\underline{i}^*_m - (\underline{i}^*_m \cdot \underline{i}^*_{VL})\, \underline{i}^*_{VL}] . \tag{4.4}$$

Finally, the measurement error variance is computed from

$$\overline{\alpha^2} = \sigma^2_A + \frac{\sigma^2_L}{r^{*2}_{VL}} , \tag{4.5}$$

where σ^2_A and σ^2_L are the assumed error variances in the SXT trunnion angle and the landmark, respectively.

4.2 Star-Horizon Measurements

The processing of star-horizon measurement data is the same as that required for star-landmark data, provided the landmark location vector $\underline{r}_L$ is replaced by a vector from the spacecraft to the horizon. The determination of this horizon vector is made according to the following geometrical arguments.

Consider first an earth-horizon measurement. The star direction $\underline{i}_m$ and the estimated CM position vector $\underline{r}_V$ determine a plane. Assuming that the horizon is at a constant altitude above the earth's surface, the intersection of this measurement plane with the horizon is approximately an ellipse. The orientation of this horizon ellipse is defined in terms of three mutually orthogonal unit vectors:

$$\left.\begin{aligned} \underline{i}_2 &= \text{Unit}(\underline{i}_m \times \underline{r}_V) \\ \underline{i}_0 &= \text{Unit}(\underline{i}_N \times \underline{i}_M) \\ \underline{i}_1 &= \underline{i}_2 \times \underline{i}_0 . \end{aligned}\right\} \tag{4.6}$$

where $\underline{i}_N$ is a unit vector in the direction of the earth's polar axis. Referring to Figure 7, it is clear that $\underline{i}_0$ and $\underline{i}_1$ coincide with the semi-major and semi-minor axes of the horizon ellipse, respectively. The plane containing the horizon ellipse is inclined with respect to the earth's equatorial plane by an angle I, where

$$\sin I = \underline{i}_1 \cdot \underline{i}_N . \tag{4.7}$$

The shape of the horizon ellipse is determined by the lengths of its major and minor axes. Assuming the contour of the earth to be well-approximated by the so-called Fischer ellipsoid, the semi-major axis of the horizon ellipse a_H is simply the sum of the semi-major axis of this ellipsoid and the constant horizon altitude. Likewise, the semi-minor axis b_H is found by adding the horizon altitude to that value of the radius of the Fischer ellipsoid which corresponds to a latitude equal to the inclination angle I.

The problem of determining the vector $\underline{r}_L$ is readily solved in the horizon coordinate system for which the x and y axes coincide with the directions $\underline{i}_0$ and $\underline{i}_1$, respectively, as illustrated in Figure 8. The matrix

$$M = \begin{pmatrix} \underline{i}_0^T \\ \underline{i}_1^T \\ \underline{i}_2^T \end{pmatrix} \tag{4.8}$$

will serve to transform vectors from the original coordinate system to the horizon system. Let $\underline{r}_H$ and $\underline{i}_{mH}$ represent the components of $\underline{r}_V$ and $\underline{i}_m$ in horizon coordinates, so that

$$\left.\begin{aligned} \underline{r}_H &= \begin{pmatrix} x_H \\ y_H \\ 0 \end{pmatrix} = M\underline{r}_V \\ \underline{i}_{sH} &= M\underline{i}_s \end{aligned}\right\} \tag{4.9}$$

Further, define $\underline{t}_0$ and $\underline{t}_1$ as vectors from the point (x_H, y_H) to the two points of tangency with the horizon ellipse.

The vectors $\underline{t}_0$ and $\underline{t}_1$ are obtained by solving simultaneously the equation of the horizon ellipse:

$$\frac{x^2}{a_H^2} + \frac{y^2}{b_H^2} = 1$$

and the equation of the line tangent to the ellipse and passing through the point (x_H, y_H):

$$\frac{xx_H}{a_H^2} + \frac{yy_H}{b_H^2} = 1 .$$

We have

$$\underline{t}_i = \frac{1}{d}\begin{pmatrix} x_H \pm \frac{a_H}{b_H} y_H \sqrt{(d-1)} \\ y_H \mp \frac{b_H}{a_H} x_H \sqrt{(d-1)} \\ 0 \end{pmatrix}, \qquad i = 0, 1 , \tag{4.10}$$

where

$$d = \frac{x_H^2}{a_H^2} + \frac{y_H^2}{b_H^2} .$$

The upper and lower signs in Equation (4.10) correspond, respectively, with $i = 0$ and $i = 1$.

The two points of tangency, $\underline{t}_0$ and $\underline{t}_1$, correspond to the two horizon points, $\underline{r}_{LH}^+$ (near horizon) and $\underline{r}_{LH}^-$ (far horizon), but not necessarily respectively. By definition, the near horizon is associated with that vector $\underline{t}_i$ which makes the smaller angle with the star vector. The horizon vector is then

$$\underline{r}_L = M^T\underline{t} , \tag{4.11}$$

using for $\underline{t}$ either $\underline{t}_0$ or $\underline{t}_1$, whichever is appropriate.

In the CMC, moon-horizon measurements are processed under the assumption that the moon has a spherical shape. The determination of the horizon vector $\underline{r}_L$ is made as for the earth-horizon case but with the mean radius of the moon used for both a_H and b_H.

5. COMMAND MODULE NAVIGATION DURING RENDEZVOUS

To accomplish rendezvous with the CM and LM in the Apollo mission, two types of navigation data are obtained with the CM-based sensors. As described in Section 1, these data are acquired both manually with the optics and automatically through the VHF range-link. They serve in complementary roles for rendezvous navigation. The optical data yields information in the directions normal to the CM-to-LM line of sight, whereas the VHF range-link gives information along the line of sight. Taken together, they provide an excellent rendezvous navigation capability.

It is recognized that the optimum method of applying the Kalman estimation technique to rendezvous navigation would be to solve the twelve-dimensional problem of simultaneous estimation of both the CM and LM states. However, the limitations of a relatively small erasable memory preclude the storage of the required twelve-dimensional error transition matrix. In the CMC rendezvous navigation procedure (and in the LMC as well), the estimated state vector of one vehicle only is altered by the navigation data. Consequently, the error transition matrix is of manageable six-dimensional size and is extrapolated with whichever state vector is being estimated.

Because of this simplification, together with some other related factors, the error transition matrix does not maintain an accurate representation of the covariances of the relative state errors between the two vehicles. This situation necessitates a re-initialization of the matrix at various times during the rendezvous, essentially to restart the solution in both the CMC and LMC.

5.1 Optical Measurements

Approximately once per minute during the navigation portions of the rendezvous phase, the Apollo navigator sights the LM using the reticule pattern in the SXT star line of sight. When the mark button is depressed, the time of the measurement, the SXT shaft and trunnion angles, and the three IMU gimbal angles, which describe the orientation of the navigation base with respect to the inertially stabilized platform, are all automatically recorded in the CMC.

From these five angles and the known orientation of the inertial platform, the measured direction, $\underline{i}_m$, of the LM as observed from the CM is obtained. It is convenient to consider the unit vector $\underline{i}_m$ as having been found by the simultaneous measurement of the angles between the lines of sight to the LM and two stars. The data are processed by selecting two convenient unit vectors (fictitious star directions), converting the vector $\underline{i}_m$ to an equivalent set of two artificial star-LM measurements, and using the measurement incorporation procedure of Section 3 twice - once for each artificial measurement. These two unit vectors are chosen such that they and the estimated line of sight vector form an orthogonal triad.

The processing of each of the two artificial measurements is similar to the cislunar-midcourse navigation procedure described in Section 4. Let $\underline{r}_{PC}$ and $\underline{r}_{PL}$ be the CM and LM position vectors, respectively, extrapolated to the time of the measurement.

For convenience, the first fictitious star direction is chosen to be

$$\underline{i}_{s1} = \text{Unit}\left[(\underline{i}_{CL} \times \underline{i}_m) \times \underline{i}_{CL}\right] , \tag{5.1}$$

where

$$\underline{i}_{CL} = \text{Unit}\,(\underline{r}_{PL} - \underline{r}_{PC})$$

is the estimated CM-to-LM line of sight. The measured deviation is then given by

$$\delta Q = \cos^{-1}(\underline{i}_{s1} \cdot \underline{i}_m) - \frac{\pi}{2} , \tag{5.2}$$

since $\underline{i}_{s1}$ and $\underline{i}_{CL}$ are perpendicular. This orthogonality also permits a simplification in the calculation of the geometry vector $\underline{h}_0$. Indeed, from Equation (4.4), we have

$$\underline{h}_0 = \pm \frac{1}{r_{CL}} \underline{i}_{s1} . \tag{5.3}$$

The plus or minus sign is selected, respectively, according as the CM or LM state vector is to be updated.

Following the alteration of the appropriate state vector, the second artificial measurement is incorporated by recomputing the vector $\underline{i}_{CL}$ and selecting the fictitious star direction as

$$\underline{i}_{s2} = \text{Unit}\,(\underline{i}_{s1} \times \underline{i}_{CL}) . \tag{5.4}$$

The measurement error variance $\overline{\alpha^2}$ for each incorporation is a constant which is the sum of the assumed error variances of the SXT and knowledge of the IMU orientation.

5.2 VHF Range-Link Measurements

Asynchronous with the manually-acquired optical data, and at approximately the same frequency, VHF range measurements are automatically taken. Again, the time of the measurement and the measured CM-to-LM range R are recorded in the CMC.

The measured deviation for this range measurement is simply

$$\delta Q = R - r_{CL} . \tag{5.5}$$

To determine the geometry vector $\underline{h}_0$, we compute the first-order differential of

$$R^2 = (\underline{r}_{PL} - \underline{r}_{PC}) \cdot (\underline{r}_{PL} - \underline{r}_{PC})$$

under the assumption that it is the CM state vector which is to be estimated. There results

$$R\delta R = -(\underline{r}_{PL} - \underline{r}_{PC}) \cdot \delta\underline{r}_{PC}$$

or

$$\delta R = -\underline{i}_{CL} \cdot \delta\underline{r}_{PC} .$$

On the other hand, if the LM state vector is to be updated, the relationship would be

$$\delta R = \underline{i}_{CL} \cdot \delta\underline{r}_{PL} .$$

Hence

$$\underline{h}_0 = \mp \underline{i}_{CL} , \tag{5.6}$$

where the upper or lower sign depends, respectively, on whether the CM or LM state vector is to be changed.

6. LUNAR MODULE NAVIGATION DURING RENDEZVOUS

While the CMC is actively acquiring and processing data during the rendezvous phase of Apollo, the same navigation problem is being simultaneously and independently solved in the LMC, using tracking information gathered from the rendezvous radar mounted on the LM vehicle. After RR tracking acquisition of the CM is established, the LMC records, at approximately one minute intervals, the measured range R and range rate $\dot{R}$ of the CM with respect to the LM, together with the shaft angle β and the trunnion angle θ of the gimballed radar dish. In addition to these four measured quantities, the time of the measurement and the three IMU gimbal angles are also noted.

As described in Section 1, eight variables are estimated as part of the navigation procedure. The usual three components of the position and velocity vectors, $\underline{r}_{PV}$ and $\underline{v}_{PV}$, of the selected spacecraft, LM or CM, with respect to the primary body P, constitute the first six components of the state vector. The estimates of the biases, $\delta\beta$ and $\delta\theta$, in the RR shaft and trunnion angles are the seventh and eighth elements. A dummy variable is used for the ninth component to facilitate three-dimensional vector operations.

The measured quantities produce four sequential alterations of the nine-dimensional state vector through four separate applications of the measurement incorporation method described in Section 3. The incorporation is, of course, performed recursively, i.e. at each stage the new components of the state vector, resulting from the previous update, are used in computing the next update in the sequence. The measurement error variances α^2 are based on *a priori* knowledge of the radar performance.

6.1 Range and Range-Rate Measurements

The measured deviation in range and the associated measurement geometry vector $\underline{h}_0$ are the same as for the VHF measurement of range in the CM and need not be further discussed. For the range-rate data, the measured deviation is easily seen to be

$$\delta Q = \dot{R} - (\underline{v}_{PC} - \underline{v}_{PL}) \cdot \underline{i}_{LC} . \tag{6.1}$$

The geometry vector $\underline{h}$ for the range-rate measurement provides the first case encountered thus far for which both the $\underline{h}_0$ and $\underline{h}_1$ partitions are different from zero. To determine the $\underline{h}$ vector, we compute the first-order differential of the relation

$$\dot{R} = \frac{(\underline{r}_{PC} - \underline{r}_{PL}) \cdot (\underline{v}_{PC} - \underline{v}_{PL})}{[(\underline{r}_{PC} - \underline{r}_{PL}) \cdot (\underline{r}_{PC} - \underline{r}_{PL})]^{1/2}}$$

with the assumption that it is the CM state vector which is to be estimated. We have

$$\delta\dot{R} = \frac{1}{r_{LC}} (\underline{v}_{LC} \cdot \delta\underline{r}_{PC} + \underline{r}_{LC} \cdot \delta\underline{v}_{PC}) - \frac{1}{r_{LC}^3} (\underline{r}_{LC} \cdot \underline{v}_{LC}) \underline{r}_{LC} \cdot \delta\underline{r}_{PC}$$

or, alternatively,

$$\delta\dot{R} = \frac{1}{r_{LC}^3} \underline{r}_{LC} \times (\underline{v}_{LC} \times \underline{r}_{LC}) \cdot \delta\underline{r}_{PC} + \frac{1}{r_{LC}} \underline{v}_{LC} \cdot \delta\underline{v}_{PC} .$$

To obtain the proper relationship if the LM state vector is to be estimated, we need only note that $\delta\underline{r}_{PL} = \delta\underline{r}_{PC}$. Thus, the non-zero partitions of the $\underline{h}$ vector are

$$\left.\begin{aligned} \underline{h}_0 &= \pm \frac{1}{r_{LC}} \underline{i}_{LC} \times (\underline{v}_{LC} \times \underline{i}_{LC}) \\ \underline{h}_1 &= \pm \underline{i}_{LC} \end{aligned}\right\} \tag{6.2}$$

with the choice in sign dependent on the particular state vector to be updated (plus for CM and minus for LM).

6.2 Radar Antenna Angle Measurements

The rendezvous radar antenna dish is at the origin of the RR cartesian coordinate system, as seen in Figure 9. Shaft motion takes place about the positive -y axis and the shaft angle β is measured from the positive x axis. Trunnion motion occurs in a plane normal to the xz plane and containing the shaft axis y. The configuration is such that the trunnion axis would coincide with the x axis for zero shaft angle with the trunnion angle θ, under those circumstances, measured from the z axis.

Let $\underline{i}_x$, $\underline{i}_y$, $\underline{i}_z$ be unit vectors along the RR coordinate axes. By means of the recorded IMU gimbal angles, together with the knowledge of the inertial orientation of the IMU, the components of $\underline{i}_x$, $\underline{i}_y$, $\underline{i}_z$ are readily

obtained in basic reference coordinates. It is easy to verify with reference to the figure that the measured deviation for the shaft angle is

$$\delta Q = \beta - \left[\tan^{-1} \left(\frac{i_x \cdot i_{LC}}{i_z \cdot i_{LC}} \right) + \delta\beta \right] . \tag{6.3}$$

while that for the trunnion angle is

$$\delta Q = \theta - [\sin^{-1} (-i_y \cdot i_{LC}) + \delta\theta] . \tag{6.4}$$

The position vector of the CM with respect to the LM is expressible in terms of components along the RR coordinate axes as

$$r_{PC} - r_{PL} = r_{LC} \begin{pmatrix} \sin\beta\cos\theta \\ -\sin\theta \\ \cos\beta\cos\theta \end{pmatrix}$$

To obtain the measurement geometry vector corresponding to the shaft angle, we again compute the differential

$$\delta r_{PC} = r_{LC} \begin{pmatrix} \cos\beta\cos\theta \\ 0 \\ -\sin\beta\cos\theta \end{pmatrix} \delta\beta$$

$$= r_{LC}\cos\theta \ \text{Unit}\,(i_y \times i_{LC})\,\delta\beta ,$$

Hence

$$\delta\beta = \frac{1}{r_{LC}\cos\theta} \ \text{Unit}\,(i_y \times i_{LC}) \cdot \delta r_{PC} .$$

Similarly, for the trunnion angle, we have

$$\delta r_{PC} = r_{LC} \begin{pmatrix} -\sin\beta\sin\theta \\ -\cos\theta \\ -\cos\beta\sin\theta \end{pmatrix} \delta\theta$$

$$= \frac{r_{LC}}{\cos\theta} (i_y \times i_{LC}) \times i_{LC}\delta\theta ,$$

so that

$$\delta\theta = \frac{1}{r_{LC}\cos\theta} (i_y \times i_{LC}) \times i_{LC} \cdot \delta r_{PC} .$$

The quantity $r_{LC}\cos\theta$, which appears in the expressions for $\delta\beta$ and $\delta\theta$, is simply the length of the projection of the r_{LC} vector in the xz plane. By denoting this length by r_{xz}, we may calculate its value from

$$r_{xz} = r_{LC}\sqrt{[1 - (i_y \cdot i_{LC})]} . \tag{6.5}$$

$$b_0 = \pm \frac{1}{r_{xz}} \ \text{Unit}\,(i_y \times i_{LC}) \tag{6.6}$$

and, for the trunnion angle vector,

$$b_0 = \pm \frac{1}{r_{xz}} (i_y \times i_{LC}) \times i_{LC} . \tag{6.7}$$

where the plus or minus sign, as before, indicates either the CM or LM state vector is to be updated.

Since the antenna angles are independent of velocity, we conclude that $b_1 = 0$ for both. However, because we are also estimating the angle biases, the b_2 partitions of the b vectors are not zero. Indeed, for the shaft angle,

$$b_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

and, for the trunnion angle,

$$b_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} .$$

The estimation of the radar biases is included regardless of whether it is the CM or LM state vector which is being estimated. Since the radar biases do not depend on which vehicle's state vector comprises the first six components of the estimated state, there is no sign selection associated with these h_2 partitions.

7. COMMAND MODULE ORBIT NAVIGATION

When the Apollo spacecraft is in either earth or lunar orbit, navigation data can be obtained by optical measurements of the lines of sight to landmarks. These planetary surface features can be either of the "known" or "unknown" variety. A known landmark is an identifiable feature whose coordinates are known and tabulated. In contrast, an unknown landmark is any surface feature which the astronaut may select and optically track in the brief period during which it is visible. The mechanics of the measuring process is quite similar to the CM-to-LM line of sight measurement procedure described in Section 5.

7.1 Known Landmark Measurements

As indicated in Section 1, orbit navigation involves a nine-dimensional state vector, the last three components of which are the coordinates of the landmark. Since the tracking period for one landmark is very short (less than one minute in the case of the earth and only two or three minutes during lunar orbit), all navigation data for any particular landmark are acquired before the processing begins. At the conclusion of the tracking, the landmark partition of the state vector is initialized from the identification data entered by the navigator into the computer.

Sufficient CMC erasable storage is allocated for five measurements on a single landmark. The data from each of these sightings consists of the time of the measurement and the set of angles described in Section 5. From these angles the measured unit vector $\underline{i}_m$ along the CM-to-landmark line of sight is computed.

Each of the measured unit vectors is converted to an equivalent set of two artificial star-landmark measurements in exactly the same manner as in the rendezvous navigation procedure discussed in Section 5. The only difference is that the geometry vector is nine-dimensional and is given by

$$\underline{h} = \frac{1}{r_{VL}} \begin{pmatrix} \underline{i}_s \\ \underline{0} \\ -\underline{i}_s \end{pmatrix}, \tag{7.1}$$

where r_{VL} is the current estimated distance between the vehicle and the landmark, and $\underline{i}_s$ denotes the direction to the artificial stars.

7.2 Unknown Landmark Measurements

The nine-dimensional orbit navigation technique provides a means of mapping on the surface of a planet a point which is designated only by a number of sets of optical tracking data. This process may be used either to locate a desired LM landing site which may have unknown coordinates or to map features on the surface of the moon.

Assume that a landmark has been tracked and N sets of optical navigation data have been acquired. If the navigator cannot -- or chooses not to -- identify the landmark, it is then treated as an unknown landmark. In this process the data from the first navigation measurement are used to compute an initial estimate of the landmark location. The nine-dimension state vector is then formed, and the data from the remaining $N-1$ sightings are incorporated exactly as if the optically-designated point had been an identified landmark.

The determination of the initial estimate of the landmark position vector $\underline{r}_L$ is accomplished as follows. Let $\underline{i}_m$ be the measured unit vector from the vehicle to the landmark calculated from the data of the first measurement. Then

$$\underline{r}_L = \underline{r}_{PV} + r_{VL}\underline{i}_m , \tag{7.2}$$

with r_{VL} obtained by applying the law of cosines to the triangle defined by Equation (7.2). We then have

$$r_{VL}^2 - (2r_{PV}\cos A_L)r_{VL} + (r_{PV}^2 - r_P^2) = 0 , \tag{7.3}$$

where r_P is the radius of the primary body, and A_L is the angle between the directions from the vehicle to the center of the primary body and the landmark and is computed from

$$\cos A_L = -\underline{i}_m \cdot \text{Unit}(\underline{r}_{PV}) .$$

Solving Equation (7.3) and selecting the appropriate sign yields

$$r_{VL} = r_{PV}\left[\cos A_L - \left(\frac{r_P^2}{r_{PV}^2} - \sin^2 A_L\right)^{1/2}\right] . \tag{7.4}$$

Thus, the initial estimated location of the unknown landmark is established. The remaining $N-1$ sets of data are then used as for a known landmark, according to Section 7.1.

7.3 Initialization of the Error Transition Matrix

Several different landmarks are tracked by the Apollo navigator to navigate in orbit about the earth or moon. As a consequence, special methods are required to reinitialize the error transition matrix each time a new landmark is acquired. This is necessary to reflect the fact that the initial landmark location errors are not correlated with the errors in the estimated CM position and velocity vectors. Therefore, before processing the measurement data associated with a new landmark, it is necessary to convert the nine-dimensional error transition matrix $\underline{M}$ to a six-dimensional matrix having the same CM position and velocity error variances and covariances. A new nine-dimensional matrix is then formed by augmenting appropriate landmark uncertainty information.

As the first stage in the initialization process it is necessary to determine a square root of the six-dimensional partition $\underline{M}_{44}$ of the nine-dimensional covariance matrix $\underline{M}$, defined in Equation (2.19). From Equations (2.21) and (3.9), it is clear that

$$\underline{M}_{44} = \begin{pmatrix} \underline{W}_3^T & \underline{W}_1^T & \underline{W}_2^T \\ \underline{W}_3^T & \underline{W}_4^T & \underline{W}_5^T \end{pmatrix} \begin{pmatrix} \underline{W}_0 & \underline{W}_3 \\ \underline{W}_1 & \underline{W}_4 \\ \underline{W}_2 & \underline{W}_5 \end{pmatrix} . \tag{7.5}$$

The square root of $\underline{M}_{44}$, i.e., a six-dimensional matrix $\underline{N}_{44}$ such that

$$\underline{N}_{44}\underline{N}_{44}^T = \underline{M}_{44} , \tag{7.6}$$

is not unique. However, a convenient method, which is used in the CMC, is to select a triangular form for $\underline{N}_{44}$, so that the resulting set of algebraic equations for the elements of $\underline{N}_{44}$ is most easily solved.

The second part of the $\underline{M}$ matrix initialization depends on whether the landmark being tracked is known or unknown. For a known landmark, the new nine-dimensional $\underline{M}$ matrix is formed as:

$$\underline{M} = \begin{pmatrix} \underline{W}_0^T & \underline{W}_1^T & \underline{0} \\ \underline{W}_3^T & \underline{W}_4^T & \underline{0} \\ \underline{0} & \underline{0} & \underline{W}_8^T \end{pmatrix} , \tag{7.7}$$

with the upper left-hand six by six partition being the square root matrix $\underline{N}_{44}$ found in part one. The three-dimensional sub-matrix $\underline{W}_8$ is given a value consistent with the expected errors in the knowledge of the landmark location. In particular, $\underline{W}_8$ is chosen so that

$$\underline{W}_8^T\underline{W}_8 = \underline{\zeta}\underline{\zeta}^T , \tag{7.8}$$

where $\underline{\zeta}$ is the vector error in the landmark position.

If, on the other hand, the measurements are made using an unknown landmark, the error in the initially computed estimate of the location of the landmark will be a function of the uncertainties in the CM position estimate, the tracking accuracy and the altitude of the landmark above the gravitational center (see Reference 4). In this case, part two for the initialization of the $\underline{M}$ matrix is more complicated.

The basic relationship among the quantities of interest is Equation (7.2). Errors in $\underline{r}_{PV}$ and $\underline{i}_m$, denoted respectively by $\underline{q}$ and $\delta\underline{i}_m$, will produce an error in the landmark location $\underline{r}_L$. However, this error is clearly in the plane of the landmark, i.e., perpendicular to $\underline{r}_L$. Also, an error in r_{VL} will result in a landmark position error in the direction of $\underline{i}_m$ and of a magnitude $\delta r_L\sqrt{(\underline{r}_L^T\underline{r}_L)}/\underline{i}_m^T\underline{r}_L$, where δr_L is the error in the landmark altitude. Thus, the total error $\underline{\zeta}$ in the landmark location is given by

$$\underline{\zeta} = \underline{U}(\underline{q} + r_{VL}\delta\underline{i}_m) + \delta r_L \frac{\sqrt{(\underline{r}_L^T\underline{r}_L)}}{\underline{i}_m^T\underline{r}_L}\underline{i}_m , \tag{7.9}$$

where

$$\underline{U} = \underline{I} - \frac{\underline{i}_m\underline{r}_L^T}{\underline{i}_m^T\underline{r}_L} \tag{7.10}$$

is the projection operator which assures that only the components of $\underline{q}$ and $\delta\underline{i}_m$ in the plane normal to $\underline{r}_L$ are related to the landmark location error.

Now, consider a coordinate system in which the direction $\underline{i}_m$ is along one of the coordinate axes. Then, if $\underline{T}$ is the transformation matrix which relates the selected axis system and the original reference system, we have

$$\underline{i}_m = T\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

The error in $\underline{i}_m$ may be expressed as:

$$\delta\underline{i}_m = T\begin{pmatrix}\alpha\cos\chi\\ \alpha\sin\chi\\ 0\end{pmatrix},$$

where α is the small random angle between the true and measured directions to the landmark. The polar angle χ is defined in the plane normal to $\underline{i}_m$ from the coordinate axis to the projection of the measured direction of $\underline{i}_m$.

Assume that α and χ are statistically independent random variables with zero means. Further, assume that χ is uniformly distributed over the interval $-\pi$ to π. Then, for the covariance matrix of the uncertainty in $\underline{i}_m$, we obtain

$$\overline{\delta\underline{i}_m\delta\underline{i}_m^T} = \tfrac{1}{2}\sigma_\alpha^2 T\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}T^T$$

$$= \tfrac{1}{2}\sigma_\alpha^2(I - \underline{i}_m\underline{i}_m^T),$$

where σ_α^2 is the variance of α.

Finally, since $\underline{\varepsilon}$ and the velocity error $\underline{\eta}$ are statistically independent of $\delta\underline{i}_m$ and $\delta\underline{r}_L$, we may calculate the following covariance matrices from Equation (7.9):

$$\left.\begin{aligned}\overline{\underline{\zeta}\underline{\alpha}^T} &= \underline{U}\overline{\underline{\varepsilon}\underline{\alpha}^T}\\ \overline{\underline{\zeta}\underline{\eta}^T} &= \underline{U}\overline{\underline{\varepsilon}\underline{\eta}^T}\\ \overline{\underline{\zeta}\underline{\zeta}^T} &= \underline{U}\overline{\underline{\varepsilon}\underline{\varepsilon}^T}\underline{U}^T + \tfrac{1}{2}\sigma_\alpha^2 r_{VL}^2\underline{U}(I - \underline{i}_m\underline{i}_m^T)\underline{U}^T + \sigma_L^2\frac{\underline{r}_L^T\underline{r}_L}{(\underline{i}_m^T\underline{r}_L)^2}\underline{i}_m\underline{i}_m^T.\end{aligned}\right\} \quad (7.11)$$

As a result, the $\underline{W}$ matrix is initialized as

$$\underline{W} = \begin{pmatrix}\underline{W}_0^T & \underline{E}_1^T & \underline{0}\\ \underline{W}_3^T & \underline{W}_4^T & \underline{0}\\ \underline{U}\underline{W}_0^T & \underline{U}\underline{W}_1^T & \underline{W}_5^T\end{pmatrix}, \quad (7.12)$$

where $\underline{W}$ is now determined as the three-dimensional triangular square root of

$$\underline{W}_5^T\underline{W}_5 = \tfrac{1}{2}\sigma_\alpha^2 r_{VL}^2\underline{U}(I - \underline{i}_m\underline{i}_m^T)\underline{U}^T + \sigma_L^2\frac{\underline{r}_L^T\underline{r}_L}{(\underline{i}_m^T\underline{r}_L)^2}\underline{i}_m\underline{i}_m^T. \quad (7.13)$$

After incorporating the data obtained from the tracking of a landmark, all 81 elements of the $\underline{W}$ matrix will again, in general, be non-zero, indicating correlations among all components of vehicle position and velocity and landmark position. This matrix is extrapolated, as described in Section 3, until a new landmark is acquired necessitating a new initialization.

8. FLIGHT EXPERIENCE OF THE APOLLO NAVIGATION SYSTEM

The first manned trip to the vicinity of the moon of Apollo 8 during December 1968 gave an excellent test of the Apollo system's on-board navigation capability (see Reference 5). Although ground tracking navigation was the primary system, the on-board navigation system had the task of confirming a safe trajectory and providing a back-up for return to earth in the remote chance that ground assistance became unavailable for on-board use.

Apollo 8 was to use sun illuminated visual horizons rather than landmarks for operational simplicity, even though, as confirmed from earth orbit by Astronaut Don Eisele in Apollo 7, the earth's horizon does not provide a distinct target for visual use. Moreover, the filter in the sextant beamsplitter, designed originally to enhance the contrast between water and land when looking down at the earth, filters out the blue in such a way as to make the horizon even more indistinct. Originally a blue sensitive photometer had been designed for horizon detection in the prototype sextant models, but was removed from the production systems, since a decision had been made that ground tracking would be the primary source of midcourse navigation. Without the photometer, interest in the earth's horizon as a visual target resulted in demonstrations on simulators that, in some subjective way, the human with a little experience can choose an altitude sufficiently repeatable, at least as good as ±3 kilometers. Accordingly, a few weeks before the Apollo 8 launch, the navigator command module pilot, Jim Lovell, spent a few

hours on the sextant earth horizon simulator at the Massachusetts Institute of Technology in Cambridge for training and to calibrate the horizon altitude he seemed to prefer. He was remarkably consistent in choosing a location 32.8 kilometers above the sea level horizon. This value was recorded in the CMC as part of the pre-launch mission load.

The plan for the mission was to examine the sextant angle measurements made early while still near the earth and, based on the spacecraft state vector determined by ground tracking, infer in real-time the horizon altitude Lovell was using. After the first eleven sightings on the earth at distances of about 30 thousand nautical miles from earth, it was estimated that he was using an 18.3 kilometers altitude and the CMC was reloaded with this new value. (Later during the mission it was agreed that a truer estimate was nearer 23 kilometers, but the value was not changed since the difference then was too small to be of concern.) Following the horizon calibration, the first midcourse correction of almost 25 ft/sec was performed. The large size of this correction was due to trajectory perturbations, resulting from the maneuvers performed in getting the spacecraft safely away from the third stage of the launch vehicle.

After the midcourse correction, the CMC state vector was made to agree with the value obtained from ground tracking. The important parameter, predicted perilune altitude, was 69.7 nautical miles - very close to the true value estimated later to be 68.8. The next 31 navigation measurements were made using the earth's horizon, modeled at 18.2 kilometers altitude. Being sufficiently far from both earth and moon during this time, it is not surprising that the initially good state vector was degraded. At the end of this period, the indicated perilune was 32 nautical miles below the moon's surface. With the next nine sightings, still using the earth's horizon, the predicted perilune increased to 92.9 nautical miles - about 32 nautical miles too high. The exact altitude of the earth's horizon was unimportant for these sightings since the distance from earth was now approximately 150,000 nautical miles, so that the 10 arc-second accuracy of the sextant was the predominant source of error.

The next group of 16 sightings was made using the moon horizon at a distance of about 80,000 nautical miles. As would be expected, the first few of these resulted in fairly large changes in the estimated state vector, while the remaining had a very small effect. At the end of this group of measurements, the indicated perilune was 67.1 nautical miles. The final set of 15 translunar sightings was made about 35,000 nautical miles from the moon with little additional effect on the perilune estimation. The final estimate was 67.5 nautical miles or about 1.3 nautical miles lower than the value later reconstructed from ground tracking data. At this time the on-board and ground tracking data were practically identical and consideration was given to using the on-board state vector for lunar orbit insertion. Although the state vector update hardly changed the on-board value, it was performed since there was no overriding argument to deviate from the flight plan.

The transearth flight of Apollo 8 after 10 lunar orbits also provided a good measure of the on-board navigation capability. The transearth injection maneuver of the service propulsion system was targetted by ground data and executed in back of the moon by the on-board digital autopilot and guidance systems. This 3522.5 ft/sec maneuver was followed by a single midcourse correction of 4.8 ft/sec 14.7 hours later, resulting in entry conditions at 400,000 feet altitude above the earth which were 0.8 ft/sec faster and 0.1° shallower than planned.

Although the primary navigation during this period was again the ground tracking network, 138 on-board navigation measurements were performed by Lovell as a monitor and back-up. In order to determine what would have happened without ground assistance, the actual on-board measurements were incorporated in a simulation with the computer initialized to the actual on-board state vector as it existed when the spacecraft emerged from behind the moon. The single transearth midcourse correction was added appropriately to this simulation, in accordance with that actually measured by the inertial guidance system. (In the actual flight a new ground determined state vector was loaded into the computer at the time of this maneuver.)

The last of the 138 measurements was completed 16 hours before entry. The incorporation of these measurements in the simulation left a hypothetical on-board estimate of entry flight path angle at 400,000 feet of -6.38°, as compared with the ground tracking estimate of -6.48°. This 0.1° difference was well within the safe tolerance of $\pm 0.5^{\circ}$. Another parameter of concern at the entry interface is the error in knowledge of altitude rate. The simulated on-board estimate of this quantity differed from that estimated by ground tracking by 236 ft/sec. However, the conservative allowable tolerance is ±200 ft/sec.

It should be emphasized that, in the event ground data were not available, the plan was to continue the on-board measurements to optimize the final midcourse correction and state vector for safe earth atmospheric entry. In the absence of actual flight data, a continuation of the simulation using the planned sighting program was made with standard deviation errors in the sextant of 10 arc-seconds and in the horizon of 3 kilometers. In addition, bias errors of 5 arc-seconds in the sextant and 4 kilometers in the horizon were included. The resulting estimation error in the entry angle at the entry interface had a standard deviation of 0.03° and a bias of 0.007°. The corresponding altitude rate uncertainty had a standard deviation of 41.1 ft/sec and a bias of 26.5 ft/sec. The capability of the on-board navigation system to bring the spacecraft safely back from the moon seems clearly to have been demonstrated.

REFERENCES

1. Miller, J.E. (Editor) et al. — *Space Navigation Guidance and Control.* Technivision Ltd, Maidenhead, England, 1966.

2. Battin, R.H. — *Astronautical Guidance.* McGraw-Hill, New York, 1964.

3. Henrici, P. — *Discrete Variable Methods in Ordinary Differential Equations.* John Wiley, New York, 1962.

4. Ballantoni, J.F. — *Unidentified Landmark Navigation.* AIAA Journal, Vol. 5, August 1967, pp. 1478-1483.

5. Hong, D.G. — *Apollo Control, Guidance and Navigation System - A Progress Report.* Presented at the National Space Navigation Meeting of the Institute of Navigation, Houston, Texas, April 1969.

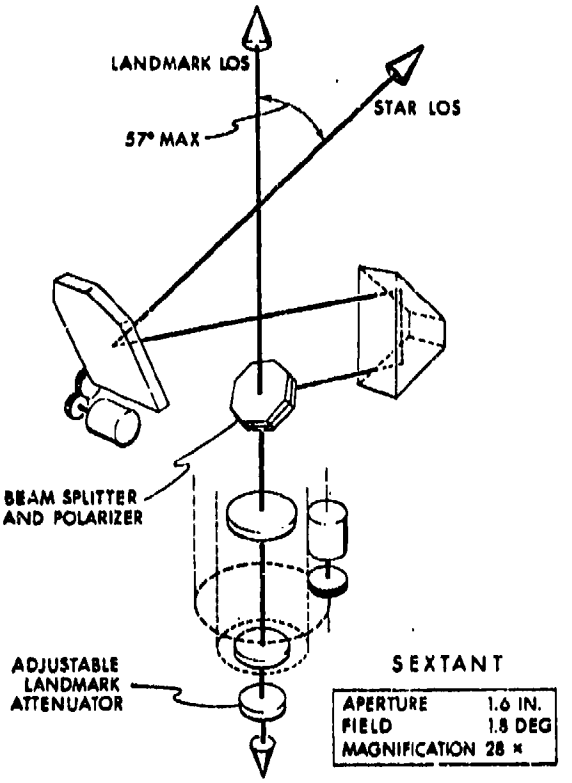


Fig. 1 Sextant schematic diagram

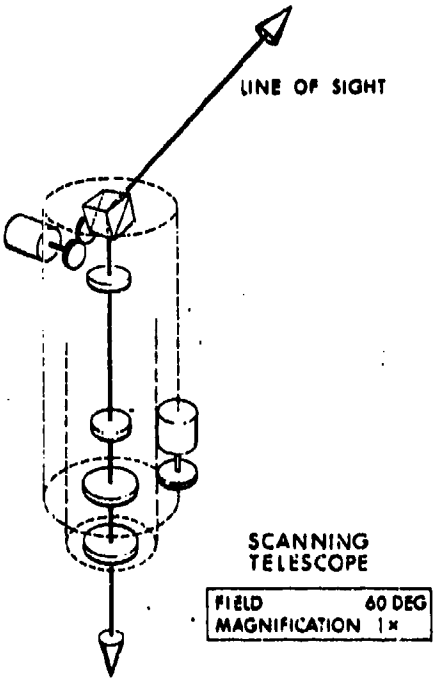


Fig. 2 Scanning telescope schematic diagram

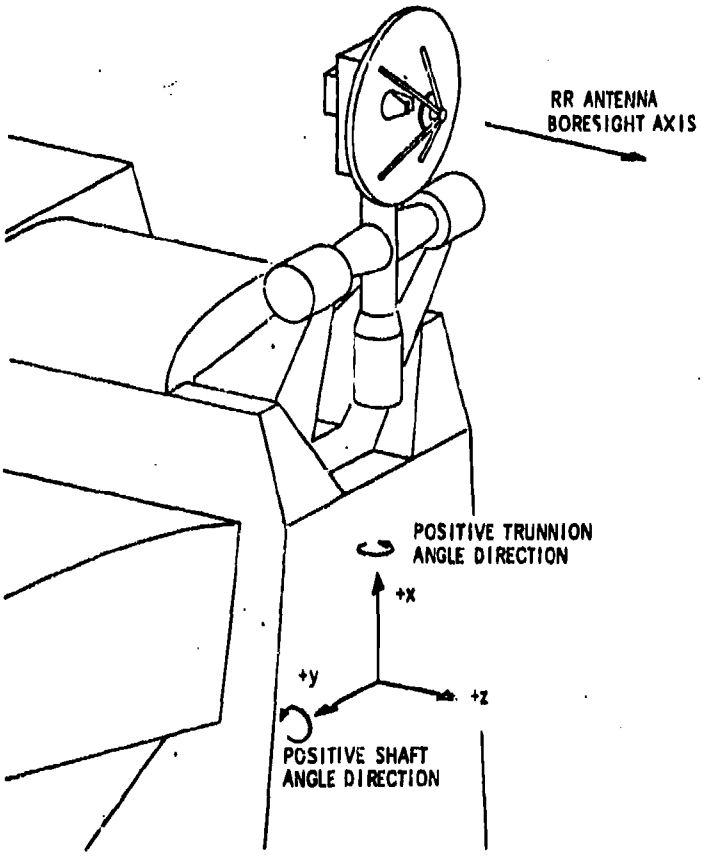


Fig.3 Rendezvous radar

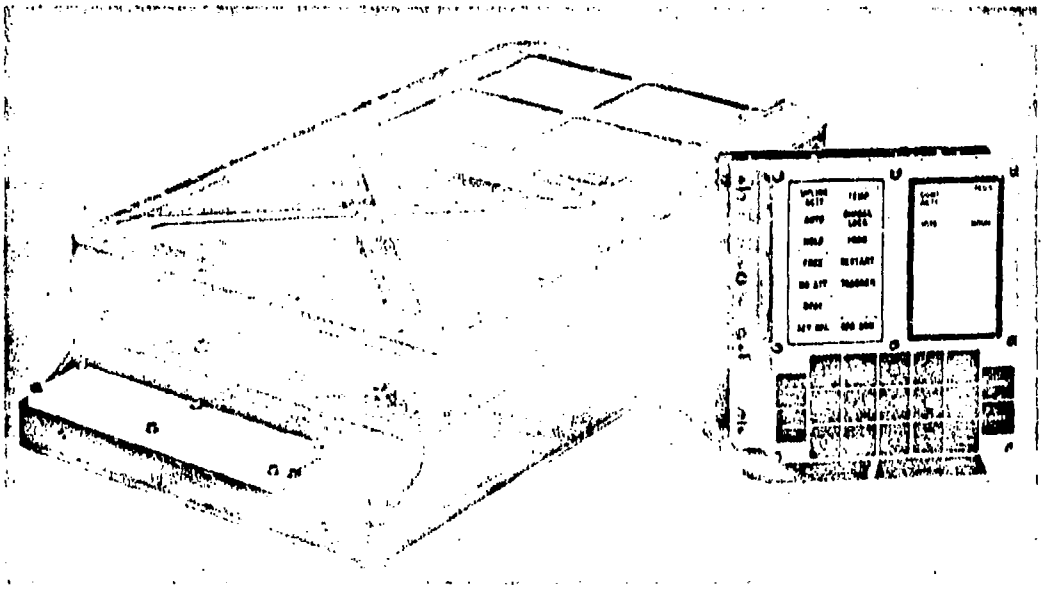


Fig.4 Apollo guidance computer and display and keyboard

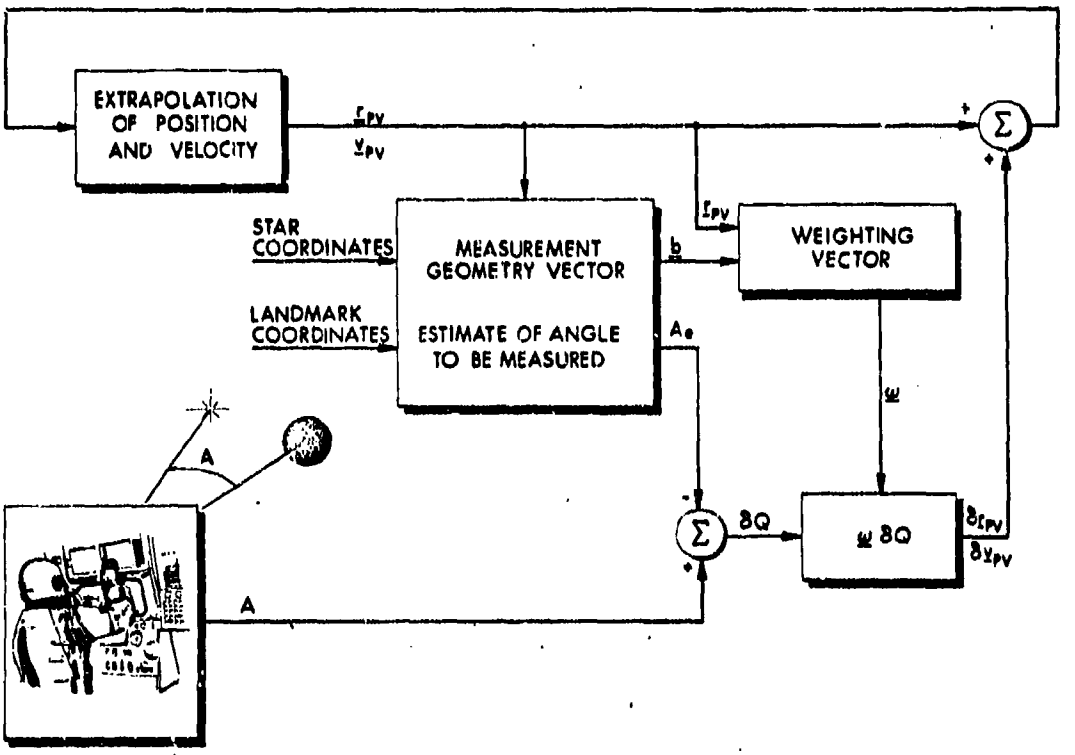


Fig. 5 Coasting flight navigation

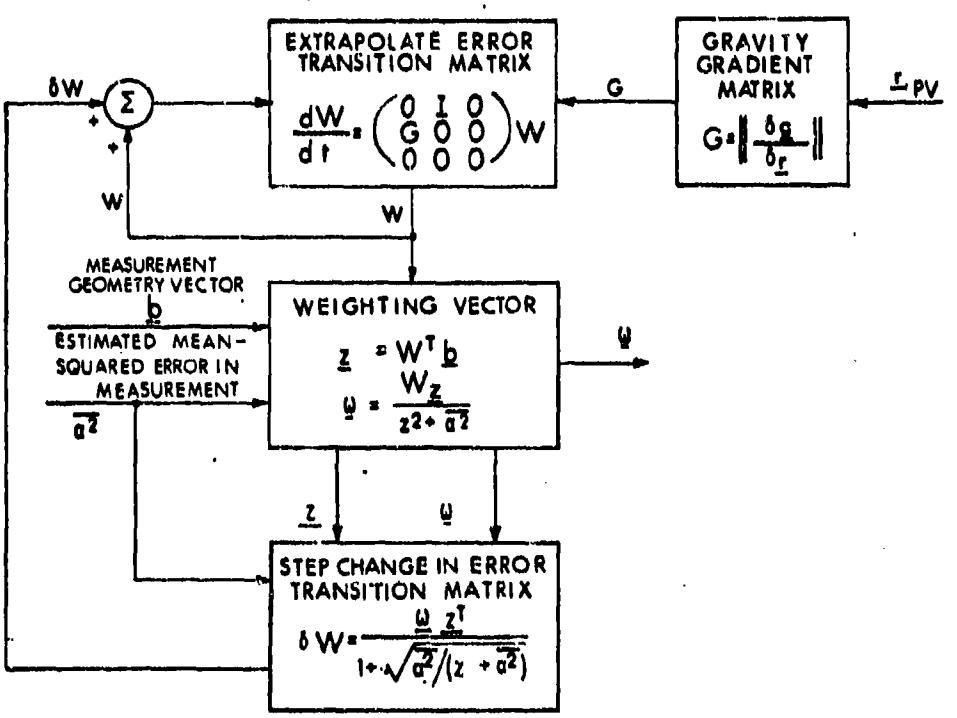


Fig. 6 Weighting vector calculation

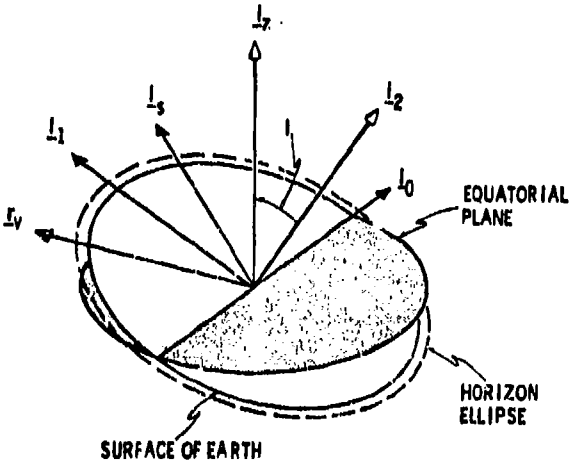


Fig. 7 Horizon coordinate system

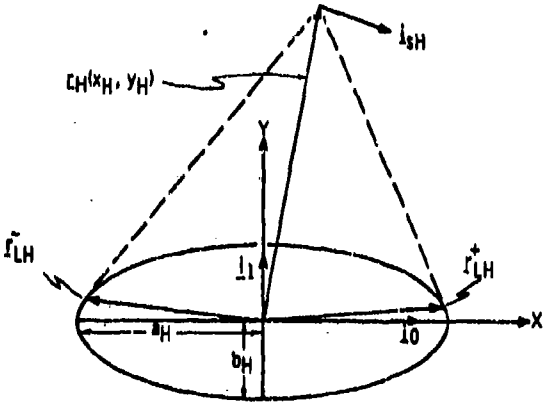


Fig. 8 Geometry of star-horizon measurement

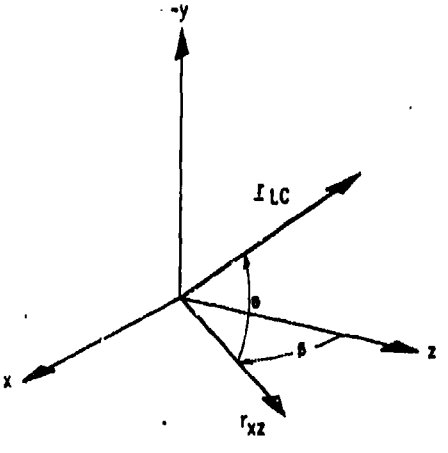


Fig. 9 Rendezvous radar coordinate system

CHAPTER 15 - SOME APPLICATIONS OF KALMAN FILTERING IN SPACE GUIDANCE

by

J.F. Bellantoni

National Aeronautics and Space Administration
Electronics Research Center
Cambridge, Massachusetts

The material in this chapter represents the views of its author and is not to be construed as an official position of the US Government or of the Electronics Research Center.

CHAPTER 15 - SOME APPLICATIONS OF KALMAN FILTERING IN SPACE GUIDANCE

J.F. Bellantoni

1. INTRODUCTION

Among the first applications of linear optimal recursive filtering were those made by Battin[1] and by McLean, Smith, Schmidt and McGee[2,22] to the estimation of the position and velocity of a spacecraft. The field of space navigation still provides a fruitful area of application of this technique. Two of the applications described in this chapter deal directly with space navigation: the error analysis of a satellite navigation system, and unidentified landmark navigation. The third application deals with initial alignment of an inertial guidance system that will indicate vehicle position and velocity during boost into a free-fall trajectory.

2. ERROR ANALYSIS OF A SATELLITE NAVIGATION SYSTEM

2.1 The Kalman Filter as a Tool in Error Analysis

One of the useful properties of the linear recursive optimal filter (Kalman filter) is that it generates its own error analysis in the process of estimating the state. In fact, the covariance matrix $P(t)$ of error in the estimate $\hat{x}(t)$, upon which the error analysis rests, may be computed recursively over the time of interest without benefit of real data. This is of importance in the preliminary design stages of aircraft and spacecraft guidance systems, since a prediction of system performance may be made knowing only the statistics of the instrument errors and the times of the observations. The basic equations are that, for the propagation of the covariance matrix between measurement times,

$$\frac{d}{dt}P(t) = F(t)P(t) + P(t)F^T(t) + G(t)\Gamma(t)G^T(t) , \tag{2.1}$$

where

$P(t) = E[x(t) - \hat{x}(t)][x(t) - \hat{x}(t)]^T$

$= $ covariance of error in estimate $\hat{x}(t)$

$F(t) = $ system matrix in state equation

$x(t) = $ state vector, given by $\dot{x}(t) = F(t)x(t) + G(t)w(t)$

$\hat{x}(t) = $ Kalman filter estimate of $x(t)$

$\Gamma(t)\delta(t) = E[w(t) - W(t)][w(t) - W(t)]^T$

$w(t) = $ white-noise driving state equation

$W(t) = Ew(t) = 0$

$\delta(t) = $ Dirac delta function of t

$G(t) = $ input matrix for $w(t)$

$E(\ ) = $ expectation of $(\ )$

$(\ )^T = $ transpose of $(\ )$

so that, for the covariance matrix P_i^+ after incorporation of data at time t_i,

$$P_i^+ = P_i^- - P_i^- H_i^T[H_i P_i^- H_i^T + R_i]^{-1} H_i P_i^- . \tag{2.2}$$

where

$$P_i^- = P(t_i^-)$$

$$P_i^+ = P(t_i^+)$$

$$R_i = R(t_i)$$

H_i = observation matrix at time t_i

t_i^-, t_i^+ = times just before and just after observation at t_i.

An alternate form of Equation (2.1), useful when the state transition matrix $\phi(t_i,t_{i-1})$ is known or may be approximated, is

$$P(t_i^-) = \phi(t_i,t_{i-1})P(t_{i-1}^+)\phi^T(t_i,t_{i-1}) + Q(t_i,t_{i-1}) \,. \qquad (2.1)'$$

where

$$Q(t_i,t_{i-1}) = \int_{t_{i-1}}^{t_i} \phi(t_i,\tau)G(\tau)\Gamma(\tau)G^T(\tau)\phi^T(t_i,\tau)\,d\tau \,. \qquad (2.3)$$

If approximations, adequate for the purposes of error analysis, are available for ϕ and Q, then Equation (2.1)' is superior to Equation (2.1) in that it avoids the integration implicit in Equation (2.1).

The assumptions underlying this type of error analysis impose limitations that may preclude its use in some applications. The major assumptions are:

(1) The state vector $x(t)$ is adequately described by a linear differential equation driven by white-noise:

$$\dot{x}(t) = F(t)x(t) + G(t)w(t) \,, \qquad (2.4)$$

where $F(t)$ and $G(t)$ are known.

(2) The observations $z(t_i)$ are linearly related to the state $x(t_i)$ by

$$z(t_i) = H(t_i)x(t_i) + v(t_i) \,, \qquad (2.5)$$

where $H(t_i)(=H_i)$ is known and $v(t_i)$ is white-noise.

(3) The mean and variance of $w(t)$ and $v(t)$ are known.

(4) The minimum variance linear filter $\hat{K}_i$ is used to improve the state estimate, as follows:

$$\hat{x}(t_i^+) = \hat{x}(t_i^-) + \hat{K}_i[z(t_i) - H_i\hat{x}(t_i^-)] \,. \qquad (2.6)$$

The last assumption (4) is not essential. If the filter is linear, but not necessarily minimum variance, Equation (2.2) may be replaced by the more general expression

$$P_i^+ = (I - K_iH_i)P_i^-(I - K_iH_i)^T + K_iR_iK_i^T \,. \qquad (2.7)$$

where K_i, a general linear filter, has been used in Equation (2.6) instead of $\hat{K}_i$. A simple derivation of Equation (2.7) is given in Appendix A. It is easy to show that Equation (2.7) reduces to Equation (2.2) when $K_i = \hat{K}_i$, where

$$\hat{K}_i = P_i^-H_i^T[H_iP_i^-H_i^T + R_i]^{-1} \,. \qquad (2.8)$$

Equation (2.7) is of use in analyzing the error of a linear, sub-optimal filter. In particular, the sub-optimal filter being analyzed may be a Kalman filter based on (1) an approximate state transition matrix, or (2) a state containing only some of the variables of the problem, or (3) some approximation in the filter designed to save computation time or space.

The choice of state vector, and the choice of models for the variables contained therein, determine the state transition matrix ϕ and the noise covariance Q in Equation (2.1)'. For error analysis purposes it is often possible to make enough approximations in the error models to allow use of the integrated form Equation (2.1)' rather than Equation (2.1). The integrated form will be used exclusively in the remainder of this section.

The choice of measurement types determines the observation matrix H and the covariance R of observation noises in Equation (2.2). Along with the assumed observation times and the initial value P_0 for P, the four matrices

$$\phi \,,\ Q \,,\ H \,,\ R$$

completely determine the computation of the state covariance matrix P and hence characterize the error analysis.

2.2 The Satellite Navigation Problem

The position and velocity of a vehicle in earth orbit are usually calculated from a series of measurements made by ground radar stations and instruments, such as horizon trackers, star trackers and landmark telescopes, located on board the vehicle. In order to compare the merits of various combinations of instruments, it is necessary to estimate the net accuracy with which any particular combination predicts position and velocity and, in some cases, vehicle attitude. If it is assumed that a minimum variance filter is employed to process the data, then Equations (2.1) or (2.1)′ and (2.2) give the second-order statistics of the position and velocity, and of any other variables included in the state vector. If the filter is linear, but not necessarily minimum variance, then Equation (2.2) is to be replaced by Equation (2.7) to achieve the same results.

The scope of the error analysis is essentially determined by the choice of (a) the state vector, (b) the measurement types, and (c) the realism desired in determining when the measurements occur. The last consideration should not be overlooked as a factor in the complexity of a digital computer program.

The most convenient place to start the analysis of a practical filter problem such as this has been found to be with (b), the measurement types, rather than with the state vector. The reason is that many of the variables in the state vector are noises or biases originating in the measuring instruments. In addition, a deft definition of the state variables, suggested by the measurement matrix, often reduces the complexity of the state transition matrix. Hence, the observation types will be discussed in Section 2.3 and the choice of state variables in Section 2.4.

The inertial coordinates X, Y, Z describing the vehicle position are shown in Figure 1, and the vehicle vertical coordinates ξ, η, ζ are shown in Figure 2.

2.3 Observation Types

The purpose of this section is to describe some of the observations that are commonly used in satellite orbit determination. Table I lists the H-matrices corresponding to these observations.

Ground Based Radar

Figure 3 shows the geometry of a tracking station. The four observations are

$$\rho = |\bar{\rho}| = \text{range} \tag{2.9}$$

$$\dot{\rho} = \frac{d}{dt}\rho = \text{range rate} \tag{2.10}$$

$$A = \tan^{-1}(\bar{\rho}\cdot\hat{e}/\bar{\rho}\cdot\hat{n}) = \text{azimuth} \tag{2.11}$$

$$E = \sin^{-1}(\bar{\rho}\cdot\hat{h}/\rho) = \text{elevation}, \tag{2.12}$$

where the symbols used are defined in Figures 1, 2 and 3. Each of these observations is affected by the vehicle position and/or velocity and can be expected to give partial orbit determination information. However, the uncertainties Δe, Δn, Δh of the station's own location are a major source of error, comparable to the noise level on range, and must be considered as variables in the state vector in any realistic error analysis. Table I, therefore, shows the H-matrices corresponding to station location errors as well as to vehicle position and velocity. Finally, the biases on range and range rate and in azimuth and elevation are often important and are also shown in Table I. These biases are denoted by $\Delta\rho$, $\Delta\dot{\rho}$, ΔA and ΔE.

Inertial Measurement Unit

Most experiments carried on a satellite require knowledge of the vehicle's orientation in inertial space in order to produce useful data. Since it is often impractical to control the vehicle's attitude precisely, a gyroscopically stabilized platform is often required. Experimental observations are then made (perhaps using the vehicle itself as an intermediate coordinate frame) relative to the gyroscopically stabilized axes. The stabilized coordinate frame itself may be referenced to inertial space by a star tracker that makes observations of two or more stars. It will be assumed that the gyroscopically stabilized axes are maintained coincident either with the inertial axes XYZ of Figure 1 or with local vertical axes $\xi\eta\zeta$ of Figure 2. Where necessary, the two cases will be distinguished.

Observations made relative to the gyroscopically stabilized (IMU) axes will contain errors due to the discrepancy between the true and the indicated orientations of the IMU. This discrepancy may be represented as small rotations θ_1, θ_2, θ_3 about the IMU gyroscopic axes, i.e. along the XYZ or $\xi\eta\zeta$ axes. These gyro angles $(\theta_1,\theta_2,\theta_3)$ will be needed in the state vector, not only because they are important errors in many of the observations to be described, but because they are a large part of the attitude error of the vehicle. In many missions, attitude has an importance equivalent to that of position and velocity. A suitable model for θ_1, θ_2 and θ_3 will be described in the section dealing with the state variables. It should be noted that only the attitude indication of the IMU is of interest, since its accelerometers, if any, give practically no output in the absence of thrust, the case under consideration.

Horizon Tracker

A horizon tracker may be pictured, for the sake of error analysis, as four independent sighting telescopes, two pointing to the forward and backward horizons in the orbital plane, and two pointing to the left and right horizons in a plane normal to the orbit (see Figure 4). The tracker measures the four angles Φ_1, Φ_2, Φ_3, Φ_4 of the telescopes relative to the $\hat{\xi}\hat{\eta}\hat{\zeta}$ coordinate frame at the vehicle. These angles are combined to give an average stadimetric angle

$$\bar{\Phi}_1 = \tfrac{1}{2}(\Phi_1 - \Phi_2 + \Phi_4 - \Phi_3) \tag{2.13}$$

and two angles, $\bar{\Phi}_2$ and $\bar{\Phi}_3$, that indicate the direction of the local vertical relative to the $\hat{\xi}\hat{\eta}\hat{\zeta}$ coordinate frame:

$$\bar{\Phi}_2 = \tfrac{1}{2}(\Phi_1 + \Phi_2) \tag{2.14}$$

$$\bar{\Phi}_3 = \tfrac{1}{2}(\Phi_3 + \Phi_4) \ . \tag{2.15}$$

One of the major errors in these observations is in the deviations h_1, h_2, h_3, h_4 of the sensible horizon heights from their assumed values H_1, H_2, H_3, H_4. Corresponding to the errors $\delta\bar{\Phi}_1$, $\delta\bar{\Phi}_2$, $\delta\bar{\Phi}_3$ in the observations of Equations (2.13), (2.14) and (2.15), there are three combined horizon deviations,

$$\bar{h}_1 = \tfrac{1}{2}(h_1 + h_2 + h_3 + h_4) = \sqrt{(R^2 - R_{EH}^2)}\ \delta\bar{\Phi}_1 \tag{2.16}$$

$$\bar{h}_2 = \tfrac{1}{2}(h_1 - h_2) = \sqrt{(R^2 - R_{EH}^2)}\ \delta\bar{\Phi}_2 \tag{2.17}$$

$$\bar{h}_3 = \tfrac{1}{2}(h_4 - h_3) = \sqrt{(R^2 - R_{EH}^2)}\ \delta\bar{\Phi}_3 \ , \tag{2.18}$$

where R is the distance from the center of the earth to the vehicle and R_{EH} is the distance from the center of the earth to the assumed horizon altitude. It is possible to take these combined horizon deviations as three independent state variables. A convenient model for them, consisting of time correlated zero-mean noises, will be discussed in the next section. In addition, it is often necessary to consider the bias errors in the $\bar{\Phi}_1$, $\bar{\Phi}_2$, $\bar{\Phi}_3$ observations produced by a constant error in the estimated horizon heights. Such an error will produce a bias in $\bar{\Phi}_1$, but not in $\bar{\Phi}_2$ and $\bar{\Phi}_3$. Hence it is often realistic to introduce a state variable $\Delta\bar{H}_1$ for horizon bias affecting the stadimetric angle $\bar{\Phi}_1$.

A second major error in the combined observations is in the tracking instrument itself. Instrument errors appear in the angles $\Phi_1, \Phi_2, \Phi_3, \Phi_4$ and hence in $\bar{\Phi}_1$, $\bar{\Phi}_2$, $\bar{\Phi}_3$. To describe them one may postulate combined instrument errors $\bar{\phi}_1$, $\bar{\phi}_2$, $\bar{\phi}_3$. These errors may be introduced as time-correlated variables in the state vector, if a suitable model is found. A rather complete model would include biases $\Delta\bar{\Phi}_1$, $\Delta\bar{\Phi}_2$ and $\Delta\bar{\Phi}_3$ and purely white-noise components. In general, the time constant of the time correlated error is small enough to allow this error to be considered as white-noise.

A third error that appears in the horizon tracker observations is that due to the IMU error angles θ_1, θ_2, θ_3. Here the two cases of IMU orientation must be distinguished. Moreover, errors in IMU orientation do not affect the stadimetric angle $\bar{\Phi}_1$, but do add directly to the observations $\bar{\Phi}_2$ and $\bar{\Phi}_3$ of the vertical.

In summary, Table 1 lists the matrices relating the effective horizon tracker observations $\bar{\Phi}_1$, $\bar{\Phi}_2$, $\bar{\Phi}_3$ to the following variables in the state vector: vehicle position errors x, y, z; effective horizon deviations $\bar{h}_1$, $\bar{h}_2$, $\bar{h}_3$; IMU error angles θ_1, θ_2, θ_3; effective tracker errors $\bar{\phi}_1$, $\bar{\phi}_2$, $\bar{\phi}_3$; effective horizon tracker biases $\Delta\Phi_1$, $\Delta\Phi_2$, $\Delta\Phi_3$; horizon deviation bias ΔH_1.

Star Tracker

A star tracker provides highly accurate attitude information to the IMU axes and, thereby, to other instruments. The observations are train angle (TR) and elevation angle (EL), relative to the $\hat{\xi}\hat{\eta}\hat{\zeta}$ reference frame, of the unit vector from the vehicle towards the star. Figure 5 shows these observations. Note that here TR and EL are measured relative to the vehicle vertical coordinates, even if the IMU axes are along XYZ. For simplicity, the same geometry is assumed for star and landmark observations.

The errors in the TR and EL observations are those in the instrument itself and in the IMU orientation, because star directions are known with high precision.

The instrument errors are divided, as usual, into two correlated errors Ω_1, Ω_2, bias errors ΔTR, ΔEL, and pure white-noise. The total instrument error is often of the order of arc-seconds.

The IMU orientation errors θ_1, θ_2, θ_3 usually dominate the total error if more than a few minutes have elapsed since the last star observations.

Landmark Sensor

Observation of a landmark of known location on the earth may be made with a telescope by a crew member.

The geometry of a landmark observation is assumed to be similar to that of a star observation (see Figure 5). But, since the landmark is not at infinity, as is a star, the observation is affected by, and hence yields information about, the vehicle's position as well as its attitude. The errors involved may be modelled as time correlated errors π_1 , π_2 , bias errors ΔTR , ΔEL , and pure white-noise, just as in the case of the star tracker.

Radar Altimeter

This instrument measures RA , the distance $|\bar{R}-\bar{R}_{ss}|$ between the vehicle and the sub-satellite point $\bar{R}_{ss}$ (see Figure 6). It involves only the position variables of the state, as shown in Table I.

The major errors in the observation are in knowledge of the altitude of the sub-satellite terrain, in tidal heights, and in the instrument itself. The characterization of the terrain uncertainty as a time-correlat first-order Markov process is valid only as a first approximation, since this error is highly orbit dependent.

The radar altimeter information is restricted to the radial direction, as may be seen in Table I. Hence cross-track and down-track errors cannot be removed by this observation.

2.4 State Vector; the ϕ and Q Matrices

2.4.1 *Partitioning for Biases*

The variables introduced in Section 2.3 from examination of the observations will be included in the state. Table II lists those state variables that are not biases; Table III lists the bias variables. The total number of variables is seen to be close to 80, if only one ground radar is used; if a computer program were constructed to calculate the basic covariance equations, Equations (2.1)′ and (2.2) as they stand, approximately 2500 computer locations would be required for each of the P_{i+1} , P_i , ϕ and Q matrices. Moreover, 0.25×10^6 operations (multiplies and adds) are required for a single matrix product such as ϕP . At 10 microseconds per operation, Equation (2.1)′ alone would require about 5 seconds of computer time. When six or more ground stations are employed simultaneously, the number of state variables is about 88, each matrix requires about 7200 locations, and Equation (2.1)′ requires about 25 seconds. It is, then, clearly impossible to simulate a one or two day mission, in which thousands of observations are made, without consuming large amounts of computer time and space.

It is possible to alleviate considerably the computer storage and time problem by (1) processing only one (scalar) observation at a time and (2) partitioning Equations (2.1)′ and (2.2) into sub-matrices for the bias variables.

(1) If the noise covariance R in Equation (2.2) is diagonal, it is possible to process one row of the H matrix at a time. (As pointed out in Reference 4, it is always possible to transform Equation (2.2) so that R is diagonal.) This avoids matrix inversion and saves on storage space for PH^T . Equation (2.2) is then

$$P^+ = P^- - (P^-H_k^T)(P^-H_k^T)^T/(H_kP^-H_k^T+R_k) \ , \tag{2.19}$$

where R_k is the k^{th} diagonal element of R and H_k is the k^{th} row of H . Equation (2.1)′ is still

$$P_i^- = \phi(t_i,t_{i-1})P_{i-1}^+\phi^T(t_i,t_{i-1}) + Q(t_i,t_{i-1}) \ . \tag{2.20}$$

(2) Pure bias states have identity elements in the state transition matrix and [0] in the Q matrix. Partitioning Equation (2.20) gives

$$\phi = \begin{bmatrix}\phi_s & 0\\ 0 & I\end{bmatrix} \qquad P = \begin{bmatrix}P_s & P_{sB}\\ P_{Bs} & P_B\end{bmatrix} \qquad Q = \begin{bmatrix}Q_s & 0\\ 0 & 0\end{bmatrix} \ . \tag{2.21}$$

Here B stands for bias, s for something else.

The equations then become

$$\begin{bmatrix}P_s(t^-) & P_{sB}(t^-)\\ P_{sB}^T(t^-) & P_B(t^-)\end{bmatrix} = \begin{bmatrix}\phi_sP_s(t_0)\phi_s^T+Q_s & \phi_sP_{sB}(t_0)\\ P_{sB}^T(t_0)\phi_s^T & P_B(t_0)\end{bmatrix} \tag{2.22}$$

and

$$\begin{bmatrix}P_s(t^+) & P_{sB}(t^+)\\ P_{sB}^T(t^+) & P_B(t^+)\end{bmatrix} = \begin{bmatrix}P_s(t^-) & P_{sB}(t^-)\\ P_{sB}^T(t^-) & P_B(t^-)\end{bmatrix} - \frac{1}{D}\begin{bmatrix}\alpha^T\alpha & \alpha^T\beta\\ \beta^T\alpha & \beta^T\beta\end{bmatrix} \ , \tag{2.23}$$

where

$$\alpha = H_s(t)P_s(t^-) + H_B(t)P_{sB}^T(t^-) \tag{2.24}$$

$$\beta = H_s(t)P_{sB}(t^-) + H_B(t)P_B(t^-) \tag{2.25}$$

$$D = H_s(t)P_s(t^-)H_s^T(t) + R(t)P_B(t^-)B^T(t) + 2H_s(t)P_{sB}(t^-)H_s^T(t) + R_k \ . \tag{2.26}$$

The first of these two equations may be summarized as

$$\boxed{\begin{aligned} P_s(t^-) &= \phi_s(t,t_0)P_s(t_0^+)\phi_s^T(t,t_0) + Q_s(t,t_0) \qquad (2.27)\\ P_{sB}(t^-) &= \phi_s(t,t_0)P_{sB}(t_0^+) \qquad (2.28)\\ P_B(t^-) &= P_B(t_0^+) \ . \qquad (2.29) \end{aligned}} \quad \text{TIME PROPAGATION}$$

The second may be summarized as

$$\boxed{\begin{aligned} P_s(t^+) &= P_s(t^-) - \alpha^T\alpha/D \qquad (2.30)\\ P_{sB}(t^+) &= P_{sB}(t^-) - \alpha^T\beta/D \qquad (2.31)\\ P_B(t^+) &= P_B(t^-) - \beta^T\beta/D \ , \qquad (2.32) \end{aligned}} \quad \text{DATA INCORPORATION}$$

where

$$\alpha = H_s(t)P_s(t^-) + H_B(t)P_{sB}^T(t^-) \tag{2.33}$$

$$\beta = H_s(t)P_{sB}(t^-) + H_B(t)P_B(t^-) \tag{2.34}$$

$$D = \alpha H_s^T(t) + \beta H_B^T(t) + R_k \ . \tag{2.35}$$

The above time and data equations are forms that may be programmed. The former are used to extrapolate the error matrices P_s, P_{sB}, P_B from t_0 to t in the absence of measurements; the latter are used to re-evaluate P_s, P_{sB}, P_B at time t, once for each measurement. It will be noticed that α and β are row matrices, $\alpha^T\alpha$, $\beta^T\beta$ are square matrices, $\alpha^T\beta$ is a rectangular matrix, and D is a scalar.

2.4.2 *Equations for* ϕ_s *and* Q_s

The state transition matrix ϕ_s for the error vector of Table II is obtained from the deterministic description of the error process, in the integrated form

$$X(t) = \phi_s(t,t_0)X(t_0) + \int_{t_0}^{t}\phi_s(t,\sigma)N(\sigma)\,d\sigma \ , \tag{2.36}$$

where X is the error, ϕ_s its state transition matrix and $N(\sigma)$ the driving noise.

The Q_s matrix is obtained as the covariance of the integral on the right side of the Equation (2.36):

$$Q_s = \int_{t_0}^{t} d\lambda \int_{t_0}^{t} d\mu\, \phi_s(t,\lambda)E[N(\lambda)N^T(\mu)]\phi_s^T(t,\mu) \ , \tag{2.37}$$

where $E[\ ]$ indicates expectation.

In practice, the ϕ_s and Q_s matrices are themselves partitioned into sub-matrices appropriate to the natural groupings of variables in the state. For the errors shown in Table II, the partitioning of Figures 7 and 8 are appropriate for ϕ_s and Q_s, respectively. Note that the dynamic biases $\delta\mu$, $\delta C_N(1)$, ..., $\delta C_N(4)$, $\delta D_0/D_0$ have been retained in ϕ_s because they affect position and velocity via ϕ_0. The various non-zero sub-matrices of Figures 7 and 8 will now be discussed.

Vehicle Position and Velocity

These are the error quantities of primary interest in most applications. A closed form for the two-body state transition matrix ϕ_{RV} is given by Koehl[5,6]. A realistic picture of error propagation, however, includes

the effects of uncertainties in the earth's gravitational parameter μ, in the zonal harmonics of the geopotential C_i, $i = 2,3,4, \ldots$, and, for low altitude satellites, the uncertainties in air drag. A method for treating these three classes of uncertainties as biases (constant errors) is given in Appendix B. A more sophisticated treatment for air drag is warranted in many cases[7-11] and extensive models have been developed for the errors in the geopotential[12-14]. The equations of Appendix B, however, are adequate to predict the correct order of magnitude of drag and potential errors for earth orbits from about 50 to 500 miles altitude. In particular, the comparatively large errors produced by the uncertainty in C_2 can be accurately predicted by these equations.

The state transition sub-matrix Φ_D may be calculated by numerical integration and filled into Φ_s as shown in Figure 7. The portion of Q_s corresponding to position and velocity is zero, as shown in Figure 8, if drag is treated as a bias in D_0. This portion of Q_s would be non-zero if a stochastic drag model were used.

Gyro Errors

The model of gyro drift is taken as a fixed drift rate plus drift rate noise:

$$\theta(t) = \theta_0 + \omega_0\tau_{GD}(1-e^{-|(t-t_0)/\tau_{GD}|}) + \omega_B(t-t_0) + \int_{t_0}^{t}\tau_{GD}(1-e^{-|(t-\sigma)/\tau_{GD}|})\eta_\omega(\sigma)\,d\sigma\,, \qquad (2.38)$$

where

$\theta(t)$ = total angular error about input axis of gyro

ω_0 = initial gyro drift rate

τ_{GD} = gyro drift noise time constant

η_ω = uncorrelated drift noise

ω_B = gyro drift rate bias.

The gyro drift rate process itself is taken to be white-noise through an RC filter:

$$\omega(t) = \omega(t_0)e^{-|(t-t_0)/\tau_{GD}|} + \int_{t_0}^{t} e^{-|(t-\sigma)/\tau_{GD}|}\eta_\omega(\sigma)\,d\sigma\,. \qquad (2.39)$$

The gyro drifts about each of the three orthogonal IMU axes are assumed to be mutually independent and each follows the model of Equations (2.38) and (2.39). A more refined model involving time-dependent drift rate would be needed for long periods of gyro operation without reset; such a case would occur if reentry or maneuver had to be made long after star tracker failure. The present model, however, is adequate for non-failure mode operation.

When the covariances of $\theta(t)$ and $\omega(t)$ are taken for the three gyros, the combined state transition matrix for gyro angles, drift rate noise and drift rate biases is

$$\Phi_{GD} = \left[\begin{array}{ccc|ccc|ccc}
1 & 0 & 0 & \tau_{GD}(1-X) & 0 & 0 & t-t_0 & 0 & 0 \\
0 & 1 & 0 & 0 & \tau_{GD}(1-X) & 0 & 0 & t-t_0 & 0 \\
0 & 0 & 1 & 0 & 0 & \tau_{GD}(1-X) & 0 & 0 & t-t_0 \\
\hline
0 & 0 & 0 & X & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & X & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & X & 0 & 0 & 0 \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right], \qquad (2.40)$$

where $X = e^{-|(t-t_0)/\tau_{GD}|}$.

The covariance matrix of the accumulated noise is

$$Q_{GD} = \left[\begin{array}{ccc|ccc|ccc} A & 0 & 0 & B & 0 & 0 & 0 & 0 & 0 \\ 0 & A & 0 & 0 & B & 0 & 0 & 0 & 0 \\ 0 & 0 & A & 0 & 0 & B & 0 & 0 & 0 \\ \hline B & 0 & 0 & C & 0 & 0 & 0 & 0 & 0 \\ 0 & B & 0 & 0 & C & 0 & 0 & 0 & 0 \\ 0 & 0 & B & 0 & 0 & C & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] , \qquad (2.41)$$

where

$$A = \sigma_{GD}^2 \tau_{GD}[2(t-t_0) + \tau_{GD}(1-X^2) - 4\tau_{GD}(1-X)]$$

$$B = \sigma_{GD}^2 \tau_{GD}(1-X)^2$$

$$C = \sigma_{GD}^2(1-X^2) ,$$

and

σ_{GD}^2 = variance of gyro drift noise

Horizon Deviations

The combined horizon deviations $\tilde{h}_1$, $\tilde{h}_2$, $\tilde{h}_3$ are dependent on the actual horizon deviations h_1 , h_2 , h_3 , h_4 at the points of tangency of the telescopic lines of sight and the sensed horizon. For a reasonable approximation, one may assume[17] that h_1 , h_2 , h_3 , h_4 are correlated in time and location with an autocorrelation function

$$CF(\tau,\delta) = \sigma_{HD}^2 \, e^{-|\tau/\tau_H| - |\delta/\delta_H|} , \qquad (2.42)$$

where τ is the time between measurements of deviations h in height of horizon points separated by a distance δ , τ_H and δ_H are constants determined by observation. Here σ_{HD}^2 is the variance of h_1 , h_2 , h_3 , or h_4 . This approximation is applicable to the CO_2 14-16 micron spectral band. Using this autocorrelation function, and taking into account the vehicle velocity but ignoring the correlations due to the adjacency of successive orbits, Wilcox[18] has calculated the variance of the position error along the vertical axis ζ as

$$\frac{R^2}{R_E^2} \frac{(1+2a+b)}{4} \sigma_{HD}^2 , \qquad (2.43)$$

and the variance of the position errors along the horizontal axes η and ξ as

$$\frac{R^2(1-b)}{(R^2-R_E^2)^{\frac{1}{2}}} \sigma_{HD}^2 , \qquad (2.44)$$

where

$$a = e^{-(1/\delta_H)\cos^{-1}(R_E/R)^{\frac{1}{2}}}$$

$$b = e^{-(\delta/\delta_H)\cos^{-1}(R_E/R)}$$

R_E = radius of earth

σ_{HD}^2 = variance of h_1 , h_2 , h_3 or h_4 .

In terms of the errors $\delta\bar{\Phi}_1$, $\delta\bar{\Phi}_2$, $\delta\bar{\Phi}_3$ in the combined tracker observations, the vertical position error is

$$-\frac{R}{R_E}(R^2-R_E^2)^{\frac{1}{2}}\delta\bar{\Phi}_1 \qquad (2.45)$$

and the horizontal position errors are

$$R\delta\Phi_s \quad . \tag{2.46}$$

Combining Equations (2.45), (2.43) and (2.16) gives the variance of $\tilde{h}_1$:

$$\text{var}(\tilde{h}_1) = \sigma^2_{HD1} = \frac{1 + 2a + b}{4}\sigma^2_{HD} \quad . \tag{2.47}$$

and the variances of $\tilde{h}_2$, $\tilde{h}_3$:

$$\text{var}(\tilde{h}_2) = \sigma^2_{HD2} = \frac{(1-b)}{2}\sigma^2_{HD} \tag{2.48}$$

$$\text{var}(\tilde{h}_3) = \sigma^2_{HD3} = \frac{(1-b)}{2}\sigma^2_{HD} \quad . \tag{2.49}$$

The three combined horizon deviations may thus be considered as the outputs of three first-order linear systems driven by three independent white-noise sources. The model for each is

$$H(t) = H(t_0)e^{-|(t-t_0)/\tau_H|} + \int_{t_0}^{t} e^{-|(t-\sigma)/\tau_H|}\eta(\sigma)\, d\sigma \quad . \tag{2.50}$$

The ϕ and Q for the three deviations are .

$$\phi_{HD} = \begin{bmatrix} e^{-|(t-t_0)/\tau_{HD}|} & 0 & 0 \\ 0 & e^{-|(t-t_0)/\tau_{HD}|} & 0 \\ 0 & 0 & e^{-|(t-t_0)/\tau_{HD}|} \end{bmatrix} \quad , \tag{2.51}$$

$$Q_{HD} = \begin{bmatrix} \sigma^2_{HD1}(1-y^2) & 0 & 0 \\ 0 & \sigma^2_{HD2}(1-y^2) & 0 \\ 0 & 0 & \sigma^2_{HD3}(1-y^2) \end{bmatrix} \quad , \tag{2.52}$$

where

$$y = e^{-|(t-t_0)/\tau_{HD}|} \tag{2.53}$$

and, as shown in Reference 18,

$$\tau_{HD} = \tau_H \delta_H/(\delta_H + 2\pi\tau_H/T_0) \quad , \tag{2.54}$$

where T_0 is the orbital period of the satellite.

Horizon Tracker

The combined instrument errors $\tilde{\phi}_1$, $\tilde{\phi}_2$, $\tilde{\phi}_3$, are illustrated in Figure 4. The deterministic model for each is that of white-noise through an RC filter, the three noise components being assumed uncorrelated with each other. The model for each error is

$$\tilde{\phi}(t) = \tilde{\phi}(t_0)e^{-|(t-t_0)/\tau_{HT}|} + \int_{t_0}^{t} e^{-|(t-\sigma)/\tau_{HT}|}\eta\tilde{\phi}(\sigma)\, d\sigma \tag{2.55}$$

$$\phi_{HT} = \begin{bmatrix} e^{-|(t-t_0)/\tau_{HT}|} & 0 & 0 \\ 0 & e^{-|(t-t_0)/\tau_{HT}|} & 0 \\ 0 & 0 & e^{-|(t-t_0)/\tau_{HT}|} \end{bmatrix} \tag{2.56}$$

$$Q_{HT} = \begin{bmatrix} \sigma^2_{HT1}(1-e^{-2|(t-t_0)/\tau_{HT}|}) & 0 & 0 \\ 0 & \sigma^2_{HT2}(1-e^{-2|(t-t_0)/\tau_{HT}|}) & 0 \\ 0 & 0 & \sigma^2_{HT3}(1-e^{-2|(t-t_0)/\tau_{HT}|}) \end{bmatrix} \quad , \tag{2.57}$$

where

σ^2_{HT1} = variance of $\tilde{\phi}_1$

σ^2_{HT2} = variance of $\tilde{\phi}_2$

σ^2_{HT3} = variance of $\tilde{\phi}_3$

τ_{HT} = time constant for the horizon tracker errors.

If a four-telescope tracker is employed, as described in Section 2.3,

$$\sigma^2_{HT1} = \sigma^2_{HT}/4 \tag{2.58}$$

$$\sigma^2_{HT2} = \sigma^2_{HT3} = \sigma^2_{HT}/2 , \tag{2.59}$$

where σ^2_{HT} is the actual variance of a single telescope angle.

Star Tracker

The time correlated errors in the star tracker train and elevation angle are shown in Figure 5. The error process for each is assumed to be white-noise through an RC filter, the two components being uncorrelated with each other:

$$\Omega(t) = \Omega(t_0)e^{-|(t-t_0)/\tau_{ST}|} + \int_{t_0}^{t} e^{-|(t-\sigma)/\tau_{ST}|}\, \eta_{ST}(\sigma)\, d\sigma . \tag{2.60}$$

The state transition matrix and noise matrix are

$$\Phi_{ST} = \begin{bmatrix} e^{-|(t-t_0)/\tau_{ST}|} & 0 \\ 0 & e^{-|(t-t_0)/\tau_{ST}|} \end{bmatrix} \tag{2.61}$$

$$Q_{ST} = \begin{bmatrix} \sigma^2_{ST1}(1-e^{-2|(t-t_0)/\tau_{ST}|}) & 0 \\ 0 & \sigma^2_{ST2}(1-e^{-2|(t-t_0)/\tau_{ST}|}) \end{bmatrix} \tag{2.62}$$

where

σ^2_{ST1} = variance of star tracker train angle error

σ^2_{ST2} = variance of star tracker elevation angle error

τ_{ST} = correlation time of star tracker errors.

Landmark Sensor

The time correlated noises in the landmark sensor train angle and elevation angle (Fig. 5), are state variables. The error process for each is assumed to be white noise through an RC filter, the two components being uncorrelated with each other:

$$\pi(t) = \pi(t_0)e^{-|(t-t_0)/\tau_{LS}|} + \int_{t_0}^{t} e^{-|(t-\sigma)/\tau_{LS}|}\, \eta_{LS}(\sigma)\, d\sigma . \tag{2.63}$$

The state transition matrix and noise matrix are

$$\Phi_{LS} = \begin{bmatrix} e^{-|(t-t_0)/\tau_{LS}|} & 0 \\ 0 & e^{-|(t-t_0)/\tau_{LS}|} \end{bmatrix} \tag{2.64}$$

$$Q_{LS} = \begin{bmatrix} \sigma^2_{LS1}(1-e^{-2|(t-t_0)/\tau_{LS}|}) & 0 \\ 0 & \sigma^2_{LS2}(1-e^{-2|(t-t_0)/\tau_{LS}|}) \end{bmatrix} \tag{2.65}$$

where

$\sigma^2_{Ls_1}$ = variance of landmark sensor train angle

$\sigma^2_{Ls_2}$ = variance of landmark sensor elevation angle

τ_{Ls} = correlation time of landmark sensor errors.

The geometry chosen for landmark sensor and star tracker is not ideal from a mechanization point of view, since it requires infinite train rates when the star or landmark is along the vertical. Nevertheless, it was chosen because it is simpler to compute and because it should give results representative of the several practical configurations possible.

Gravity, Oblateness and Drag Biases

The state transition matrix relating fixed uncertainties in the gravity, oblateness and drag parameters to position and velocity is given by ϕ_0, described in Appendix B. This matrix is obtained most conveniently by numerical integration of the differential equations. The state transition matrix relating the biases at time t_0 to the same biases at time t is just the identity matrix I. The Q_s matrix for these biases is zero.

2.5 Interpreting the Covariance Matrix

2.5.1 *Vehicle Position Error Covariances*

The vehicle position error covariances contained in P_s are in inertial coordinates. They may be converted to local vertical coordinates. If T indicates transpose,

$$\text{cov}(\bar{X}_{LV}\bar{X}_{LV}^T) = \text{cov}\{(S^T\bar{X}_I)\cdot(S^T\bar{X}_I)^T\} = S^T\text{cov}(\bar{X}_I\bar{X}_I^T)S$$

$$= S^T P_{s(1-3)} S\,, \tag{2.66}$$

where $\bar{X}_{LV}$ is the position error vector in local vertical coordinates, $\bar{X}_I$ the same vector in inertial coordinates, $P_{s(1-3)}$ is the upper left 3×3 of P_s, and S is an orthogonal transformation from local vertical to inertial coordinates:

$$S_{11} = \cos\Psi\cos\Omega - \sin\Psi\sin\Omega\cos i = \cos\beta \tag{2.67}$$

$$S_{12} = -\sin\Psi\cos\Omega - \cos\Psi\sin\Omega\cos i = -\sin\beta \tag{2.68}$$

$$S_{13} = \sin\Omega\sin i = 0 \tag{2.69}$$

$$S_{21} = \cos\Psi\sin\Omega + \sin\Psi\cos\Omega\cos i = \sin\beta \tag{2.70}$$

$$S_{22} = -\sin\Psi\sin\Omega + \cos\Psi\cos\Omega\cos i = \cos\beta \tag{2.71}$$

$$S_{23} = -\sin i\cos\Omega = 0 \tag{2.72}$$

$$S_{31} = \sin\Psi\sin i = 0 \tag{2.73}$$

$$S_{32} = \cos\Psi\sin i = 0 \tag{2.74}$$

$$S_{33} = \cos i = 1\,, \tag{2.75}$$

where Ψ and i, given in Figure 9(a) are used if the inclination is not zero, and β, given in Figure 9(b), is used if the inclination is zero.

2.5.2 *Vehicle Velocity Error Covariances*

The velocity error covariances contained in P_s may be transformed from inertial to local vertical coordinates. The computation is

$$\text{cov}(\bar{V}_{LV}\bar{V}_{LV}^T) = \text{cov}\{(S^T\bar{V}_I)(SV_I)^T\} = S^T\,\text{cov}(\bar{V}_I\bar{V}_I^T)S$$

$$= S^T P_{s(4-6)} S\,, \tag{2.76}$$

where $\bar{V}_{LV}$ is the velocity error in local vertical coordinates, $\bar{V}_I$ is the velocity error in inertial coordinates and $P_{s(4-6)}$ is $P_{s(I,J)}$ for $I = 4,5,6$ and $J = 4,5,6$. The matrix S is that given above for position covariances.

2.5.3 *Gyro Misalignment Covariances*

The elements of $P_{s(I,J)}$ for $I,J = 7,8,9$ are the covariances of the gyro misalignments along platform axes 1,2,3. If the platform axes coincide with the local vertical no transformation is needed. If the platform axes coincide with the inertial coordinate frame a transformation to local vertical axes is needed. The calculation is as follows:

$$\begin{aligned} \mathrm{cov}(\vec{\theta}_{LV}\vec{\theta}_{LV}^{T}) &= \mathrm{cov}((S^{T}\vec{\theta}_{I})(S^{T}\vec{\theta}_{I})^{T}) \\ &= S^{T}P_{s(7,8,9)}S \ , \end{aligned} \tag{2.77}$$

where S is as above, $P_{s(7,8,9)}$ is $P_{s(I,J)}$ for $I,J = 7,8,9$ and θ is the vector gyro misalignment.

2.5.4 *Vehicle Attitude Covariances*

The vehicle attitude error $\vec{\gamma}$ is the sum of gyro misalignment θ and the rotation error ϵ due to position and velocity errors:

$$\vec{\gamma} = \vec{\theta} + \vec{\epsilon} \tag{2.78}$$

$$\simeq \theta + \left(\frac{\vec{R}\times\delta\vec{R}}{R^{2}} + \frac{\hat{\xi}\cdot\delta\vec{V}}{\eta\cdot V}\hat{\xi}\right) , \tag{2.79}$$

where $\hat{\xi}\hat{\eta}\hat{\zeta}$ are the local vertical unit vectors of Figure 2 and $\delta\vec{R}$, $\delta\vec{V}$ are position and velocity error. Factoring out δR and δV gives

$$\begin{aligned} \vec{\gamma} &= \vec{\theta} + \left[[\vec{R}\times]\frac{1}{R^{2}} + \frac{\vec{R}\times\vec{V}}{(\vec{R}\times\vec{V})^{2}}\vec{R}\right]\cdot\begin{Bmatrix}\delta\vec{R}\\ \delta\vec{V}\end{Bmatrix} \\ &= \vec{\theta} + [\theta_{A}]\cdot\vec{X} \ , \qquad \text{where } \vec{X} = \begin{Bmatrix}\delta\vec{R}\\ \delta\vec{V}\end{Bmatrix}_{(\text{INERTIAL})} \end{aligned} \tag{2.80}$$

and $[\theta_{A}]$ is known in terms of $\vec{R}$ and $\vec{V}$, in inertial coordinates. If $\vec{\theta}$ is in a local vertical system, the attitude covariance, in local vertical coordinates, is

$$\begin{aligned} \mathrm{cov}(\vec{\gamma}_{LV}\vec{\gamma}_{LV}) &= \mathrm{cov}\{(S^{T}\vec{\theta}_{I} + S^{T}[\theta_{A}]\vec{X})(S^{T}\vec{\theta}_{I} + S^{T}[\theta_{A}]\vec{X})^{T}\} \\ &= \mathrm{cov}\{S^{T}(\vec{\theta}_{I} + [\theta_{A}]\vec{X})(\vec{\theta}_{I}^{T} + \vec{X}^{T}[\theta_{A}]^{T})S\} \\ &= S^{T}\,\mathrm{cov}\{\vec{\theta}_{I}\vec{\theta}_{I} + [\theta_{A}]\vec{X}\vec{\theta}_{I}^{T} + \vec{\theta}_{I}\vec{X}^{T}[\theta_{A}]^{T} + [\theta_{A}]\vec{X}\vec{X}^{T}[\theta_{A}]^{T}\}S \\ &= S^{T}\left[P_{s(\theta\theta)} + [\theta_{A}]P_{s(x\theta)} + P_{s(\theta x)}[\theta_{A}]^{T}\right] + \left[\theta_{A}P_{s(xx)}[\theta_{A}]^{T}\right]S \ , \end{aligned} \tag{2.81}$$

where

$$\begin{aligned} P_{s(\gamma\gamma)} &= P_{s(I,J)} \quad \text{for} \quad I = 7\text{-}9 \ , \quad J = 7\text{-}9 \\ P_{s(x\gamma)} &= P_{s(I,J)} \quad \text{for} \quad I = 1\text{-}6 \ , \quad J = 7\text{-}9 \\ P_{s(\gamma x)} &= P_{s(I,J)} \quad \text{for} \quad I = 7\text{-}9 \ , \quad J = 1\text{-}6 \\ P_{s(xx)} &= P_{s(I,J)} \quad \text{for} \quad I = 1\text{-}6 \ , \quad J = 1\text{-}6 \ . \end{aligned}$$

If θ is in an inertial frame, the attitude covariance in local vertical coordinates is

$$\begin{aligned} \mathrm{cov}(\vec{\gamma}_{LV}\vec{\gamma}_{LV}^{T}) &= \mathrm{cov}\{(\vec{\theta}_{LV} + S^{T}[\theta_{A}]\vec{X})(\vec{\theta}_{LV} + S^{T}[\theta_{A}]\vec{X})^{T}\} \\ &= \mathrm{cov}(\vec{\theta}_{LV}\vec{\theta}_{LV}^{T}) + S^{T}[\theta_{A}]\ \mathrm{cov}(\vec{X}\vec{\theta}_{LV}^{T}) + \mathrm{cov}(\vec{\theta}_{LV}\vec{X}^{T})[\theta_{A}]^{T}S + S^{T}[\theta_{A}]\mathrm{cov}(\vec{X}\vec{X}^{T})[\theta_{A}]^{T}S \\ &= P_{s(\gamma\gamma)} + S^{T}[\theta_{A}]P_{s(x\gamma)} + P_{s(\gamma x)}[\theta_{A}]^{T}S + S^{T}[\theta_{A}]P_{s(xx)}[\theta_{A}]^{T}S \ . \end{aligned} \tag{2.82}$$

The remainder of the covariance matrix is readily interpreted and needs no transformation.

2.6 Determining the Times of the Observations

The logic and search procedures that must be programmed to simulate realistically the times at which observations are made may constitute almost half of the computer program coding.

2.6.1 *Ground Based Radar*

Most radar observations taken with the antenna at a low elevation angle are unreliable and are discarded. As a result any one radar can take readings only when the vehicle is within a cone extending out from the radar station and with its axis along the local vertical of the station. Typically, the half-cone angle is between 60° and 85°. The maximum range of the radar puts a spherical cap on this cone of visibility. A search procedure must determine, for each radar in the tracking net, whether the vehicle falls within this cone of visibility. The rotation of the earth complicates the picture in a minor way. A variable search step is found to be indispensible, in many cases, to avoid excessive computation time.

2.6.2 *Horizon Tracker*

This instrument is limited in operation only by the power and data processing capacity of the vehicle.

2.6.3 *Star Tracker*

The major considerations in determining the times of observation for a star tracker are (a) stellar magnitude, (b) obscuration of the star by the earth, and (c) adequate angular separation of stars.

2.6.4 *Landmark Sensor*

The considerations here are very similar to those for the ground-based radar observations. Indeed, the same search procedure (coding) may be used for landmarks as for radars, with suitably different parameters.

2.6.5 *Radar Altimeter*

It is usually desired to investigate the effect of operating the altimeter only over water, since the tidal uncertainty is much less than the terrain uncertainty and easier to determine. Storing a facsimile of the land-sea configuration may be avoided, for error analysis purposes, by allowing altimetric readings for a fraction of the time corresponding to the percentage of water underlying the orbit.

2.7 Results of Typical Computer Runs

The results of a typical computer program for on-board measurements alone are recorded and discussed in Reference 18. The instruments employed are horizon tracker, star tracker, radar altimeter. The computer runs indicate, among other things, the oscillatory nature of cross-track error, the value of a radar altimeter in restricting the accumulation of down-track errors, and the futility of increasing the horizon tracker data rate beyond the limit set by the horizon deviation time constant.

The results of a typical computer program that combines both on-board and ground-based data are shown in Figure 10. This computer run assumed three radar stations tracking a satellite in a 250 nautical miles circular orbit. Ground radar range readings were taken at 10, 22 and 24 hours. They have a bias of 35 ft, white noise of 10 ft, station location biases of $\Delta e = \Delta n = 75$ ft , and $\Delta h = 25$ ft . The on-board horizon tracker was assumed to have a bias of 0.015 deg and a noise variance of 0.25 deg^2 with a 1 second time constant. The horizon deviations were assumed to have $R_E\delta_H = 2500$ nautical miles, and a variance $\sigma^2_{HD} = 1.4(\text{nautical miles})^2$. A star tracker with 0.16 deg bias, 1 second time constant and $4 \times 10^{-6}(deg)^2$ variance was assumed, along with a radar altimeter of 0.01 nautical mile white-noise and 0.005 nautical mile bias. These results are not realistic beyond about one day of orbital time, because no drag or uncertainties in the gravity model were included.

When drag bias $\delta D_0/D_0$ and gravitational bias $\delta\mu/\mu$, but not biases in C_n , are accounted for, an error analysis of ground radar tracking yields results such as shown in Figure 11. The radar location biases of 200 ft in each of Δe , Δn , Δh were estimated during the run and at 13.2 hours were, on the average, 141 ft. Radar range readings were taken at times indicated by a solid triangle in Figure 11. The program also attempted to estimate $\delta\mu$, δC_2 , δC_4 , δC_6 and δC_8 , but no significant changes in their covariances were observed because of the brevity of the run.

3. SELF-CONTAINED ALIGNMENT OF AN INERTIAL GUIDANCE SYSTEM

3.1 Description of Problem

As the accuracy and sensitivity of gyroscopes and accelerometers increase, so will the importance of self-contained alignment of inertial navigators. This section describes the application of the Kalman filter to the problem of estimating the orientation of an inertial sensor coordinate frame using only the signals from the gyroscopes and accelerometers. The vehicle in which the instruments are located is undergoing random, but limited, translations and rotations, such as a launch vehicle before lift-off, or an aircraft being loaded preparatory to take-off. The procedure to be described applies both to sensors stabilized by means of gimbals and to sensors mounted directly on the vehicle body. For simplicity it is assumed that one gyro input axis and one accelerometer input axis is along each of three mutually orthogonal directions. These directions define the sensor coordinate frame.

At the time that the fine alignment procedure commences, the sensor coordinate frame will have a small vector misalignment $\vec{\theta}$ relative to some intended inertially-fixed alignment frame (see Figure 12). The vector $\vec{\theta}$ is

a random variable with known mean and covariance at t_0, determined by the errors in some coarse alignment scheme that has been completed before t_0 and that need not concern us. The orientation of the intended alignment frame is accurately known in inertial space; changes in orientation of the sensor frame relative to its actual alignment at t_0 are also accurately known (in the case of body-mounted sensors, because of continuous calculation of the direction cosine matrix, and, in the case of a gimballed system, because of precisely known torques applied to the gyros). The purpose of the Kalman filter then is to estimate the initial misalignment vector $\vec{\theta}$ from the accelerometer readings. The data are the vectors $\vec{Z}(t_i)$ $(i = 1, 2, \ldots N_f)$:

$$\vec{Z}(t_i) = -\vec{g}(t_i) - \vec{g}(t_i) \times \vec{\theta} + \vec{V}(t_i) , \tag{3.1}$$

where $\vec{g}(t_i)$ is the local gravity vector at the alignment site at time $t_i > t_0$ and $\vec{V}(t_i)$ is the sum of vehicle acceleration and accelerometer noise. In order to simplify the filter equations, $\vec{V}(t_i)$ is taken to be zero mean noise. It should be noted that $\vec{g}(t_i)$ is known in any given inertial coordinate system because the geodetic latitude, longitude and altitude of the alignment site are precisely known for all t, $t_0 \leq t \leq t_n$.

3.2 Recursive Filter for the Problem

If one denotes by $\hat{\theta}_i^-$ and $\hat{\theta}_i^+$ the filter's estimate of θ just before and just after the incorporation of the data $\vec{Z}(t_i)$, then the data incorporation equations are

$$\hat{\theta}_i^+ = \hat{\theta}_i^- + K_i[\vec{Z}(t_i) - \hat{Z}(t_i)] \tag{3.2}$$

$$P_i^+ = P_i^- - K_i H_i P_i^- , \tag{3.3}$$

where

$$K_i = P_i^- H_i^T [H_i P_i^- H_i^T + R_i]^{-1} \tag{3.4}$$

$$= \text{Kalman filter matrix at time } t_i ,$$

$$\hat{Z}(t_i) = -\vec{g}(t_i) - \vec{g}(t_i) \times \hat{\theta}_i^- \tag{3.5}$$

$$= \text{estimated observation at } t_i$$

$$H_i = [\vec{g}(t_i) \times] \tag{3.6}$$

$$= \text{antisymmetric matrix corresponding to cross product by } \vec{g}(t_i)$$

$$R_i = \text{covariance of } \vec{V}(t_i)$$

$$= E(\vec{V}(t_i)\vec{V}(t_i)^T) . \tag{3.7}$$

The time propagation equations are particularly simple in inertial space:

$$\hat{\theta}_i^- = \hat{\theta}_{i-1}^+ ; \qquad \hat{\theta}_0^+ = 0 \tag{3.8}$$

$$P_i^- = P_{i-1}^+ ; \qquad P_0^+ = \text{cov}(\hat{\theta}_0^+) . \tag{3.9}$$

The observation matrix H, however, is time-varying if viewed in an inertial frame, because $\vec{g}(t)$ is rotating at the earth's rate:

$$\vec{g}(t_i) = [\theta(t_i, t_0)]\vec{g}(t_0) , \tag{3.10}$$

where $[\theta(t_i, t_0)]$ is a matrix representing the earth's rotation from t_0 to t_i. To be specific, let us express all vectors and matrices in an inertial coordinate system having one axis (u_3) along the earth's spin axis and the other two pointing to points fixed on the celestial equator (Fig. 13). In this coordinate frame, with u_1 directed to the vernal equinox, one has

$$[\theta(t_i, t_0)]_{u_1 u_2 u_3} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.11}$$

$$[g(t_1)\times]_{u_1u_2u_3} = \begin{bmatrix} 0 & -g_3 & g_2 \\ g_3 & 0 & -g_1 \\ -g_2 & g_1 & 0 \end{bmatrix} = H_1 , \tag{3.12}$$

where

$$\theta = (t_1 - t_0)|\vec{\Omega}_E| \tag{3.13}$$

$|\vec{\Omega}_E|$ = magnitude of earth's spin

$$g_1 = -|\vec{g}| \sin\lambda \tag{3.14}$$

$$g_2 = -|\vec{g}| \cos\lambda \cos L^* \tag{3.15}$$

$$g_3 = -|\vec{g}| \cos\lambda \sin L^* \tag{3.16}$$

$|\vec{g}|$ = magnitude of local gravitational acceleration

λ = local geodetic latitude

L^* = local hour angle of Aries.

The selection of state variables in this alignment problem results in a time-varying observation matrix, but trivial time propagation equations for the state and its covariance matrix. In general, such a compromise is desirable in applying the Kalman filter because

(a) the covariance time propagation equation is a matrix equation, requiring integration or the calculation of a state transition matrix,

(b) the observation matrix H is calculated only when data is received, and need not be stored between data times, and

(c) no numerical inaccuracy is introduced in propagating the state or its covariance between data times.

3.3 Batched Filter for the Problem

The recursive nature of the Kalman filter lends itself to an alignment procedure that may have to be terminated at any time. Such is the case when a quick reaction capability is needed, as in rescue launches or when the usual alignment procedure has failed a short time before launch. At the other extreme, the operation of a recursive filter over an extended period of time can result in numerical and geometric difficulties[19-21], adequately discussed elsewhere in this publication, that make its continued use of dubious value.

If adequate time is available to collect, say, N pieces of data and if it is desired to examine these data all at once after the last is received, batch processing formulas are more appropriate than recursive ones. In that case one may assemble the data $Z(t_i)$, $i = 1,2,\ldots N$, into a single data vector

$$Z_{N,1} = \begin{bmatrix} Z(t_1) \\ Z(t_2) \\ \vdots \\ Z(t_N) \end{bmatrix} , \tag{3.17}$$

calculate the estimated observations, using only the estimated state at t_0,

$$\hat{Z}_{N,1} = \begin{bmatrix} -\bar{g}(t_1) - \bar{g}(t_1)\times\hat{\delta}_0^+ \\ -\bar{g}(t_2) - \bar{g}(t_2)\times\hat{\delta}_0^+ \\ \vdots \\ -\bar{g}(t_N) - \bar{g}(t_N)\times\hat{\delta}_0^+ \end{bmatrix} , \tag{3.18}$$

and apply the Kalman Filter equations in the form

$$\hat{\delta}_N^+ = \hat{\delta}_0^+ + P_N^+ H_{N,1}^T R_{N,1}^{-1}(Z_{N,1} - \hat{Z}_{N,1}) \tag{3.19}$$

$$P_N^+ = [(P_0^+)^{-1} + H_{N,1}^T R_{N,1}^{-1} H_{N,1}]^{-1} , \tag{3.20}$$

where,

$$H_{N,1}^T = [H_1^T, H_2^T, \ldots, H_N^T] \tag{3.21}$$

$$R_{N,1}^{-1} = \begin{bmatrix} R_1^{-1} & 0 & \cdots & 0 \\ 0 & R_2^{-1} & & 0 \\ \vdots & \vdots & & \\ 0 & 0 & 0 & R_N^{-1} \end{bmatrix} \tag{3.22}$$

When the matrix multiplications are carried out it is found that no more than 3×3 matrices are involved for

$$H_{N,1}^T R_{N,1}^{-1}(Z_{N,1} - \hat{Z}_{N,1}) = \sum_{i=1}^{N} H_i^T R_i^{-1}[Z(t_i) + \bar{g}(t_i) + \bar{g}(t_i) \times \hat{\theta}_0^+] \tag{3.23}$$

and

$$H_{N,1}^T R_{N,1}^{-1} H_{N,1} = \sum_{i=1}^{N} H_i^T R_i^{-1} H_i \ . \tag{3.24}$$

These formulas reduce to a weighted least squares estimate, as expected[22], when $\hat{\theta}_0^+ = 0$ and $(P_0^+)^{-1} = 0$.

3.4 Physical Significance and Accuracy of Filtering

The physical significance of the alignment filtering is easily seen in Figure 13. The true gravity vector at the alignment site rotates in a predetermined cone about the earth's spin axis. This gravity vector, as seen through the misaligned inertial sensors, is estimated to be $\bar{g}_{est} = \bar{g}(t) + \bar{g}(t) \times \hat{\theta}$, where $\hat{\theta}$ is the estimate of misalignment, while in actuality it is $\bar{g}_{true} = \bar{g}(t) + \bar{g}(t) \times \theta$, where θ is the true misalignment. Both $\bar{g}_{est}$ and $\bar{g}_{true}$ rotate at the same rate as $\bar{g}$ about the earth's spin axis and sweep out parts of cones from t_0 to t_N. The function of the filter is to bring these two partial cones into near coincidence by bringing $\hat{\theta}$ closer to θ. In the case of the recursive form, an adjustment is made at each data point, while in the batch form the estimate cone piece is adjusted to fit the data (and $\hat{\theta}_0^+$) in a single step. This picture of adjusting a portion of a cone to fit the observations explains why, for given $t_1, t_2, \ldots, t_N$ and given $\bar{V}(t_i)$, the alignment process is most accurate at the equator and inoperative at the poles: the conic surface generated by $\bar{g}$ in a given time is greatest at the equator and nil at the poles. Analytically, the observability matrix for the problem is just Equation (3.21), which may be written

$$H_{N,1} = \begin{bmatrix} [\bar{g}(t_1) \times] \\ [\bar{g}(t_2) \times] \\ \vdots \\ [\bar{g}(t_N) \times] \end{bmatrix} \tag{3.25}$$

and this is seen to be singular if $\bar{g}(t_1) = \bar{g}(t_2) = \ldots = \bar{g}(t_N)$, which occurs at the poles of the earth.

If a general measure of the accuracy of the alignment is desired, one may employ the ratio of the determinant of $P_{N,1}^+$ to the determinant of P_0^+. Potter and Fraser[23] have shown that this ratio is, in the present case,

$$|P_{N,1}^+| / |P_0^+| = |R_{N,1}| / |R_{N,1} + H_{N,1} P_0^+ H_{N,1}^T| \ . \tag{3.26}$$

From this, it is seen that a large value of $H_{N,1} P_0^+ H_{N,1}^T$ will result in a large reduction of initial alignment error. This matrix, however, is composed of N^2 3×3 sub-matrices, of which the $(i,j)^{th}$ sub-matrix is

$$\begin{aligned} H_i P_0^+ H_j^T &= [\bar{g}(t_i) \times] E(\hat{\theta}_0^+ \hat{\theta}_0^{+T}) [\bar{g}(t_j) \times]^T \\ &= E[(\bar{g}(t_i) \times \hat{\theta}_0^+)(\bar{g}(t_j) \times \hat{\theta}_0^+)^T] \ , \end{aligned} \tag{3.27}$$

where E indicates expectation. If the earth were not rotating, $\bar{g}(t_i) = \bar{g}(t_j) = \bar{g}(t_0)$ for $0 \leq i, j \leq N$. In this case an initial misalignment $\hat{\theta}_0^+$ parallel to $\bar{g}(t_0)$ would make all the sub-matrices zero. Hence one may conclude, as expected, that the earth's rotation is necessary to remove initial misalignments about the vertical. Similarly, one can conclude that initial misalignments about axes normal to $\bar{g}(t_0)$ are reduced even without the earth's rotation.

4. UNIDENTIFIED LANDMARK NAVIGATION

4.1 Introduction

Unidentified landmark navigation[24-27] is an orbital navigation technique that relies heavily upon the recursive properties of the Kalman filter. The method produces improvement of vehicle position and velocity information

using only passive, on-board measurements. Since the essential information required is the direction from the vehicle to an unidentified ground point, relative to inertial space, the typical instrumentation comprises a landmark telescope, a star tracker, and a computer. It is not necessary to store an extensive list of landmark locations and identification aids. Moreover, the complete elimination of the astronaut's participation is also a possibility, because it is necessary only to track patterns and not to identify them.

The methods available fall into two groups: (1) those such as Reference 24, in which the landmark location biases are introduced as state variables, resulting in a nine-dimensional state vector; and (2) those similar to Reference 25, in which the landmark biases do not appear explicitly in the state vector, which is then the usual six-vector of position and velocity deviations.

The first type, comprising what will be termed the one-sight and the two-sight methods will be discussed here because they clearly illustrate the role of the Kalman filter. A discussion of both types of methods is given in Reference 26.

Assuming that a computer on the orbiting vehicle is available and that a Kalman filter is employed to process navigational data from other than landmarks, the following additional computations must be performed in order to implement the unidentified landmark method under consideration: (1) a larger transition matrix of the state vector; (2) the estimated value of the landmark observation minus the true value; (3) the measurement matrix, relating the landmark observation to the state; (4) the best initial estimate of the landmark's position; and (5) the initial covariance of the state when a new landmark is located. These five required calculations will be discussed in turn.

4.2 State Transition Matrix

The nine state variables $\bar{r}$, $\bar{v}$, $\bar{l}$, representing deviations from nominal in position, velocity, and landmark location, have a state transition matrix from t_0 to t given by

$$\phi(t,t_0) = \begin{bmatrix} \phi_{rr} & \phi_{rv} & 0 \\ \phi_{vr} & \phi_{vv} & 0 \\ 0 & 0 & \phi_{ll} \end{bmatrix} , \qquad (4.1)$$

where the four upper left sub-matrices constitute the usual error propagation matrix for vehicle position and velocity, and the lower right sub-matrix gives the transition between landmark biases $\bar{l}_0$ at time t_0 and their value $\bar{l}$ at time t. If $\bar{l}$ is in a planet-fixed coordinate system, then ϕ_{ll} is the identity matrix I; if $\bar{l}$ is in inertial space, then

$$\phi_{ll} = (\Omega) ,$$

where (Ω) is the rotational transformation representing planetary rotation from time t_0 to t. The most convenient coordinate system for $\bar{l}$ is one fixed in the planet, at least from the point of view of the state transition matrix computation.

4.3 Estimated Observations

The observation is some function of vehicle position $\bar{R}$, velocity $\bar{V}$, and landmark location $\bar{L}$ determined by the instrumentation. The computer calculates the observation Y_E using estimated $\bar{R}$, $\bar{V}$, and $\bar{L}$ and subtracts it from the actual observation Y_T supplied by the instruments. The difference $Y_T - Y_E$, weighted by the linear filter matrix, is used to correct the deviations $\bar{r}$, $\bar{v}$, $\bar{l}$. In practice, it is often computationally expedient to consider the "observation" to be some function of the actual instrument readouts.

In Table IV the first column gives the instrument readouts, the second column gives the computation of the observation Y_T in terms of these readouts, and the third column gives the computation of the observation Y_E from estimated $\bar{R}$, $\bar{V}$, and $\bar{L}$. It will be noticed that the landmark sensor angles T (train angle) and D (depression angle) may be processed together to yield the unit vector measurement $\hat{\rho}$ or processed separately to give two scalar measurements. Also, the calculation of $(S)_E$, since it does not concern landmark navigation exclusively, is not given in detail.

4.4 Measurement Matrices

The third calculation required for the linear filter is the matrix M relating the state X to the observation deviations $\bar{y}$. In the present case, X includes deviations $\bar{l}$ of the estimated landmark location from the nominal location. The measurement matrix may be obtained by differentiation of the estimated observation function, shown in column 3 of Table IV, with respect to $\bar{R}$, $\bar{V}$, $\bar{L}$. The results for the two choices of landmark observation Y_E are given in Table V.

4.5 Initial Estimate of Landmark Position

From the estimated vehicle position and the observed direction in inertial space from vehicle to landmark, the on-board computer must extract a first estimate of landmark location. Two ways of doing this will be described.

One-sight method

If the moon or planet is assumed spherical of radius R_E near the landmark, then the observed line of sight $\hat{\rho}$ may be extrapolated from the vehicle position $\bar{R}$ to where it intersects the sphere. This is the estimated position $\bar{L}$ of the landmark. It is computed as

$$\bar{L} = \bar{R} + \rho\hat{\rho} , \tag{4.2}$$

where

$$\rho = -\hat{\rho}\cdot\bar{R} - [R_E^2 - \bar{R}^2 + (\bar{R}\cdot\hat{\rho})^2]^{1/2} .$$

The major inaccuracy of this method is in R_E. Even if an ellipsoidal formula is used to compute R_E for the earth, or a table stored, the terrain on which the landmark is situated may depart from these approximations by a significant distance.

Two-sight method

If two sightings $\hat{\rho}_1$ and $\hat{\rho}_2$ are taken at times t_1 and t_2 and the estimated vehicle positions $\bar{R}_1$ and $\bar{R}_2$ recorded, then the on-board computer may estimate the location of the landmark with an error proportional to the average sighting distance times the instrumentation angular uncertainty. In this case, Reference 24 gives the formulas for $\bar{L}$ for a non-rotating planet. In vector form, they are

$$\bar{L}_1 = (\bar{L}_1' + \bar{L}_1'')/2 , \tag{4.3}$$

where

$\bar{L}_1$ = two-sight estimate of $\bar{L}$ at t_1

$$\bar{L}_1' = \bar{R}_1 + (\bar{R}_2 - \bar{R}_1)\cdot(\hat{\rho}_1 - c\hat{\rho}_2)\hat{\rho}_1/(1-c^2)$$

$$\bar{L}_1'' = \bar{R}_2 - (\bar{R}_2 - \bar{R}_1)\cdot(\hat{\rho}_2 - c\hat{\rho}_1)\hat{\rho}_2/(1-c^2)$$

and

$$c = \hat{\rho}_1\cdot\hat{\rho}_2 .$$

If the planet is rotating with constant angular velocity Ω, then, as shown in Appendix C, it is necessary only to replace $\bar{R}_2$ by $(\Omega)^{-1}\bar{R}_2$ and $\hat{\rho}_2$ by $(\Omega)^{-1}\hat{\rho}_2$, where (Ω) is the matrix representing planetary rotation from t_1 to t_2.

Use of the two-sight method introduces a slight sequencing problem. It is obvious from the geometry that the two sights upon which the initial landmark estimate is based should be well-separated, preferably with 90° between $\hat{\rho}_1$ and $\hat{\rho}_2$. But the linear filter cannot process readings until the initial landmark estimate is made. Hence, in this method, it is necessary to collect all the readings on a given landmark, select two well-separated for the two-sight estimation, and then process the rest in order. If twelve T and D readings and their times are taken in the process, then about 36 storage locations are required. The sequence of operations is as follows:

(1) Store $\bar{R}_1$.

(2) Collect and store $\hat{\rho}_i$ and t_i, $i = 1,\ldots,N$.

(3) Extrapolate $\bar{R}_1$ to $\bar{R}_N$.

(4) Calculate $\bar{L}_1$ from Equation (4.3), using $\bar{R}_N$ and $\hat{\rho}_N$ in place of $\bar{R}_2$, $\hat{\rho}_2$.

(5) Process the data $\hat{\rho}_i$, t_i for $i = 2,\ldots,N-1$ as though they occurred in real time. The result is $\bar{R}_{N-1}$ and $\bar{V}_{N-1}$, the best estimates at t_{N-1}.

4.6 Initial Covariance Matrix

The initialization of the covariance matrix of landmark and vehicle errors when a new landmark is placed can be a critical on-board computation, as the computer runs to be discussed later will show. The extent of the on-board computation depends on the method employed to estimate the landmark location when it is first sighted. The two methods just described lead to different covariance matrices. For the first, the one-sight method, the proper covariance is

$$\mathrm{cov}(X) = \begin{bmatrix} P_{rr} & P_{rv} & P_{rl} \\ P_{rv}^T & P_{vv} & P_{vl} \\ P_{rl}^T & P_{vl}^T & P_{ll} \end{bmatrix} , \tag{4.4}$$

where

$$\left.\begin{aligned} P_{rl} &= P_{rr}U^T, \qquad P_{vl} = P_{vr}U^T \\ P_{ll} &= UP_{rr}U^T + \tfrac{1}{2}\rho^2\sigma_\phi^2 U[I - \hat{\rho}\hat{\rho}^T]U^T + \frac{\bar{\rho}\bar{\rho}^T\sigma_T^2}{(\bar{\rho}^T\bar{L}/L)^2} \end{aligned}\right\} \tag{4.5}$$

and

$$U = I - \bar{\rho}\bar{L}^T/\bar{\rho}^T\bar{L}, \qquad \bar{\rho} = \bar{L} - \bar{R}, \qquad \hat{\rho} = \bar{\rho}/\rho \tag{4.6}$$

σ_ϕ^2 = variance of instrumentation angular noise

σ_T^2 = variance of terrain uncertainty.

For the second method, the two-sight method, the proper covariance is the same form as Equation (4.4), but where

$$\left.\begin{aligned} P_{rl} &= (AP_{rr} + BP_{vr})^T, \qquad P_{vl} = (AP_{rv} + BP_{vv})^T \\ P_{ll} &= AP_{rr}A^T + AP_{rv}B^T + BP_{rv}^T + BP_{vv}B^T + D_1K_1D_1^T + D_2K_2D_2^T \end{aligned}\right\} \tag{4.7}$$

and

$$\left.\begin{aligned} A &= (\tfrac{1}{2})(I + \phi_{rr}) \\ B &= (\tfrac{1}{2})\phi_{rv} \\ D_1 &= (\tfrac{1}{2})\rho_1[I + \hat{\rho}_2^*\hat{\rho}_1^{*T}]U_1 \\ D_2 &= (\tfrac{1}{2})\rho_2[I + \hat{\rho}_1\hat{\rho}_1^T]U_2 \\ K_1 &= (\tfrac{1}{2})\sigma_{\phi 1}^2[I - \hat{\rho}_1\hat{\rho}_1^T] \\ K_2 &= (\tfrac{1}{2})\sigma_{\phi 2}^2[I - \hat{\rho}_2^*\hat{\rho}_2^{*T}] \end{aligned}\right\} \tag{4.8}$$

In the preceding, $\hat{\rho}^*$ is written for $(\Omega)^{-1}\hat{\rho}$, and the subscripts 1 and 2 refer to the times of the first and last sightings. The quantities I, ϕ, ρ, σ_ϕ^2, U, and (Ω) have the same meanings as previously. Note that σ_T^2, the variance of terrain roughness, does not appear in the second method, as to be expected. The derivations of these initial covariances are rather involved and will not be presented here.

4.7 Computer Simulation

Because the one-sight method appears to require less computation than other methods, a program was constructed[44] to try it out. The program calculated the covariance of the nine-variable state by the linear filter equations. It simulated white-noise on the readings and the terrain error in estimating the landmarks. The error in the optimal estimate of the state was calculated and compared to the covariance matrix as an approximate check.

The first significant result obtained is that the one-sight method does not work unless the covariance matrix is properly initialized for each new landmark. This result was obtained in several computer runs, a typical pair of which is shown in Figures 14 and 15. The conditions for these runs are listed in Table VI.

Figure 14 shows the reduction in the magnitude of the actual position error and the square root of the trace of the position covariance when the covariance matrix was initialized for each landmark as:

$$\text{cov}(\bar{x}) = \begin{bmatrix} P_{rr} & P_{rv} & 0 \\ P_{rv}^T & P_{vv} & 0 \\ 0 & 0 & D \end{bmatrix} = 9\times 9, \tag{4.9}$$

where P_{rr}, P_{rv}, P_{vv} are the 3×3 sub-matrices of the vehicle position and velocity at the time of sighting the landmark, and D is a diagonal matrix of elements each equal to the trace of P_{rr}. Although this initialization procedure seems conservative, the results are disappointing; while the covariance matrix gradually approaches zero, the actual error, which it supposedly represents, goes to infinity.

In contrast, Figure 15 shows the identical computer run (same instrument noise, same terrain error, etc.), but with the covariance matrix initialized as given by Equations (4.4), (4.5) and (4.6). The actual error here

is in much better agreement with the covariance matrix. From these and other similar pairs of runs, it is concluded that the covariance of the landmark error and cross correlation of landmark and vehicle errors must be calculated correctly at each new landmark for the one-sight method.

The computer runs just discussed are for a lunar satellite. Figure 18 shows the results of the one-sight method for an earth satellite. Seven landmarks are observed, all on the dayside. About six pairs of readings (T and D), spaced 7.2 seconds apart, are taken on each landmark. This measurement schedule is judged to be within the capability of a human operator, cloud cover and other duties allowing.

With the run shown in Figure 18 as a nominal case, the following parameters were perturbed: sensor noise, orbital altitude ($e = 0$), terrain error. The results are shown in Figures 17-19. These curves represent the square root of the trace of the covariance matrix immediately after, and one-half hour after, the last landmark observation.

5. CONCLUDING REMARKS

The facility with which the Kalman filter dispatched the complicated estimation problems described in Sections 2, 3 and 4 should not obscure two serious difficulties in its use, of which only a very brief discussion will be given here:

(1) The need to linearize the problem without loss of accuracy.

(2) The presence of computational difficulties.

5.1 Nonlinearity

The linearity problem enters, generally, through the approximation made in the state transition matrix or that made in the observation matrix. These two matrices represent linear approximations to what may be nonlinear relations.

In the case of orbital navigation, the measurement nonlinearity often produces a more serious error than does the state transition matrix approximation. This is usually true for radar range readings[3]. The problem of measurement nonlinearity may be overcome in an error analysis, such as described in Section 2, by using data simulated from a linear model, rather than real data or data generated by a nonlinear model. An error analysis thus constructed will not have nonlinearity errors, but will not be a reliable guide to the nonlinearity effects to be encountered in a program that processes real data.

In the case of inertial system alignment, Section 3, nonlinearity can enter only in the measurement matrix, since the state transition matrix is exactly the identity. The extent of the nonlinearity error depends on the size of the estimates $\hat{\epsilon}$.

Nonlinearity errors can be expected to occur in unidentified landmark navigation, both because of the measurement model and because of the state equation. The simulations described in Section 4, however, were based on a nonlinear measurement process. For this reason it would seem that measurement nonlinearity does not seriously affect the landmark navigation methods of Section 4.

5.2 Computational Difficulties

All three applications discussed in this chapter are susceptible to the numerical inaccuracies caused by a finite computer word length. These enter in many ways and have varied mechanisms and cures[19-21], the discussion of which is taken up elsewhere in this publication. Only two areas of difficulty, which have appeared in the simulation programs of Sections 2 and 4 of this chapter, will be mentioned. They are the calculation of the covariance matrix

$$P^+ = P^- - KHP^- ,$$

and the calculation of the matrix product

$$PH^T .$$

The first of these leads to an inaccurate result if P^- and KHP^- are almost equal. This is because the magnitude of the error (which is determined by the number of digits employed in the computer), divided by the magnitude of the answer (which may be extremely small if P^- and KHP^- are almost equal), yields a large fractional error in the new P^+ matrix. This situation arises, as it did in the simulations of Section 2, when a series of highly accurate range readings are taken by a ground radar on a satellite whose position uncertainty in approaching the station is very large. The initially large covariance matrix is reduced by a series of subtractions of the form $-KHP$ to one that is smaller by several orders of magnitude, at least in the direction of the line of sight from the station to the vehicle. If these orders of magnitude equal or exceed the number of digits in the computer word, all numerical significance is lost.

The second type of computation, PH^T, causes numerical inaccuracies in the same circumstances as the first. In fact it can be seen that the covariance calculation of Equation (2.2), for the case of $R_i = 0$, leads to a new covariance matrix P^+ such that $P^+H^T = 0$. But neither P^+ nor H^T is zero; each zero in the product must be the result of a computation of the form

$$\sum_k (P^+)_{ik} (H^T)_{kj} .$$

If the elements of P^+ and of H^T are large, the zero result can come about only by the subtraction of equal, large numbers with the attendant loss of numerical significance. In the case of radar range readings, H^T is just $\hat{\rho}_0$, the unit vector from ground station to vehicle; after many highly accurate range readings the vehicle's position uncertainty, as represented by P, may still be large, but is distributed almost exclusively in a plane normal to $\hat{\rho}_0$. Hence the product $P\hat{\rho}_0$ is almost zero, even though the individual elements of P may be large. This situation is typical of many in which highly accurate, closely-spaced observations are made.

ACKNOWLEDGMENTS

The work on which Sections 2 and 4 of this chapter are based was largely carried out while the author was employed in the Sperry Rand Systems Group, Division of the Sperry Rand Corporation, Great Neck, NY. The efforts of E. Wu and R. Massey in constructing the computer program of Section 2 are gratefully acknowledged.

REFERENCES

General

1. Battin, R.L. — *A Statistical Optimizing Navigation Procedure for Space Flight.* American Rocket Society Journal, Vol. 32, No. 11, Nov. 1962, pp. 1681-1686.

2. McLean, J.D. et al. — *Optimal Filtering and Linear Prediction Applied to a Space Navigation System for the Circumlunar Mission.* NASA TN D-1208, March 1962.

3. Denham, W.F. Pines, S. — *Sequential Estimation When Measurement Function Nonlinearity is Comparable to Measurement Error.* AIAA Journal, Vol. 4, No. 6, June 1966, p. 1071ff.

References for Section 2

4. Andrews, A. — *A Square Root Formulation for the Kalman Covariance Equation.* AIAA Journal, Vol. 6, 1968, pp. 1165-1166.

5. Kochi, K.C. — *Exact First-Order Navigation-Guidance Mechanization and Error Propagation Equations for Two-Body Reference Orbits.* AIAA Journal, Vol. 2, 1964, pp. 385-388.

6. Kochi, K.C. — *Exact Two-Body Error Sensitivity Coefficients.* AIAA Journal, Vol. 2, 1964, p. 1503.

7. Maxwell, J.A. Dorfman, M. — *The Accuracy of Satellite Orbit Determination.* AIAA Paper 63-155, June 1963.

8. Moe, K. — *Stochastic Models for Errors in Orbit Predictions for Artificial Earth Satellites.* American Rocket Society Journal, Vol. 32, No. 11, Nov. 1962, pp. 1726-1728.

9. Karrenberg, H.K. et al. — *Variations of Satellite Position with Uncertainties in the Mean Atmospheric Density.* American Rocket Society Journal, Vol. 32, No. 4, April 1962, p. 576ff.

10. Rauch, H.E. — *Optimum Estimation of Satellite Trajectories Including the Effect of Random Variations in Drag.* AIAA Paper 64-396, July 1964.

11. Wengert, R.E. Jones, J.F. — *Orbit Determination Accuracies Using Systematic Error Adjustments.* AIAA Paper 64-395, July 1964.

12. Guttmann, P.T. — *The Effects of Uncertainties in Some Zonal Harmonics of the Earth's Gravity Potential on Satellite Position Detection.* SSD TDR 63-174 TDR-269 (4110)-2, AD415437.

13. Kaula, W.M. — *Determination of the Earth's Gravitational Field.* Review of Geophysics Vol. 1, No. 4, Nov. 1963, pp. 507-551. (Also NASA RP-132.)

14. Kaula, W.M. — *The Comparison and Combination of Satellite with Other Determination of Geodetic Parameters.* Goddard Space Flight Center Contributions to the COSPAR Meeting, May 1962, pp. 119-124.

15. Kaula, W.M. — *A Review of Geodetic Parameters.* NASA TN D-1847, May 1963.

16. Proskurin, V.F. Batrakov, V.V. — *Perturbations in the Motion of Artificial Satellites Due to the Oblateness of the Earth.* American Rocket Society Journal, Vol.31, No.1, Jan.1961, pp.117-123. (Translated by Research Information Service, New York, from Byulleten' Instituta Teoreticheskoy Astronomoi, Vol.7, No.7 (90), 1960, pp.537-548.)

17. McArthur, W.G. — *Horizon Sensor Navigation Errors Resulting from Statistical Variations in the CO_2 14-16 Micron Radiation Band.* Published in the Proceedings of the 5th Symposium on Ballistic Missile and Space Technology, San Diego, California, August 12-14, 1964.

18. Wilcox, J.C. — *Self-Contained Orbital Navigation Systems with Correlated Measurement Errors.* Published in the Proceedings of the AIAA/ION Guidance and Control Conference, Minneapolis, Minnesota, 16-18 August, 1965. (The variances referred to are the diagonal elements of Equation (42) of this Reference.)

References for Section 3

19. Schlee, F.H. et al. — *'ergence in the Kalman Fi'ter.* AIAA Journal, Vol.5, 1967, pp.1114-1120.

20. Mendelsohn, J. O'Dwyer, W. — *Astrody. mics of Lunar Satellites, Part I: Orbit Determination. Vol.I - Summary and Conclusions, Vol.II - Derivations* Grumman Final Report on Contract NAS 8-5015, NASA Report RE-168, Oct.1963.

21. Bellantoni, J.F. Dodge, K.W. — *A Square-Root Formulation of the Kalman-Schmidt Filter.* AIAA Journal, Vol.5, No.7, Aug.1967, p.1309ff.

22. Smith G.L. et al. — *Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle.* NASA TR R-135, Appendix F, 1962.

23. Potter, J.E. Fraser, D.C. — *A Formula for Updating the Determinant of the Covariance Matrix.* AIAA Journal, Vol.5, No.7, July 1967, pp.1352-1354.

References for Section 4

24. DeNessa, E.J. Dittrich, M.S. — *Orbit and Landmark Determination During Lunar Orbit.* Master's Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, May 17, 1963.

25. Levine, G.M. — *A Method of Orbital Navigation Using Optical Sightings to Unknown Landmarks.* Published in the Proceedings of the AIAA/ION Guidance and Control Conference, New York, 1965.

26. Bellantoni, J.F. — *Unidentified Landmark Navigation.* AIAA Journal, Vol.5, No.8, 1967, pp.1478-1483.

27. Toda, N.F. Schlee, F.H. — *Autonomous Orbital Navigation by Optical Tracking of Unknown Landmarks.* Published in the Proceedings of the AIAA/JACC Guidance and Control Conference, New York, 1966.

APPENDIX A

Derivation of Equation (2.7)

If the linear filter K is used to update the state estimate $\hat{x}^-$, what is the covariance P^+ of the error in the resulting estimate, $\hat{x}^+$? Let

$$\hat{x}^+ = \hat{x}^- + K(z - H\hat{x}^-) ,$$

where

$$z = Hx + v$$

$$E(v) = 0$$

$$E(vv^T) = R = \text{known}$$

$$E(x - \hat{x}^-)v^T = Ev(x - \hat{x}^-)^T = 0 .$$

A direct calculation of P^+ from its definition gives

$$\begin{aligned} P^+ &= E(x - \hat{x}^+)(x - \hat{x}^+)^T \\ &= E[x - \hat{x}^- - K(z - H\hat{x}^-)][x - \hat{x}^- - K(z - H\hat{x}^-)]^T \\ &= E[e^- - KHe^- - Kv][e^- - KHe^- - Kv]^T \\ &= E[(I - KH)e^- - Kv][(I - KH)e^- - Kv]^T \\ &= (I - KH)E[(e^-)(e^-)^T](I - KH)^T + KE[(v)(v)^T K^T] - KE[(v)(e^-)^T](I - KH)^T - (I - KH)E[(e^-)(v)^T]K^T \\ &= (I - KH)P^-(I - KH)^T + KRK^T , \end{aligned}$$

where e has been used for $x - \hat{x}$ and P for $E(ee^T)$.

APPENDIX B

State Transition Matrix for Air Drag, Gravity and Oblateness Coefficients

This Appendix derives the differential equations for vehicle position and velocity errors produced by constant errors in the drag parameter, oblateness coefficients and gravitational parameter of an earth satellite.

The vehicle position $\bar{R}$ and velocity $\bar{V}$ are given by

$$\frac{d}{dt}\begin{bmatrix} \bar{R} \\ \bar{V} \end{bmatrix} = \begin{bmatrix} \bar{V} \\ \bar{F}_\mu + \bar{F}_{OS} + \bar{D} \end{bmatrix} , \tag{B.1}$$

where

$$\bar{F}_\mu = -\mu\bar{R}/R^3$$

$$\bar{F}_{OS} = -\partial U_{OS}/\partial\bar{R}$$

$$\bar{D} = \tfrac{1}{2}D_0 V\bar{V}e^{-K_0(R-R_0)}$$

$$U_{OS} = -\mu\sum_{n=1}^{\infty} C_n(R_E^n/R^{n+1})P_n(z/R)$$

and

P_0, K_0, R_0 = drag constants

R_E = radius of earth corresponding to C_0

D_0 = drag parameter

C_n = coefficient of n^{th} zonal harmonic in geopotential

$P_n(s)$ = n^{th} Legendre polynomial

μ = gravitational parameter of earth.

It is easily shown that

$$\bar{F}_{OB} = (\mu/R^3)\sum_{n=1}^{\infty} C_n(R_E/R)^n\left[\bar{K}R\frac{dP_n(s)}{ds} - (n+1)P_n(s)\bar{R} - s\frac{dP_n(s)}{ds}\bar{R}\right], \tag{B.2}$$

where

s = z/R

$\bar{K}$ = unit vector normal to equator.

The deviations $\delta\bar{R}$ and $\delta\bar{V}$ in position and velocity corresponding to biases $\delta\mu$, δC_i, $\delta D_0/D_0$, in gravitational parameter μ, oblateness coefficients C_i, and drag parameter D_0, follow the differential equations

$$\frac{d}{dt}\begin{bmatrix}\delta\bar{R}\\ \delta\bar{V}\end{bmatrix} = \begin{bmatrix}0 & I\\ M_R & M_V\end{bmatrix}\begin{bmatrix}\bar{R}\\ \bar{V}\end{bmatrix} + \begin{bmatrix}0\\ \delta\bar{F}_\mu + \delta\bar{F}_{OB} + \delta\bar{D}\end{bmatrix}. \tag{B.3}$$

where

$M_R = \partial\bar{F}_\mu/\partial\bar{R} + \partial\bar{F}_{OB}/\partial\bar{R} + \partial\bar{D}/\partial\bar{R}$

$M_V = \partial\bar{D}/\partial\bar{V}$

and

$$\partial\bar{F}_\mu/\partial\bar{R} = -\mu[I - 3\bar{R}\bar{R}/R^2]/R^3$$

$$\partial\bar{F}_{OB}/\partial\bar{R} = (3\mu/2)R_E^2\frac{C_i}{R^7}\left\{\left[(R^2-z^2)I + \left(\frac{7z^2}{R^2}-5\right)\bar{R}\bar{R}\right] - 2z[\delta\bar{K}\bar{R}+\bar{R}\bar{K}] + 2R^2\bar{K}\bar{K}\right\}$$

$$\partial\bar{D}/\partial\bar{R} = -K_0D_0V\bar{V}\bar{R}e^{-K_0(R-R_0)}/2R_0$$

$$\partial\bar{D}/\partial\bar{V} = \tfrac{1}{2}(D_0/V)[\bar{V}\bar{V} + V^2I]e^{-K_0(R-R_0)}$$

$$\delta\bar{F}_\mu = -\bar{R}\delta\mu/R^3$$

$$\delta\bar{F}_{OB} = \frac{\mu}{R^3}\sum_{n=1}^{\infty}\left(\frac{R_E}{R}\right)^n\left[-P_n(n+1) + s\frac{dP_n}{ds}\bar{R} + R\frac{dP_n}{ds}\bar{K}\right]\delta C_n$$

$$\delta\bar{D} = (\tfrac{1}{2})D_0V\bar{V}e^{-K_0(R-R_0)}\delta D_0/D_0 .$$

The state transition matrix, Φ_0, is obtained by integrating from t_0 to t the matrix differential equation corresponding to Equation (B.3):

$$\left.\begin{aligned}\frac{d}{dt}[\Phi_D] &= \begin{bmatrix}0 & I\\ M_R & M_V\end{bmatrix}[\Phi_D] + [A]\\ [\Phi_D(0)] &= [0] .\end{aligned}\right\} \tag{B.4}$$

where

$$[A] = \left[\begin{array}{c|c|c|c|c|c} 0 & 0 & 0 & 0 & 0 & 0\\ \hline \dfrac{\delta\bar{F}_\mu}{\delta\mu} & \dfrac{\delta\bar{F}_{OB}}{\delta C_j} & \dfrac{\delta\bar{F}_{OB}}{\delta C_j} & \dfrac{\delta\bar{F}_{OB}}{\delta C_k} & \dfrac{\delta\bar{F}_{OB}}{\delta C_l} & \dfrac{\delta\bar{D}_0}{\delta D_0/D_0}\end{array}\right]$$

The typical quantity $\delta\bar{F}_{osn}$ in the above is

$$\frac{\delta\bar{F}_{osn}}{\delta C_n} = \frac{\mu}{R^3}\left(\frac{R_E}{R}\right)^n \left\{ -\left[(n+1)P_n/s) + s\,\frac{dP_n(s)}{ds}\right]\bar{K} + R\,\frac{dP_n(s)}{ds}\,\bar{K}\right\},$$

where

$$P_{n+1}(s) = \frac{2n+1}{n+1}\,sP_n(s) - \frac{n}{n+1}\,P_{n-1}(s)$$

$$\frac{dP_{n+1}(s)}{ds} = \frac{2n+1}{n+1}\left[s\,\frac{dP_n(s)}{ds} + P_n(s)\right] - \frac{n}{n+1}\,\frac{dP_{n-1}(s)}{ds}$$

and

$$P_0 = 0\,, \qquad \frac{dP_0}{ds} = 0$$

$$P_1 = s\,, \qquad \frac{dP_1}{ds} = 1\,.$$

Equation (B.4) will yield the transition matrix relating position and velocity errors to the normalized biases $\delta\mu$, δC_i, δC_j, δC_k, δC_l and $\delta D_0/D_0$. The choice of i, j, k, l depends on which oblateness coefficients are of interest.

APPENDIX C

Two-Sight Landmark Location for a Rotating Planet

The problem is to obtain an estimate $\bar{L}_1$ of landmark position at time t_1 given observation $\hat{\rho}_1$ and vehicle position $\bar{R}_1$ at t_1, and $\hat{\rho}_2$, $\bar{R}_2$ at time t_2.

The landmark position at t_1 may be estimated as

$$\bar{L}_1 = \bar{R}_1 + \rho_1\hat{\rho}_1\,, \tag{C.1}$$

where ρ_1 is unknown. The landmark location at time t_2 may be estimated as

$$\bar{L}_2 = \bar{R}_2 + \rho_2\hat{\rho}_2\,. \tag{C.2}$$

where ρ_2 is unknown. Since $\bar{L}_2 = (\Omega)\bar{L}_1$, where (Ω) is the rotational transformation of the planet from t_1 to t_2, the last equation becomes

$$\bar{L}_1 = (\Omega)^{-1}\bar{R}_2 + \rho_2(\Omega)^{-1}\hat{\rho}_2\,. \tag{C.3}$$

Equations (C.1) and (C.3) are six scalar equations in five unknowns, and hence two solutions, $\bar{L}_1'$ and $\bar{L}_1''$, may be obtained by straightforward algebra:

$$\bar{L}_1' = \bar{R}_1 + (\bar{R}_2^* - \bar{R}_1)\cdot(\hat{\rho}_1\hat{\rho}_1 - c\hat{\rho}_2^*\hat{\rho}_1)/(1-c^2)\,, \tag{C.4}$$

$$\bar{L}_1'' = \bar{R}_2^* - (\bar{R}_2^* - \bar{R}_1)\cdot(\hat{\rho}_2^*\hat{\rho}_2^* - c\hat{\rho}_1\hat{\rho}_2^*)/(1-c^2)\,, \tag{C.5}$$

where

$$\left.\begin{aligned} \bar{R}_2^* &= (\Omega)^{-1}\bar{R}_2\,, \qquad \hat{\rho}_2^* = (\Omega)^{-1}\hat{\rho}_2 \\ c &= \hat{\rho}_1\cdot\hat{\rho}_2^*\,. \end{aligned}\right\} \tag{C.6}$$

These solutions are not identical because of noise on the data $\hat{\rho}_1$ and $\hat{\rho}_2$. Their average is a better estimate than either one alone. Thus,

$$\bar{L}_1 = (\bar{L}_1' + \bar{L}_1'')/2\,. \tag{C.7}$$

It is seen that the only effect of the planet's rotation is to introduce $(\Omega)^{-1}$ in front of $\bar{R}_2$ and $\hat{\rho}_2$.

TABLE I

Observation Matrices

Observation	*State variables*	*Observation matrix*
ρ	x , y , z	$\hat{\rho}_0$
	Δe , Δn , Δh	$(-\hat{\rho}_0 \cdot \hat{e},\ -\hat{\rho}_0 \cdot \hat{n},\ -\hat{\rho}_0 \cdot \hat{h})$
	$\Delta\rho$	1
$\dot{\rho}$	x , y , z	$[(\bar{V} - \bar{V}_0) - (\bar{V} - \bar{V}_0) \cdot \hat{\rho}_0 \hat{\rho}_0] / \rho_0$
	$\dot{x}$, $\dot{y}$, $\dot{z}$	ρ_0 / ρ_0
	Δe , Δn , Δh	$(\hat{e} \cdot \bar{\omega}_0,\ \hat{n} \cdot \bar{\omega}_0,\ \hat{h} \cdot \bar{\omega}_0)$
	$\Delta\dot{\rho}$	1
A	x , y , z	$(\hat{e} \cos A - \hat{n} \sin A) / \rho_0 \cos E$
	Δe , Δn , Δh	$\left\{\hat{e}(K_1 \rho_e^2 - \rho_n / \rho_{en}^2) + \hat{n}(K_1 \rho_e \rho_n + \rho_e / \rho_{en}^2) + \bar{\rho}_0 x [\hat{h} \rho_n K_1 + (\hat{n} \rho_n + \hat{e} \rho_e) / R_0 \rho_{en}^2]\right\}$
	ΔA	1
E	x , y , z	$(-\hat{e} \sin A \sin E - \hat{n} \cos A \sin E + \hat{h} \cos E) / \rho_0$
	Δe , Δn , Δh	$(\hat{e} \rho_e / R_0 + \hat{n} \rho_n / R_0 - \hat{h} + \hat{\rho}_0 \hat{\rho}_h / \rho_0) / \rho_{en}$
	ΔE	1
$\tilde{\Phi}_1$	x , y , z	$-\bar{R} R_{EH} / R^2 \sqrt{(R^2 - R_{EH}^2)}$
	$\tilde{h}_1$	$1 / \sqrt{(R^2 - R_{EH}^2)}$
	$\tilde{\dot{\Phi}}_2$	1
	Δh_1	$1 / \sqrt{(R^2 - R_{EH}^2)}$
	$\Delta\Phi_1$	1
$\tilde{\Phi}_2$	x , y , z	$-\hat{\eta} / R$
	θ_1 , θ_2 , θ_3	(0,0,-1) for IMU on $\hat{\xi}\hat{\eta}\hat{\zeta}$ axes
	θ_1 , θ_2 , θ_3	$-\zeta$ for IMU on XYZ axes
	$\tilde{h}_2$	$1 / \sqrt{(R^2 - R_{EH}^2)}$
	$\tilde{\dot{\Phi}}_2$	1
	$\Delta\Phi_2$	1
$\tilde{\Phi}_3$	x , y , z	$-\hat{\zeta} / R$
	θ_1 , θ_2 , θ_3	(0,1,0) for IMU on $\hat{\xi}\hat{\eta}\hat{\zeta}$ axes
	θ_1 , θ_2 , θ_3	$\hat{\eta}$ for IMU on XYZ axes
	$\tilde{h}_3$	$1 / \sqrt{(R^2 - R_{EH}^2)}$
	$\Delta\Phi_3$	1
	$\tilde{\dot{\Phi}}_3$	1

TABLE I (continued)

Observation	*State variables*	*Observation matrix*
TR_{STAR}	θ_1 , θ_2 , θ_3	$\hat{\xi} - \rho_{s\xi}(\rho_{s\eta}\hat{\eta} + \rho_{s\zeta}\hat{\zeta})/\delta'$, for IMU on XYZ axes
	θ_1 , θ_2 , θ_3	$(1, -\rho_{s\xi}\rho_{s\eta}/\delta', -\rho_{s\xi}\rho_{s\zeta}/\delta')$, IMU on $\hat{\xi}\hat{\eta}\hat{\zeta}$ axes
	Ω_1	1
	ΔTR_{STAR}	1
EL_{STAR}	θ_1 , θ_2 , θ_3	$-\hat{\rho}_s \times \hat{\xi}/\sqrt{\delta'}$, IMU on XYZ axes
	θ_1 , θ_2 , θ_3	$(0, -\rho_{s\zeta}/\sqrt{\delta'}, \rho_{s\eta}/\delta')$, IMU on $\hat{\xi}\hat{\eta}\hat{\zeta}$ axes
	Ω_2	2
	ΔEL_{STAR}	1
TR_{LMK}	x , y , z	$\hat{\xi} \times \hat{\rho}_L/\delta''$
	θ_1 , θ_2 , θ_3	$\hat{\xi} - \rho_{L\xi}(\rho_{L\eta}\hat{\eta} + \rho_{L\zeta}\hat{\zeta})/\delta''$, IMU on XYZ axes
	θ_1 , θ_2 , θ_3	$(1, -\rho_{L\xi}\rho_{L\eta}/\delta'', -\rho_{L\xi}\rho_{L\zeta}/\delta'')$, IMU on $\hat{\xi}\hat{\eta}\hat{\zeta}$ axes
	π_1	1
	ΔTR_{LMK}	1
EL_{LMK}	x , y , z	$\hat{\xi}/\sqrt{\delta''}$
	θ_1 , θ_2 , θ_3	$-\hat{\rho}_L \times \hat{\xi}/\sqrt{\delta''}$, IMU on XYZ axes
	θ_1 , θ_2 , θ_3	$(0, -\rho_{L\zeta}/\sqrt{\delta''}, \rho_{L\eta}/\sqrt{\delta''})$, IMU on $\hat{\xi}\hat{\eta}\hat{\zeta}$ axes
	π_2	1
	ΔEL_{LMK}	1
RA	x , y , z	$\bar{R}'/R$
	ΔRA	1

NOTE: Auxiliary quantities for Table I are given in Figures 1-5 and as follows:

$R = |\bar{R}|$

$\bar{V}_0 = \bar{\omega}_E \times \bar{R}_0$

ω_E = angular velocity of earth

$\rho_0 = |\bar{\rho}_0|$

$\dot{\rho}_0 = \bar{\omega}_E \times \hat{\rho}_0 + (\bar{V} - \bar{V}_0) \cdot \hat{\rho}_0 \hat{\rho}_0/\rho_0 - (\bar{V} - \bar{V}_0)/\rho_0$

$K_1 = \tan \lambda_0 / R_0 \rho_{en}^2$

$R_0 = |\bar{R}_0|$

$\rho_e = \hat{e} \cdot \bar{\rho}_0$

$\dot{\rho}_n = \hat{n} \cdot \bar{\rho}_0$

$\rho_{en}^2 = \rho_e^2 + \rho_n^2$

$\rho_h = \hat{h} \cdot \bar{\rho}_0$

R_{EH} = distance from center of earth to horizon

$\rho_{s\xi} = \hat{\xi} \cdot \hat{\rho}_s$

$\rho_{s\eta} = \hat{\eta} \cdot \hat{\rho}_s$

$\rho_{s\zeta} = \hat{\zeta} \cdot \hat{\rho}_s$

$\delta' = \rho_{s\zeta}^2 + \rho_{s\eta}^2$

$\hat{\rho}_L = \bar{\rho}_L/|\bar{\rho}_L|$

$\rho_{L\xi} = \hat{\rho}_L \cdot \hat{\xi}$

$\rho_{L\eta} = \hat{\rho}_L \cdot \hat{\eta}$

$\rho_{L\zeta} = \hat{\rho}_L \cdot \hat{\zeta}$

$\delta'' = \rho_{L\eta}^2 + \rho_{L\zeta}^2$

TABLE II

List of State Variables (Excluding Biases)

X	vehicle X inertial position error
Y	vehicle Y inertial position error
Z	vehicle Z inertial position error
$\dot{X}$	vehicle inertial velocity error along X axis
$\dot{Y}$	vehicle inertial velocity error along Y axis
$\dot{Z}$	vehicle inertial velocity error along Z axis
θ_1	gyro 1 angular error, due to drift
θ_2	gyro 2 angular error, due to drift
θ_3	gyro 3 angular error, due to drift
ω_1	gyro 1 random drift rate, time correlated
ω_2	gyro 2 random drift rate, time correlated
ω_3	gyro 3 random drift rate, time correlated
ω_{B1}	gyro 1 drift rate bias
ω_{B2}	gyro 2 drift rate bias
ω_{B3}	gyro 3 drift rate bias
$\tilde{h}_1$	horizon phenomenological deviation affecting altitude
$\tilde{h}_2$	horizon phenomenological deviation affecting vertical
$\tilde{h}_3$	horizon phenomenological deviation affecting vertical
$\tilde{\phi}_1$	horizon sensor time correlated error in $\tilde{\Phi}_1$
$\tilde{\phi}_2$	horizon sensor time correlated error in $\tilde{\Phi}_2$
$\tilde{\phi}_3$	horizon sensor time correlated error in $\tilde{\Phi}_3$
Ω_1	star tracker time correlated error in TR_S
Ω_2	star tracker time correlated error in EL_S
π_1	landmark sensor time correlated error in TR_L
π_2	landmark sensor time correlated error in EL_L
$\delta\mu$	bias in earth's gravity, μ
$\delta C_{N(1)}$ ⋮ $\delta C_{N(4)}$	biases in zonal harmonics of gravitational potential, $2 \leq N(1), N(2), N(3), N(4)$
$\delta D_0/D_0$	bias in drag parameter, D_0

TABLE III

List of Bias State Variables

Δe	ground radar station location error, easterly
Δn	ground radar station location error, northerly
Δh	ground radar station location error, vertical
$\Delta\rho$	ground radar range bias
$\Delta\dot{\rho}$	ground radar range rate bias
ΔA	ground radar azimuth bias
ΔE	ground radar elevation bias
$\Delta\tilde{\Phi}_1$	horizon tracker stadimetric angle bias
$\Delta\tilde{\Phi}_2$	horizon tracker vertical angle bias on $\hat{\xi}$ axis
$\Delta\tilde{\Phi}_3$	horizon tracker vertical angle bias on $\hat{\eta}$ axis
$\Delta\tilde{H}_1$	horizon deviation bias, affecting stadimetric altitude
ΔTK_S	star-IMU train angle bias
ΔEL_S	star-IMU elevation angle bias
ΔTK_L	landmark-IMU train angle bias
ΔEL_L	landmark-IMU elevation angle bias
ΔRA	radar altimeter bias

TABLE IV

Observations for Unidentified Landmark Navigation

Instrument readout	*True observation,* Y_T	*Estimated observation,* Y_E
Star tracker α_i, δ_i $(i = 1, 2, 3)$	$(Y)_T = (S)$. where $(S) = (\hat{S}_{I1}, \hat{S}_{I2}, \hat{S}_{I3})(\hat{S}_{V1}, \hat{S}_{V2}, \hat{S}_{V3})^{-1}$ $\hat{S}_{Vi} = (\cos\delta_i \cos\alpha_i, \cos\delta_i \sin\alpha_i, \sin\delta_i)^T$, $i = 1, 2, 3$	$(Y)_E = (S)_E$
Landmark sensor T, D	Vector observation: $\hat{\rho}_T = (S)_R(\rho_V)$, where $\rho_V = (\cos D \cos T, \cos D \sin T, \sin D)^T$	$\hat{\rho}_E = \dfrac{L - R}{\lvert L - R \rvert}$
Landmark sensor T	Scalar observations: $Y_T = T$	$Y_E = \tan^{-1} \dfrac{(L - \bar{R}) \cdot \hat{S}_2}{(L - \bar{R}) \cdot \hat{S}_1}$
D	$Y_T = D$	$Y_E = \sin^{-1} \dfrac{(L - R) \cdot \hat{S}_3}{\lvert L - R \rvert}$

NOTES:
- $\alpha_i \delta_i$ = azimuth and elevation of i^{th} star relative to vehicle frame, v
- $\hat{S}_{Ii}$ = unit column vector of i^{th} star relative to inertial space
- (S) = matrix transformation from vehicle frame to inertial space = $(\hat{S}_1, \hat{S}_2, \hat{S}_3)$
- T = train angle of landmark, measured about vehicle S_1 axis
- D = depression angle of landmark, measured normal to vehicle $S_2 S_3$ plane
- $(S)_E$ = (S) estimated
- T = transpose, when used as a superscript

TABLE V

Measurement Matrices for Unidentified Landmark Navigation

Observation	*Measurement matrix*
$\hat{Y}_E = \hat{\rho}$	$M = \left[\frac{-1}{\rho}(I-\hat{\rho}\hat{\rho}),\ \bar{0},\ \frac{1}{\rho}(I-\hat{\rho}\hat{\rho})\right]_{3\times 9}$
Scalar observations:	
$Y_E = \tan^{-1}\frac{\bar{\rho}\cdot\hat{S}_1}{\bar{\rho}\cdot\hat{S}_2}$	$M = \left[-(\rho_2\hat{S}_1-\rho_1\hat{S}_2)^T/(\rho_2^2+\rho_1^2),\ \bar{0},\ (\rho_2\hat{S}_3-\rho_3\hat{S}_2)^T/(\rho_2^2+\rho_3^2),\bar{0}\right]$
$Y_E = \sin^{-1}\frac{\hat{\rho}\cdot\hat{S}_1}{\rho}$	$M = \left[-(\hat{S}_1-\hat{\rho}\rho_1/\rho)^T/(\rho_2^2+\rho_3^2)^{1/2},\ \bar{0},\ (\hat{S}_1-\hat{\rho}\rho_1/\rho)^T/(\rho_2^2+\rho_3^2)^{1/2}\right] = 1\times 9$

NOTES: $\bar{\rho} = \bar{L}-\bar{R}$, $\hat{\rho}=\bar{\rho}/\rho$, $\rho=|\bar{\rho}|$

$\hat{S}_1\hat{S}_2\hat{S}_3$ = unit vectors along vehicle axes 1,2,3

$\bar{L}$ = estimated landmark location vector

$\bar{R}$ = estimated vehicle position vector

ρ_1,ρ_2,ρ_3 = components of $\hat{\rho}$ along vehicle axes 1,2,3

$\bar{0}$ = (0,0,0)

TABLE VI

Conditions for Computer Runs of Figures 14 and 15

Orbit: 100-nautical-mile circular lunar orbit

Gravitational parameter	1.7313×10^{14} ft^3/sec^2
Semi-major axis	1038.5 nautical miles
Eccentricity	0.0
Inclination to lunar equator	0.0 deg
Start time, t_0	0.0 hr

Measurements: Unit vector from vehicle to landmark

Time of first measurement	0.258 hr
Time between measurements	0.003 hr
Time of last measurement	0.999 hr
Landmark sensor white-noise	0.001 rad, r.m.s., actual 0.001 rad, r.m.s., *a priori*

Landmark estimation method: one-sight

Terrain altitude variation	0.250 nautical mile, r.m.s., actual 0.250 nautical mile, r.m.s., *a priori*

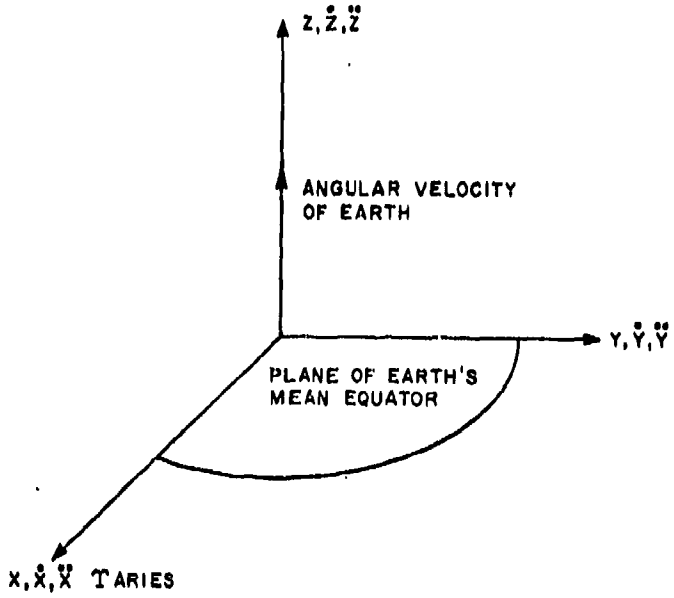
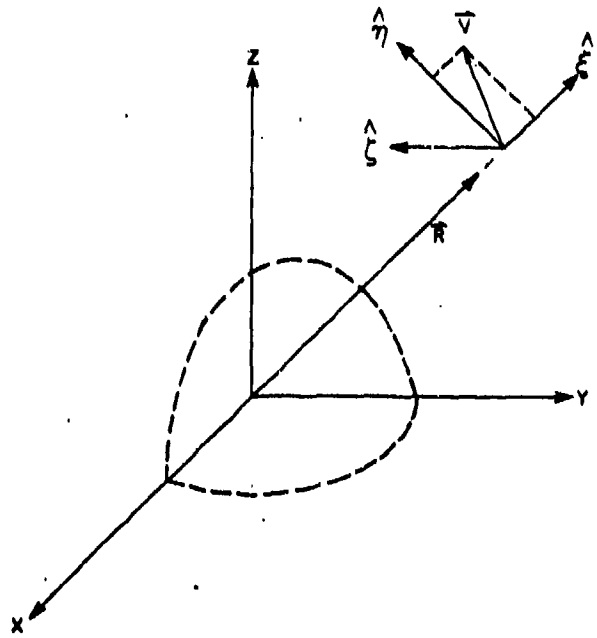
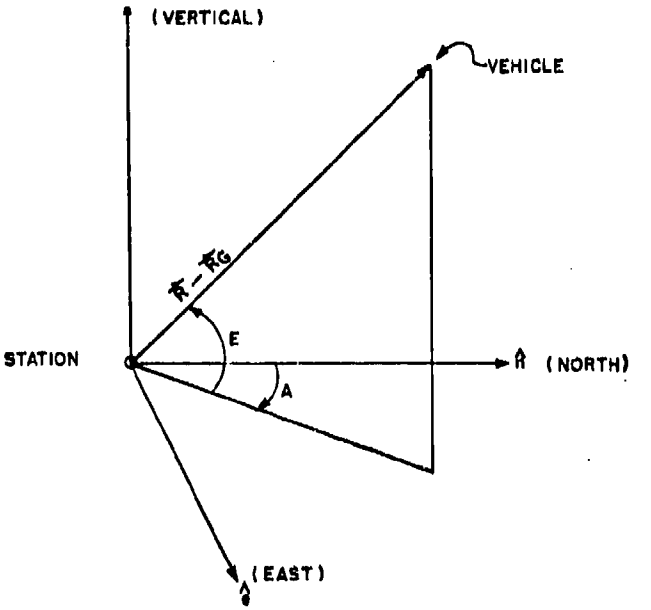


Fig. 1 Inertial coordinates



$\vec{R}$: POSITION VECTOR OF VEHICLE
$\vec{V}$: VELOCITY VECTOR OF VEHICLE
$\hat{\xi}$: UNIT VECTOR PARALLEL TO $\vec{R}$
$\hat{\zeta}$: UNIT VECTOR PARALLEL TO $\vec{R} \times \vec{V}$
$\hat{\eta}$: UNIT VECTOR PARALLEL TO $\hat{\zeta} \times \hat{\xi}$

Fig. 2 Vehicle geocentric vertical coordinates $\hat{\xi}\,\hat{\eta}\,\hat{\zeta}$



$\hat{e}, \hat{n}, \hat{r}$ = UNIT VECTORS EAST, NORTH, AND VERTICAL AT THE STATION

$\vec{R}$ = VECTOR FROM CENTER OF EARTH TO VEHICLE, = (X,Y,Z)

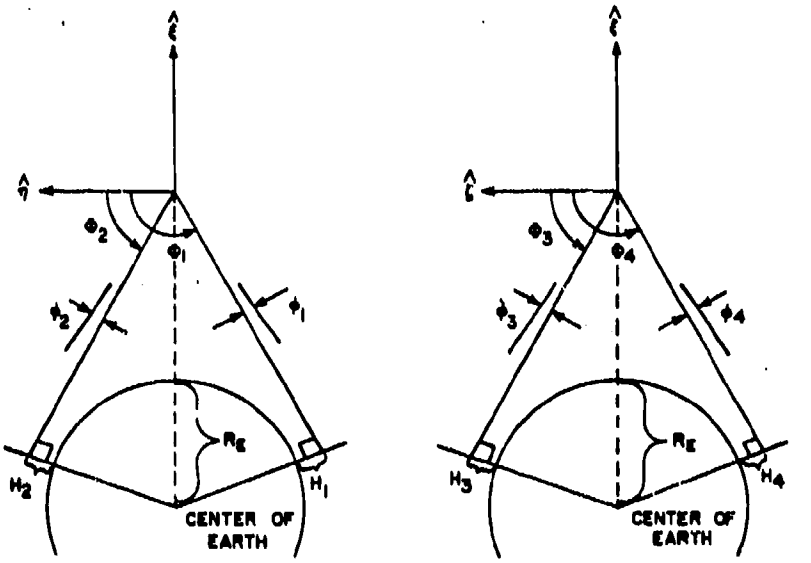
$\vec{V}$ = VELOCITY OF VEHICLE RELATIVE TO CENTER OF EARTH = $(\dot{X},\dot{Y},\dot{Z})$

$\vec{R}_G$ = VECTOR FROM CENTER OF EARTH TO STATION, = (X_G,Y_G,Z_G)

$\vec{\rho}_G$ = VECTOR FROM STATION TO VLHICLE, = $\vec{R}-\vec{R}_G$

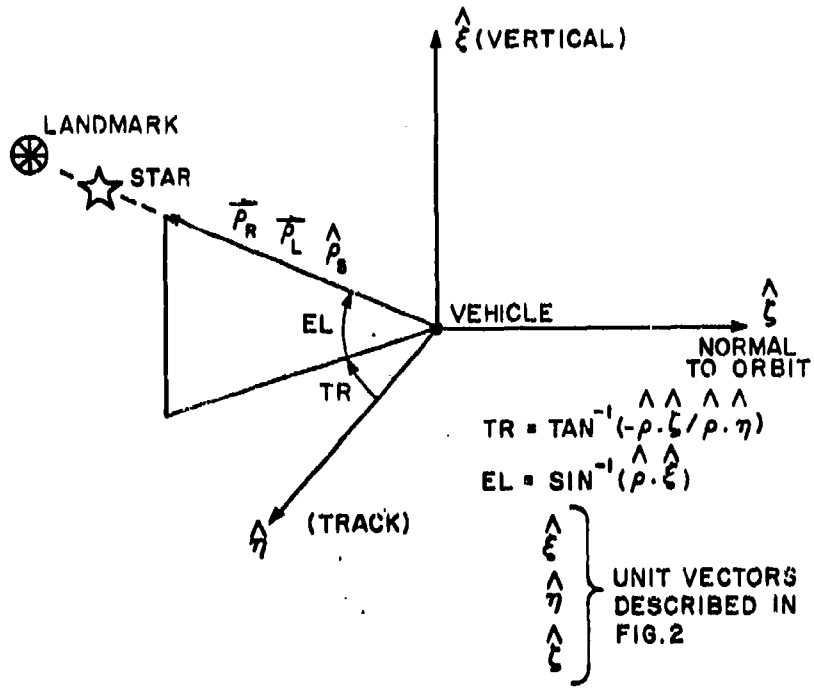
λ_G = STATION GEODETIC LATITUDE

Fig. 3 Ground radar geometry



VIEW IN ORBITAL PLANE VIEW NORMAL TO ORBITAL PLANE

Fig. 4 Horizon and horizon tracker geometry



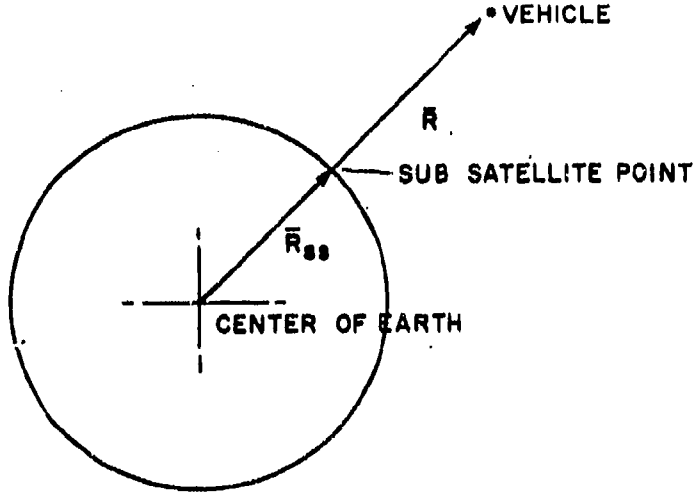
$\vec{\rho}_L$ = VECTOR FROM VEHICLE TO LANDMARK

$\hat{\rho}_s$ = UNIT VECTOR FROM VEHICLE TO STAR

TR_L, EL_L = TRAIN ANGLE, ELEVATION ANGLE TO LANDMARK

TR_s, EL_s = TRAIN ANGLE, ELEVATION ANGLE TO STAR.

Fig. 5 Landmark sensor, star tracker geometry



$RA = |\bar{R} - \bar{R}_{ss}|$

Fig. 6 Geometry of radar altimeter measurement

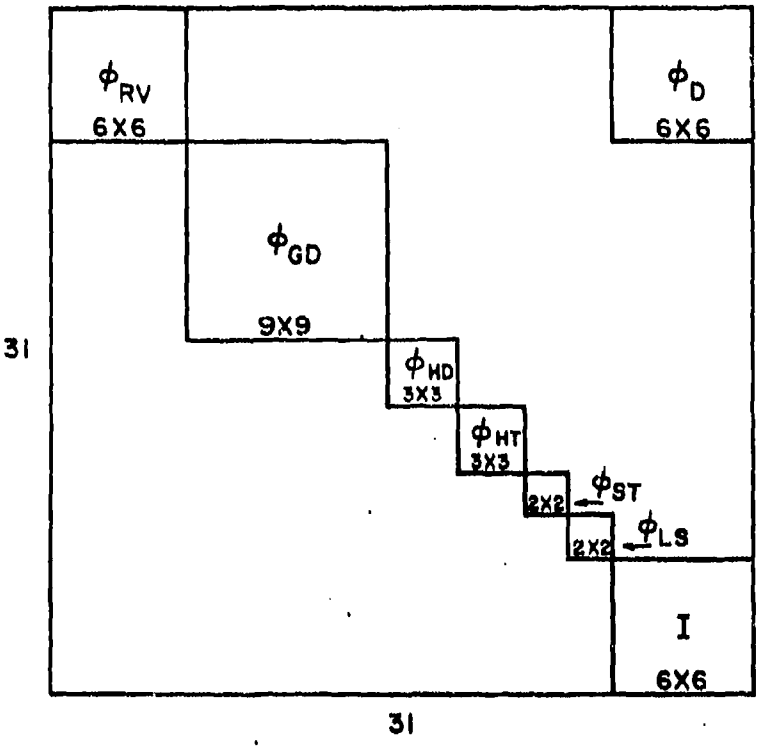


Fig.7 Partitioning of ϕ_u

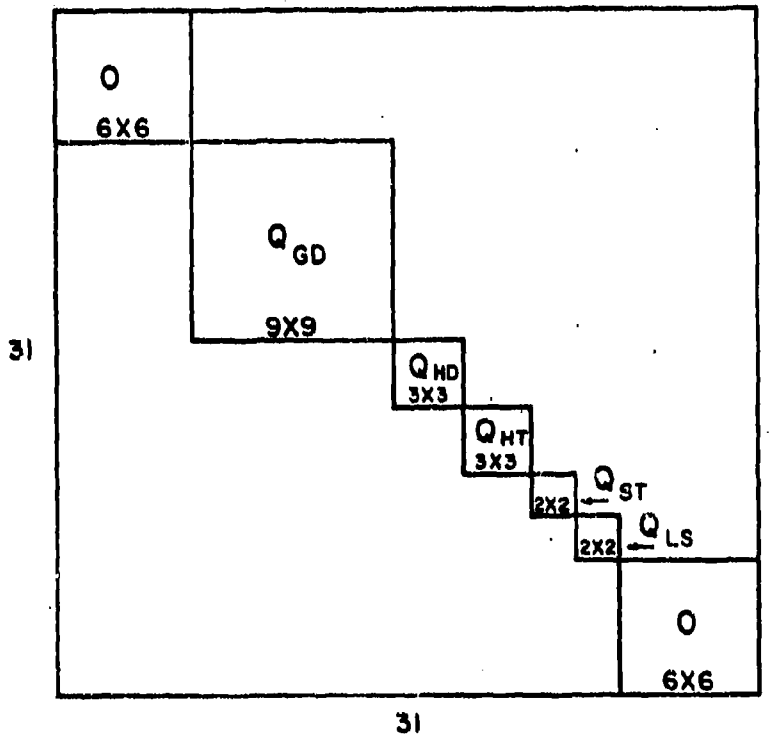


Fig.8 Partitioning of Q_u

FIG. 9A

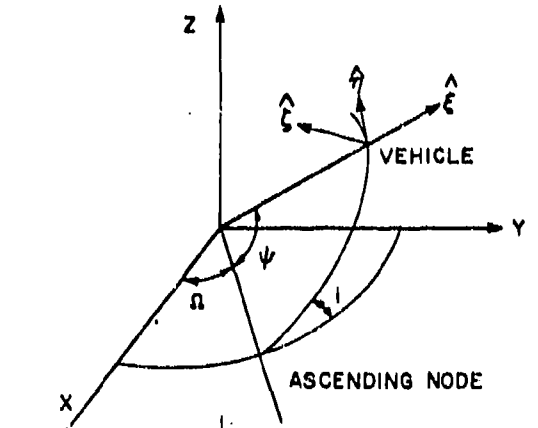


FIG. 9B

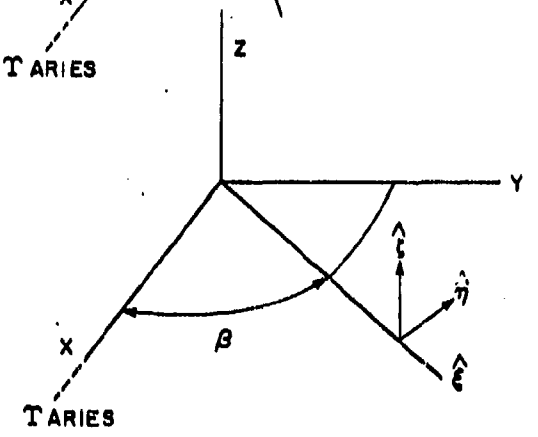


Fig. 9 Geometry of local vertical to inertial transformation

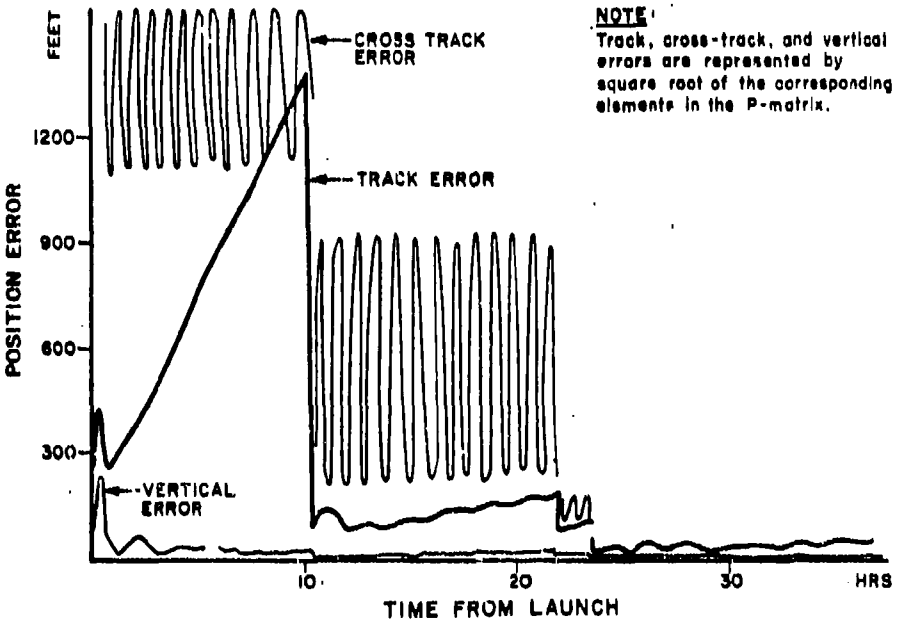


Fig. 10 Typical navigation error plot for both ground and on-board observations

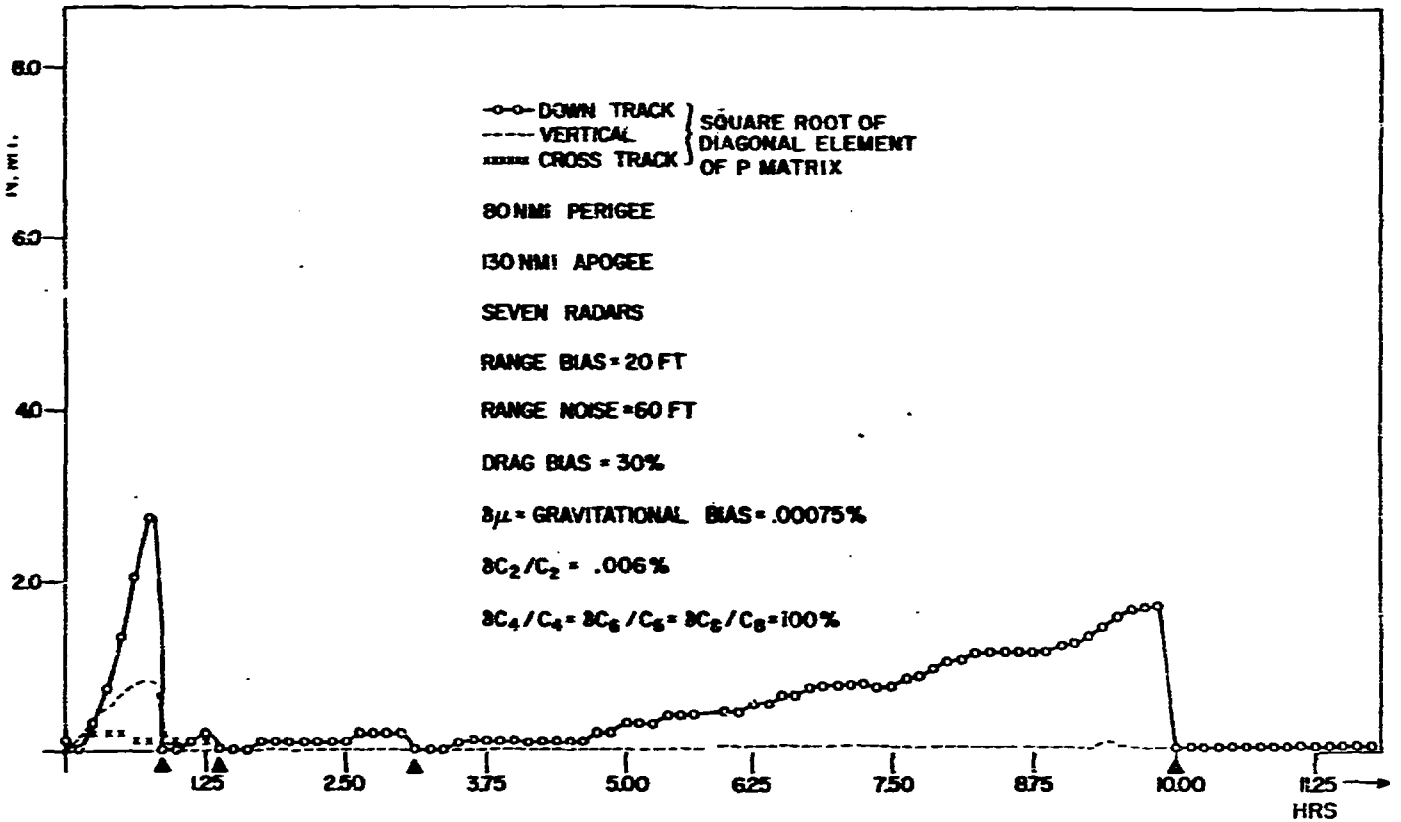


Fig. 11 Effect of radar range observations on position, assuming biases in drag, oblateness, and gravity

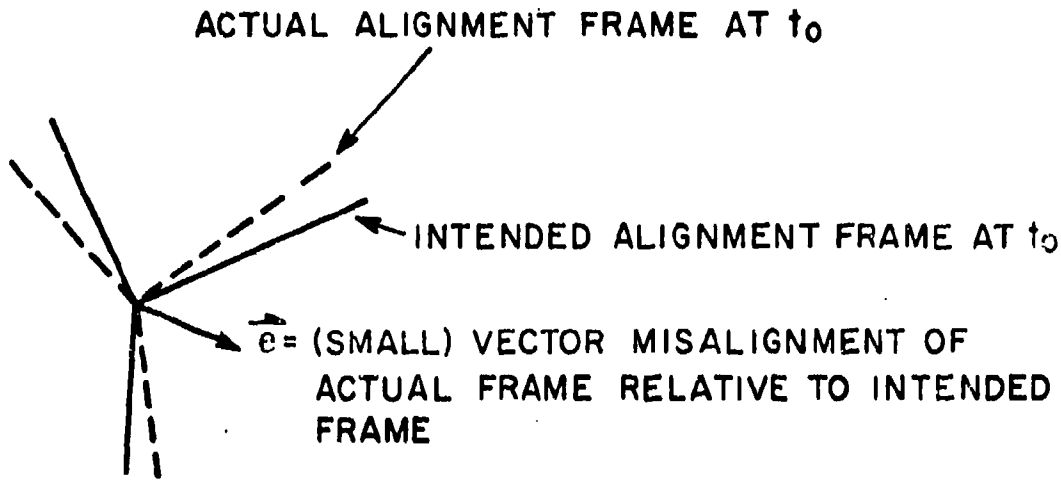


Fig. 12 Initial alignment frames of an inertial navigator

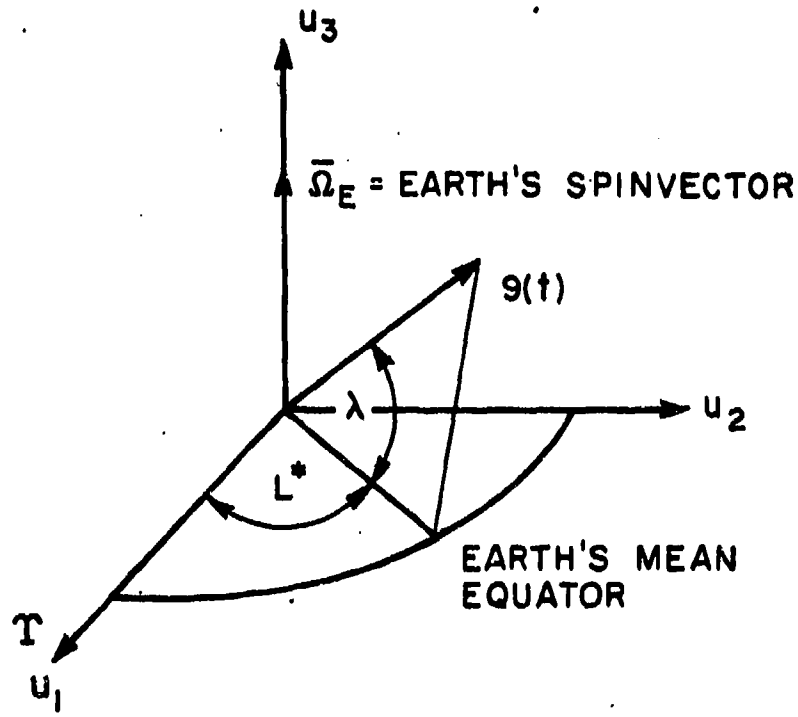


Fig. 13 Geometry of self-contained inertial alignment

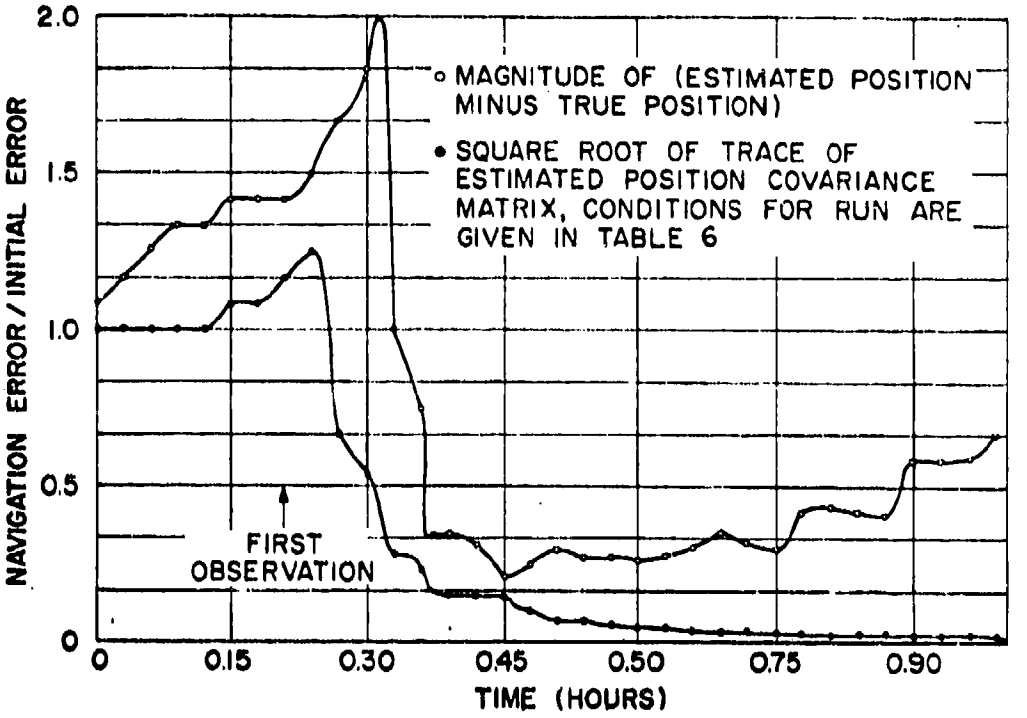


Fig. 14 Actual and estimated position error for diagonal initialization of covariance matrix at each landmark, lunar orbit

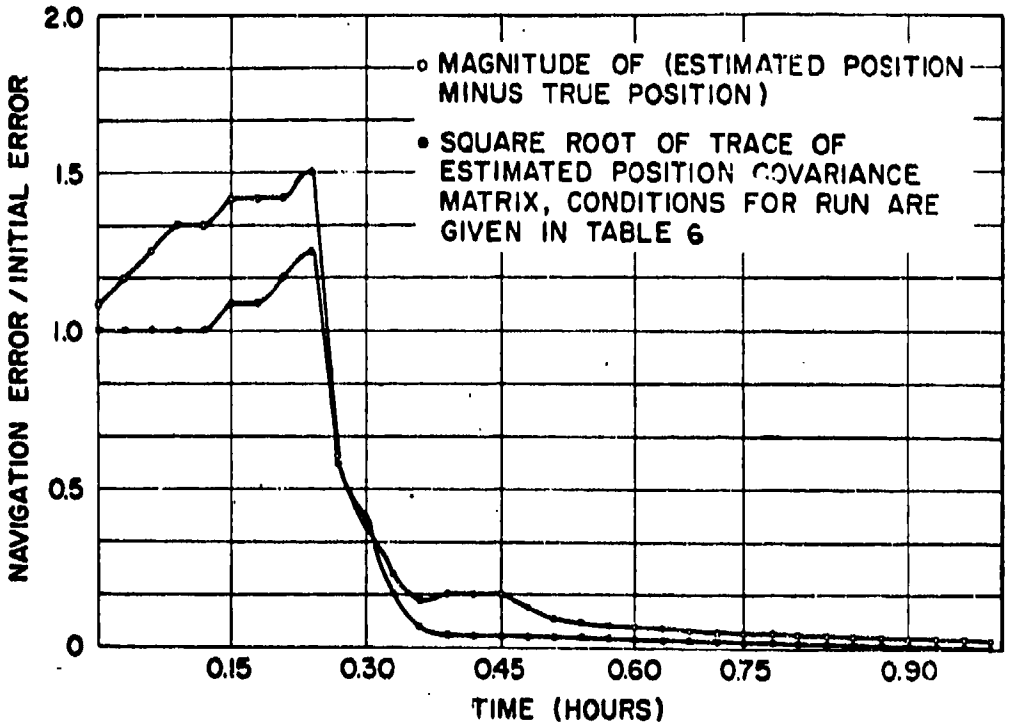


Fig. 15 Actual and estimated position error for complete initialization of covariance matrix at each landmark, lunar orbit

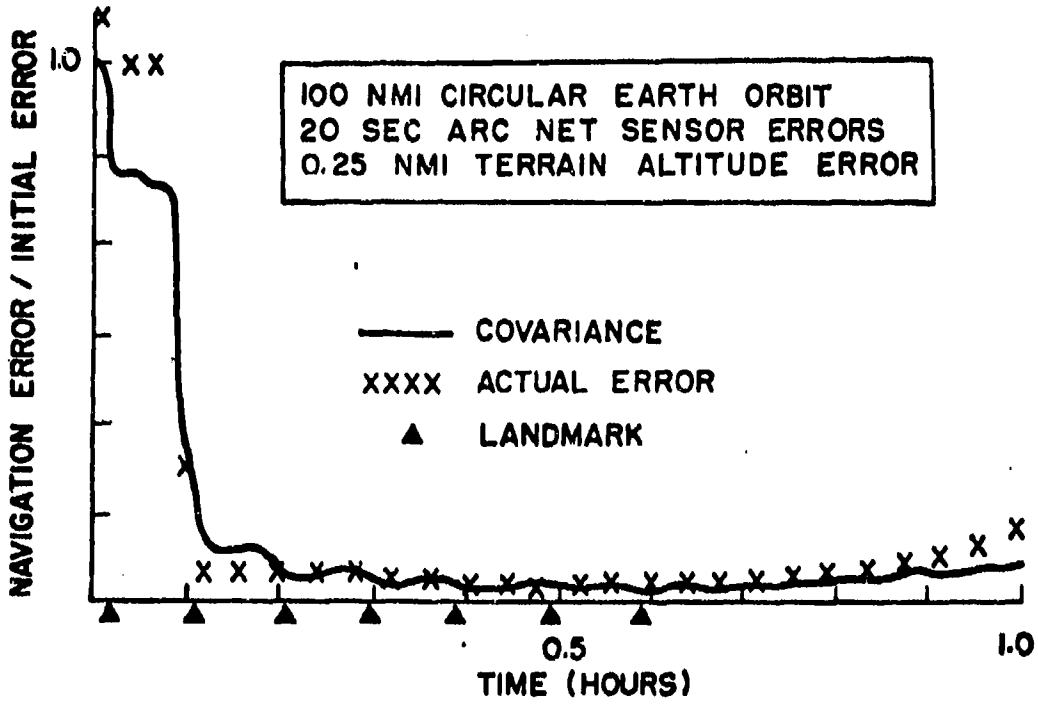


Fig. 16 Actual and estimated position error for complete initialization of covariance matrix at each landmark, earth orbit

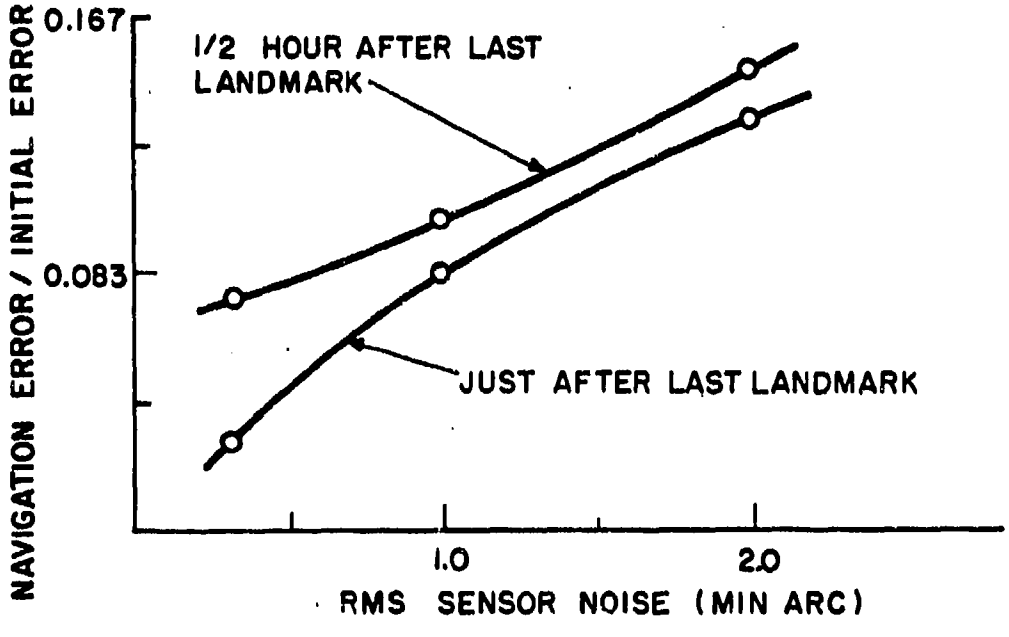


Fig. 17 Accuracy against sensor noise, earth orbit

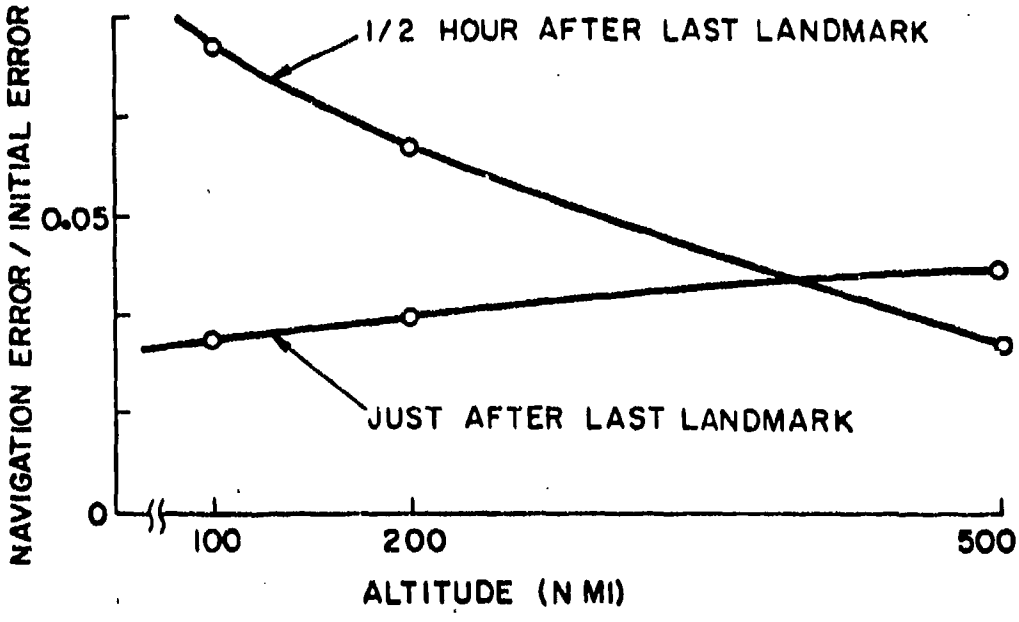


Fig. 18 Accuracy against altitude, earth orbit

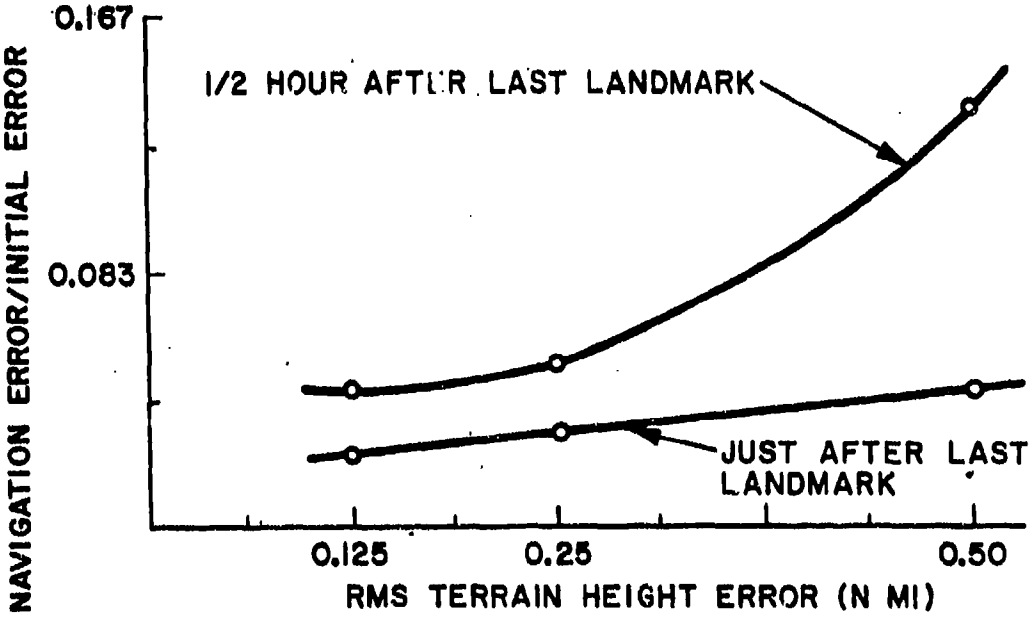


Fig. 19 Accuracy against terrain errors, earth orbit

CHAPTER 16

SECTION I – APPLICATION OF KALMAN FILTERING FOR THE ALIGNMENT OF CARRIER AIRCRAFT INERTIAL NAVIGATION SYSTEMS

by

James T.Kouba

Litton Systems, Inc.,
Woodland Hills, California, USA

CHAPTER 16

SECTION I - APPLICATION OF KALMAN FILTERING FOR THE ALIGNMENT OF CARRIER AIRCRAFT INERTIAL NAVIGATION SYSTEMS

James T. Kouba

1. INTRODUCTION

Because missions generally flown by military aircraft require covertness in a hostile environment, the need for a self-contained passive navigation system has long been recognized. Being self-contained and passive means respectively that the system does not have to rely on any external equipment such as LORAN, Omega, etc., nor radiate to perform the navigation function. Since covertness is not necessarily required for all phases of a mission, the military navigation system generally comprises of a multitude of subsystems, including active and non-self-contained elements, so that maximum use is made of all available information. However the system invariably has and requires the inherent capability of performing the navigation function in a self-contained, passive manner.

This requirement is emphatic for aircraft operating from an aircraft carrier. The military functions performed by these aircraft require flight over water for extended periods of time without the benefit of external aids to navigation and often in a hostile environment. Additionally these aircraft are required to return to the aircraft carrier, which may be concealed by weather and electromagnetic silence. Without a self-contained, passive navigation system these naval tactics would be impossible.

The only navigation system satisfying these requirements is the inertial navigation system. It provides in a self-contained, passive manner all the required cruise navigation information. It is, however, recognized that certain functions to be performed will require contact in one form or another with the outside world, so that the inertial navigator cannot be the only element of the navigation system.

Since the performance of an inertial navigator is highly correlated with the initialization accuracy, it is imperative that a highly reliable and accurate initialization process be devised. The recognition of this problem and the need for its solution as applied to carrier aircraft has, in the past few years, led to intensive research and analysis to resolve the multitude of operational and performance problems associated with it. The following sections discuss and analyze the carrier aircraft initialization process and logically develop the requirements. Because of the classified nature of the problem, the material will be limited to a qualitative description of the processes involved. The actual quantitative performance requirements and the details of the solutions are omitted.

Section 2 is a brief review of the essential features of inertial navigation and includes a generalized mechanization diagram for four different local level mechanizations. An essential ingredient in the design of a Kalman filter is the error model of the system being controlled. To this end Section 3 presents a generalized error model for the four mechanizations described in Section 2. Section 4 defines the initialization process and Sections 5, 6, and 7 discuss the multitude of problems encountered in carrier deck, catapult, and in-flight alignment. Finally the chapter concludes by illustrating the various features of a Kalman filter for carrier deck alignment.

2. THE INERTIAL NAVIGATION SYSTEM

The essential elements of the inertial navigation system are the instrument cluster (platform) (consisting of gyros and accelerometers) that is gimbaled to the airframe and a digital computer. The instrument cluster's geographic orientation is controlled by torquing signals supplied by the computer, which in turn accepts the outputs of the accelerometers and computes velocity and position.

More specifically consider a right-handed orthogonal coordinate system (x, y, z) associated with, and rigidly attached to, the instrument cluster. The accelerometers and gyros are arranged so that

(i) the positive sensitive axes of the accelerometers coincide with the x , y , z axes,

(ii) torquing signals applied to the gyros will selectively produce angular rates of the platform about the x , y , z axes.

For aircraft inertial navigation systems a local-level mechanization is generally employed. In this mechanization the platform is controlled, ideally, via its gyros so that its z axis is always collinear with the local perpendicular of the reference geoid, independent of the motion of the aircraft. There ar four different types of local-level mechanizations that can be considered. One of these, called north-pointing, is such that the y axis of the platform ideally always points north (x axis points east and z up). For the remaining three mechanizations, called wander azimuth, the y axis is not constrained to remain north but is permitted to "wander" at a

controlled rate relative to north. The computer in turn computes the wander angle by essentially (not directly) integrating the wander angle rate. Actually the wander angle, latitude, and longitude are all implicitly computed by solving a set of differential equations for a set of direction cosines.

Comparatively the north-pointing mechanization is considerably simpler than the wander azimuth mechanizations. However it does have the disadvantage of not having a polar capability. This results from the fact that, as the navigation system approaches the pole, the azimuth gyro (z gyro) torquing rate required to maintain the platform north-pointing becomes excessively large. Such is not the case for the wander azimuth mechanization since the platform is not constrained geographically.

In the sequel it will be shown that the wander azimuth mechanization is particularly advantageous for the solution of the alignment problem while the aircraft is on the flight deck of the aircraft carrier. This results principally because the wander azimuth platform can initially be at any angle relative to north and the computer's wander angle adjusted to agree with it.

A complete, detailed derivation of these four local-level mechanizations can be found elsewhere[1]. For future reference Figure 1 diagrammatically presents a generalized mechanization block diagram for local-level free-inertial navigation systems. Although it theoretically represents all four local-level mechanizations, in actual practice many variations are used. This is especially true for the north-pointing mechanization wherein many simplifications arise and, instead of computing direction cosines, latitude and longitude are obtained directly. For simplicity the vertical channel mechanization has been omitted. Table I lists the definitions of the symbols used in Figure 1, and Table II lists the vertical platform angular rates for the various mechanizations. It should be noticed that for any of the wander azimuth mechanizations the explicit reduction of the direction cosines to latitude, longitude, and wander angle is required for display purposes only.

The next section presents an error model for these mechanizations.

3. THE INERTIAL ERROR MODEL

The object of the initialization process is to remove, insofar as possible, the deviation of the actual system from an ideal, error-free system. In general the estimation of the deviation is made by comparing (subtracting) a set of reference variables with (from) a set of commensurate variables computed by the actual system and inferring from these all deviations of interest. The inference process requires a model relating the quantities that are observed with all quantities of interest.

Adequate estimate and control of the initialization process must take into account both the deviation of reference variables and the system variables from the error-free system variables. Thus a necessary requirement for the design of an alignment technique is a model of the error propagations of both the inertial system and the reference system, i.e. the differential and algebraic equations governing the deviations of the actual systems from the ideal, error-free systems.

The reference system will vary, depending upon what is available and desirable to use at the time the initialization process is being invoked. For example, if the initialization process is carried out while the aircraft is on the ground, no external information is required except initial position, and the initialization process is carried out in a completely self-contained manner. This is the case because the *a priori* knowledge that the aircraft is stationary on the ground implies that the velocity is zero (to within the small oscillations caused by wind buffeting) so that the reference variables are chosen to be the initial manually inserted position and zero velocity. In contrast, if the initialization process is carried out on the deck of an aircraft carrier, the velocity and position of the aircraft will vary as a function of time because of the ship's motion. In this case a system external to the aircraft is required so that adequate information is available to describe the motion of the aircraft system.

Of singular interest here is the error model for the aircraft inertial system, since it is required in all cases. Figure 2 displays a generalized error block diagram for the local-level free-inertial navigation system. It is a linear error model derived by essentially subtracting the error-free system equations from the actual system equation and retaining only first-order terms. For simplicity the vertical channel has been omitted. As can be seen, all four local-level mechanizations are included in this figure. And, as might be expected, various simplications can result by choosing one particular mechanization of interest; e.g., for the free azimuth mechanization, $\rho_z = 0$. Also it should be noted that $\delta\rho_z$ can be eliminated from the error model, since the errors ϕ_z and $\delta\theta_z$ always appear in the form $\phi_z - \delta\theta_z = \psi_z$ as inputs into the various nodes.

At times it may be of interest to know the latitude, longitude, and wander angle errors themselves. For nonpolar regions they are given by

$$\delta\lambda = \lambda_C - \lambda_T = (\delta\theta_x \sin\alpha + \delta\theta_y \cos\alpha)\sec\phi \tag{3.1}$$

$$\delta\phi = \phi_C - \phi_T = \delta\theta_y \sin\alpha - \delta\theta_x \cos\alpha \tag{3.2}$$

$$\delta\alpha = \alpha_C - \alpha_T = \delta\theta_z - \delta\lambda \sin\phi\,, \tag{3.3}$$

when the subscript C means computed and the subscript T means true (error free). For polar regions the relationships become nonlinear functions of $\delta\theta_x$, $\delta\theta_y$ and $\delta\theta_z$ because $\delta\lambda$ and $\delta\alpha$ do not remain small.

In the derivation of this error model, which is not presented here[2], three coordinate systems are used:

(i) True, defined by the error-free system.

(ii) Platform, defined by the actual instrument cluster.

(iii) Computer, defined by the actual computed values of the direction cosines.

To maintain a linear error model, the angular separation between these coordinate systems must remain small (less than 10 degrees).

The angular separation between the true and the computer coordinate systems is called the position error, $\delta\theta_x$, $\delta\theta_y$, and $\delta\theta_z$. That is the position errors (angular) are the small angular rotations $\delta\theta_x$, $\delta\theta_y$, and $\delta\theta_z$ about the true x, y, z axes necessary to bring the true axes into coincidence with the computer axes.

The angular separation between the true and the platform coordinate system is called the platform tilt error, ϕ_x, ϕ_y, and ϕ_z. That is the tilt errors are the small angular rotations ϕ_x, ϕ_y, and ϕ_z about the true x, y, z axes necessary to bring the true axes into coincidence with the platform axes.

The angular separation between the computer and the platform is defined by the angles Ψ_x, Ψ_y, Ψ_z, which are given by the components of the vector difference $\Psi = \phi - \delta\theta$. That is the computer-platform errors are the small angular rotations, Ψ_x, Ψ_y, Ψ_z, about the computer x, y, z axes necessary to bring the computer axes into coincidence with the platform axes.

The velocity error and the platform angular rate error are defined by the vector difference between the computed values and the true values, i.e.,

$$\underline{\delta v} = \underline{v}_C - \underline{v}_T \tag{3.4}$$

$$\underline{\delta \rho} = \underline{\rho}_C - \underline{\rho}_T \; . \tag{3.5}$$

Table III lists the definitions of the remaining variables appearing in Figure 2.

4. THE INITIALIZATION PROCESS

The initialization process for the inertial navigation system requires both the initialization of the mechanization differential equations with the proper velocity, position, and wander angle and the physical alignment of the platform to the required geographically referenced orientation. Each type of mechanization considered here requires (ideally) that the platform be aligned so that its z axis coincides with the local perpendicular of the reference geoid. The azimuth orientation for north-pointing systems requires that the y axis of the platform point north; whereas for the wander azimuth systems a dichotomy exists in that the platform azimuth must agree with the computed wander azimuth angle or vice versa.

In principle the initialization process could occur before, during, or after the aircraft is catapulted from the carrier. In each of these cases the aircraft is in motion relative to the earth, so that some internal or external (to the aircraft system) reference sensor is required. Without any further definition or restriction on the availability of reference sensors, each case can have a multitude of different possible solutions depending upon the reference sensors chosen. To limit the ensuing discussions it is therefore necessary to limit the choice of reference sensors considered. For alignment while on the carrier deck the reference sensor chosen is the ship's inertial navigation system (SINS). Alignment during catapulting is preceded by position and velocity initialization and fine leveling, using SINS data, and employs ship heading or transponder ranging to perform the azimuth alignment while catapulting. For in-flight alignment the reference considered is the Doppler radar velocity sensor. A discussion of each of these techniques is given in the following sections.

Before examining these techniques it is necessary to discuss the quantitative meaning of alignment. Since in any case the initialization process, after a practical amount of time, will result in a residual deviation from the ideal (error-free) system, it is fair to ask when is a system considered to be aligned? To answer this question it must be realized that what is actually of interest is the accuracy of the subsequent navigation. Thus it is possible to quantitatively define alignment implicitly by specifying the accuracy of the subsequent navigation. However this is usually not done, and a redundant specification of both the alignment and navigation accuracy is made. Incompatibility does not present a problem, because the alignment accuracy is derived from the navigation accuracy; and there exists a high correlation between the two. Thus the system is considered to be aligned when it has met the quantitative specification dictated by the user.

5. CARRIER DECK ALIGNMENT

It is generally recognized that alignment while on the deck of the carrier will produce the most accurate ensuing navigation. However, because of the complexity of handling the multitude of aircraft on the flight deck,

severe restrictions are placed on the acceptable solutions to the alignment procedure and technique. Thus, in the design of an alignment technique while the aircraft is on the carrier deck, full account must be taken of the operational, environmental, and control constraints imposed.

Before the strict imposition of the operational constraints two techniques had been proposed that, for historical reasons, are worth mentioning. The first technique, called transfer alignment[3] employs a portable master inertial system physically and electrically mated with the aircraft system. The close proximity of the two platforms and the angular precision of the mating permit accurate gimbal matching for azimuth alignment and acceleration matching for level alignment. The second technique, called insertion[4], employs a portable aircraft inertial navigator (PINS) that is prealigned at a shipboard console. By including a battery and sufficient computational capability in the portable unit, it can be removed from the alignment console, carried to a waiting aircraft, and inserted, ready for operation.

As will be seen, the operational constraints imposed rule out these techniques as possible solutions and implicitly require that the alignment take place with the inertial navigator in the aircraft. The combination of the three aforementioned constraints determines the form and size of the error controller employed in the alignment technique. A discussion of each of these constraints and their ramifications follows.

5.1 Operational Constraints

To avoid interference with any of the normal deck-handling procedures such as fueling, arming, and taxiing, the alignment is to be initiated and completed without any manual intervention or restriction on the location and movement of aircraft. This succinctly summarizes the operational constraints. Two immediately obvious ramifications are that no power cable and no data cable can be used for the alignment technique. In turn, the lack of the power cable requires that the alignment be achieved during the epoch of engine turn-on to catapulting. Also, solutions involving special alignment stations on the carrier deck to afford sufficient time with engines on for the alignment to be completed are implicitly ruled out, i.e., no loitering for the express purpose of alignment is permitted since this would interfere with the normal deck-handling procedures.

Without an additional specification the time allotted for alignment is ambiguous. Therefore, after consideration of the normal movements of an aircraft subsequent to engine turn-on, an alignment time is specified (by the user) excluding any taxiing periods.

The imposition of a specific alignment time results in an instrument performance problem[5,6]. Both the accelerometers and gyros will function within specification if they are operated within a relatively narrow range about a design temperature. This normal operating temperature is, for present day instruments, considerably above expected ambient temperatures. Thus, when the platform is initially turned on (energizing the heaters), the instruments will cycle through a thermal transient until equilibrium is achieved. During this thermal transient, the instruments will exhibit transient drift and bias characteristics that are large enough and last long enough to degrade the alignment process.

Additionally other portions of the system will exhibit undesirable characteristics during this thermal transient phase. The cluster structure may exhibit excessive warpage due to thermal gradients and this may result in instrument non-orthogonalities and consequently degrade alignment. Also the multitude of electronic packages required to perform such functions as pulse shaping and scaling for gyro torquing, accelerometer proof mass restoration, and formation of the velocity output pulses, may exhibit abnormal behavior during the thermal transient.

The solution to these problems requires a careful mechanical and thermal design of the various instruments, structures and electronic packages involved so that the resulting undesirable thermal transient behavior, if not reduced to a negligible value, is repeatable and predictable as a function of suitably measured temperatures.

Since the aircraft system requires the SINS data to accomplish alignment and no data cable is permitted, it is necessary to resort to a telemetering system to obtain the required information. This leads to a problem relating to the existence of an electromagnetic radiator on the aircraft carrier. It is obvious that at times it would be deemed necessary to maintain the carrier electromagnetically silent. If these times of silence coincide with the times that alignment is to be performed, an incompatibility arises. The solution to this problem is one of more specifically defining what is meant by electromagnetic silence. Electromagnetic silence is desirable to avoid the possibility of the carrier revealing its presence to an enemy. Certainly, if the radiation pattern and power are so designed that, for all practical purposes, it is detectable only on the carrier deck, use of an electromagnetic radiator during a period of silence will not endanger the carrier. Thus appropriate achievable design constraints on the telemetering system will increase its utility to cover all times.

Additional ramifications do result because of the operational constraints imposed. However, they result because of the coupling of the operational, environmental and control constraints and will therefore be discussed as they arise.

5.2 Environmental Constraints

Carrier deck alignment is to be performed while the carrier is pitching, rolling, yawing, heaving, surging, and swaying. In addition to these motions the carrier may be making heading changes to perform evasive maneuvers or to head into the wind in anticipation of launching or receiving aircraft. Additionally the aircraft itself will be buffeted by winds over the flight deck, thus causing motion that is not sensed by SINS. The design of the alignment technique requires that each of these environmental conditions be taken into account.

The principal effect of these motions is related to the reference data used for the alignment process. The reference data supplied by the SINS to the aircraft system via the data link consist in part of position and velocity, as measured by the SINS itself. Since the aircraft may be situated anywhere on the flight deck, the SINS reference data must be compensated for the relative motion before they can be adequately used as alignment reference data. The data required to perform this compensation consist of the vector distance between the SINS and the aircraft plus the carrier's pitch, roll and heading, and their rates. (See Figure 3.) Since the SINS inherently provides these angles and angular rates it is reasonable to require that these be included in the data telemetered to the aircraft. On the other hand the distance between the SINS and the aircraft will vary from aircraft to aircraft and even change for a given aircraft if, during alignment, it taxies to a new deck position. Prohibition of the manual insertion of lever arm components, a restriction imposed by the operational constraints, means that the alignment technique must include a method for the estimation of the lever arm components.

3.3 Control Constraints

The control constraints are concerned with the physical limitations of the instruments involved and the limitations imposed by the present state of development of the control theory to be used to perform the alignment.

The primary physical limitation of the instruments that affects the alignment problem is the maximum torquing rate that can be applied to the gyros. In reality a dual phenomenon is of concern. First there exists an absolute maximum torquing rate that, if exceeded, will essentially destroy the gyro torquing coils. Second, as this maximum is approached from the lower side, the accuracy with which the desired angular rate is achieved decreases, i.e., torquing nonlinearities become significant.

In the application of modern control theory to the carrier deck alignment problem the limitation that significantly affects the solution is the imposition of linearity. The reason for this will become apparent in the following discussion.

The gyro torquing limitation coupled with the hands-off operational requirement and the fact that the instrument cluster's azimuth is initially completely unknown leads to a nonlinear estimation problem. To understand how this results it is necessary to review the technique of coarse alignment of the instrument cluster.

In order to apply linear control processes to align the platform, it is necessary to bring the platform close to the desired orientation, so that linearization of the angular deviations can be assumed. Theoretically linearity can be assumed for angular deviations up to 10 degrees. From a practical point of view it is desirable to achieve a considerably smaller deviation, since then the required angular rates to align have more reasonable magnitudes. For aircraft systems the platform can be approximately leveled by nulling out the pitch and roll gimbal readouts. This process is carried out while the gyros are being brought up to speed by simply torquing the gimbals with their own synchro readout signals. This technique is sufficiently accurate to justify linearization of the control problem, since it results in a deviation of approximately 1 degree from level for aircraft on the ground and 2 to 3 degrees for aircraft on the deck of a carrier. The larger level coarse alignment errors on the carrier result from the pitching and rolling of the ship.

The azimuth of the platform is usually slewed to approximately north (and for wander azimuth mechanizations the computer wander angle is initialized at zero) by torquing the azimuth gimbal synchro with the difference between the output of the azimuth gimbal synchro and a synchro signal obtained from a magnetic compass indicating the magnetic north heading of the aircraft. Additionally the magnetic heading can be compensated for magnetic variation to make the coarse azimuth alignment more accurate.

After the coarse alignment phase has been completed and the gyros are brought up to speed, the platform is placed under the control of the gyros. At this time the fine alignment phase can begin, or alternatively additional coarse alignment in level can be achieved by simply torquing the gyros with the output of the accelerometers. The additional coarse level alignment reduces the design problem for the fine alignment phase, since it reduces the dynamic range over which it must operate.

The application of this technique for shipboard coarse alignment unfortunately fails. This is because the magnetic heading device does not function with sufficient accuracy while in the vicinity of the massive steel deck and superstructure of the aircraft carrier. Thus, although the platform can be brought to a reasonably small coarse level alignment, no direct information is available to slew its azimuth to an approximate geographic referenced direction. If manual insertion of data were permitted the problem could be easily solved by visually estimating (to within 10 degrees) the angle between the ship's heading and the aircraft's heading (the spotting angle, see Figure 3). This angle, along with the ship's heading from SINS, would suffice for coarse azimuth alignment by either slaving the azimuth gimbal to a pre-assigned direction (such as north) or by initializing the computer wander azimuth angle.

It should be pointed out that if the mechanization is such that the platform is required to have a specific azimuth direction (as, for instance, in the north-pointing mechanization) then the platform must be close to that desired azimuth before the gyros have assumed control. The reason for this is that, because of the restriction on maximum gyro torquing, an excessive amount of time would be consumed if it were required to torque the platform through a large angle.

The solution to this problem is achieved by choosing one of the wander azimuth mechanizations and employing, after the gyros have assumed control, a nonlinear estimation process to estimate the platform's wander angle and initialize the computer.

For the ground alignment case (zero reference velocity) the wander angle of the platform is estimated from the residual velocity error after leveling has achieved steady state[7]. A similar procedure can be used for carrier deck alignment, but it must be dynamical, since the aircraft will be accelerating because of the carrier's motion.

This section has examined carrier deck alignment, detailing the multitude of ramifications of the operational, environmental, and control constraints imposed. In summary, carrier deck alignment must take place to a specified accuracy and within a specified time under the following conditions:

(i) No power cable.

(ii) No data cable.

(iii) No manual insertion of spotting angle.

(iv) No manual insertion of lever arm components.

(v) No loitering for the express purpose of alignment.

(vi) No restriction on location.

(vii) Taxiing permitted.

In Section 8 the filtering techniques employed in the solution of this problem are discussed.

6. CATAPULT ALIGNMENT

The most time-consuming phase of the initialization process is the azimuth alignment of the platform, principally because of the rate at which the azimuth information propagates into measurable observables. In an unaccelerated environment the azimuth misalignment propagates into the level channel at a rate proportional to either earth rate or spatial rate. This represents a relatively weak coupling; and if the measurable observables are contaminated with extraneous information, considerable time must be spent extracting the misalignment. In an accelerated environment, the azimuth misalignment propagates into the level channel proportional to the level acceleration which, if large enough, provides a much stronger coupling and reduces the azimuth alignment time. This line of reasoning leads to the concept of performing the azimuth alignment portion of the initialization process while the aircraft is being catapulted from the carrier.

Several different approaches have been devised to perform catapult alignment. One approach, developed by NADC* Johnsville, compares the velocity increment direction as computed by the aircraft system with the catapult direction as obtained from the SINS-computed ship's heading plus the angle between the ship's heading line and the catapult. A second approach employs two transponders mounted on the carrier and compares the aircraft computed and transponder-measured relative position to estimate the azimuth misalignment. Various modifications and generalizations of these two techniques are obviously possible.

Before catapulting the aircraft both techniques require the same amount of preparation, i.e. alignment up to and including fine leveling. If possible it is desirable to adhere to all the previous operational constraints. However, at its present state of development, the velocity increment approach requires one manually inserted piece of data: the catapult from which the aircraft is to be launched. This results because the technique requires the angle between the ship's heading line and the catapult. Since there are four catapults, a simple manual switch selection is necessary so that the computer can select the correct pre-stored angle. Other than this exception the techniques are encumbered with the same tasks as carrier deck alignment, i.e. lever arm estimation, automatic coarse azimuth alignment, relative velocity compensation, permission of aircraft movement, and fine level alignment. It should be pointed out that, since the azimuth alignment is delayed until catapulting occurs, the thermal transient problem is somewhat relieved in that the platform will have more time to reach thermal equilibrium.

The transponder technique, albeit requiring additional equipment both on the aircraft and on the aircraft carrier, does not require any manual intervention. In fact the problem of lever arm estimation is reduced to simply computing the aircraft's location from the transponder data received, thus relieving the alignment technique of the burden of implicit estimation of the lever arm.

With respect to the velocity increment approach, there are several significant effects that take place during catapulting, which must be considered in the design of the technique and in assessing the accuracy of the process. The more significant of these will now be discussed.

While being catapulted the aircraft is forcibly constrained to the deck and to follow the catapult track. However, because the track will not precisely constrain the lateral motion of the aircraft, an error in the reference direction arises. This is because the premise of the technique is that the aircraft (actually its platform) moves along a precisely defined line relative to the ship. Inherently the technique has no method for compensating or measuring this motion. The problem is additionally complicated by the fact that the lateral motion of the wheel in the track can possibly cause an amplified lateral motion of the platform depending upon its location in the aircraft. Without a technique for compensating or measuring this motion it represents an irremovable error.

Superimposed on the lateral track motion is the motion of the aircraft's inertial system due to the ship's pitching, rolling and yawing (during the launching of aircraft the heading of the ship is wilfully maintained

---

* *US Naval Air Development Center, Johnsville, Warminster, Pennsylvania.*

constant). Since the inertial system measures earth-referenced velocities, it will be sensitive to the velocities induced by these angular motions coupled through the lever arm. This motion can easily be compensated for by using the pitch, roll, and yaw rates as provided by SINS through the telemetering system and the estimated lever arm.

Finally it should be recognized that the accuracy of the ship's heading and the surveyed angle of the catapult enter prominently and directly into the accuracy of the azimuth alignment achievable. Senility of the ship's heading data is also of concern. Any lags induced because of the finite data rate of the telemetering system will induce errors proportional to the angular rate. Senility can be compensated for by increasing the data rate or introducing prediction techniques before using the data.

The transponder technique similarly has several effects to be considered. Since it is utilizing information at the position level to estimate the azimuth misalignment, it is not plagued by the irremovable catapult error nor the necessity for relative velocity correction. However, since the position as measured by the transponder is relative to the ship's body coordinates, it is necessary to resolve the measured position into the local-level coordinate system. This, of course, requires the pitch, roll, and heading of the ship.

Because of the relatively short base line available for the transponding system, small-distance errors can lead to relatively large azimuth errors. Thus it is important to obtain an accurate survey of the transponders relative to the SINS. Another significant possible source of error arises because the transponder's measured position is not precisely the position of the platform. However, this displacement error, to a large extent, can be compensated for by using the inertial system data.

The single most significant drawback of the transponder technique is the need to add additional equipment to the already overcrowded aircraft avionic system and to require an additional radiator to the aircraft carrier. Design constraints, similar to those imposed on the telemetering system, are required to maintain safety and security under electromagnetic silence.

7. IN-FLIGHT ALIGNMENT

Appreciating the numerous difficulties associated with carrier deck and catapult alignment it is worthwhile to consider that alignment be delayed until after the aircraft has left the carrier. This delay would eliminate the problem of interference with the deck handling procedures and significantly reduce the problem associated with the platform thermal transients. However it is encumbered with problems not heretofore considered. The discussion will be limited to performing in-flight alignment using a Doppler radar as the reference system. The details of the synergistic Doppler-inertial system have been adequately discussed in Chapter 11, so that only problems peculiar to the carrier aircraft application will be considered here.

To establish the state of the navigation function and the inertial platform at the time in-flight alignment begins, it is necessary to discuss the sequence of events relative to the navigation process from engine turn-on until the start of alignment. It is apparent that this epoch should be kept as small as possible since relatively poor, if any, navigation would be available and would introduce position errors that are irremovable without additionally complicating the navigation function by requiring position updating.

At the time in-flight alignment begins two conditions must be met. First, since it is used as the reference system during alignment, the Doppler must be functioning properly. This is specifically pointed out because a common, although not necessary, characteristic of a Doppler radar is that it will not track the aircraft velocity, until it has exceeded some specified minimum altitude. Thus, until the aircraft has attained this altitude and the Doppler begins to track the velocity, alignment cannot begin. Second, the platform must be coarsely aligned so that the Doppler data can be transformed, at least approximately, into the navigation coordinate system and also to permit linearization of the alignment process.

One technique for achieving these two conditions is to require the aircraft, after leaving the carrier and climbing to a safe altitude, to fly straight and level. This flight condition is maintained long enough to achieve coarse platform alignment and set the platform under gyro control. Coarse alignment is performed by nulling the pitch and roll gimbals and slaving the azimuth gimbal to magnetic north. This procedure has several disadvantages. It increases the time before the beginning of alignment and thus the time during which relatively poor navigation is performed. Also it places a constraint on the maneuvers that the aircraft can perform after takeoff from the carrier. Lastly for the pilot to fly straight and level he must have some reference of the vertical. This could be the horizon, if visibility permits, or some ancillary device.

A second technique would be to perform a coarse alignment on the carrier deck and, after the gyros have assumed control of the platform, to maintain the coarse alignment by pendulous erection and magnetic heading slaving. This procedure avoids the necessity of restricting the aircraft's maneuvers and permits alignment to begin as soon as the Doppler begins tracking the velocity. Careful consideration must be given to the mechanization of this technique, since the pendulously erected platform will tilt exorbitantly during large horizontal accelerations (catapulting and turning) and the magnetic heading device, as mentioned previously, will have large errors while on the carrier deck.

Excessive platform tilting can be avoided by inhibiting erection when the level accelerometer outputs exceed some specified value. The specific value would be ascertained by an analysis of the error dynamics. The coarse

azimuth alignment problem is handled as follows. Since an erroneous magnetic north will be indicated on the deck of the carrier, completion of coarse azimuth alignment must wait until the aircraft has left the ship. However, at this time large azimuth changes of the platform cannot take place within a reasonable amount of time because of the limit on gyro torquing. It is therefore necessary to employ a wander azimuth mechanization and perform coarse azimuth alignment by proper adjustment of the computer wander angle and not the platform. Thus, while on the carrier deck, the platform is at first coarsely aligned to the aircraft heading; and, after the gyros have resumed control, the computed wander angle is slaved to the difference between the platform heading and indicated magnetic heading (see Figure 4). Since large alterations in the computed wander angle can easily be made, no particular difficulty arises when the indicated magnetic heading changes significantly as the aircraft leaves the aircraft carrier.

The foregoing discussion establishes the state of the navigation function and the inertial platform at the time in-flight alignment begins. The problems associated with the over-water alignment process have been adequately discussed in Chapter II and will be briefly reiterated here:

(i) Distortion of Doppler spectrum return over water.

(ii) Loss of Doppler return over very smooth water surface.

(iii) Doppler radar measures velocity relative to water current rather than earth.

(iv) Alignment process equivalent in form to ground alignment and therefore requires a relatively long time.

(v) During alignment, navigation performance is degraded relative to a pre-aligned system.

Because of the unpredictability of the sea conditions and the length of time required for in-flight alignment, there is a preference to perform the alignment on the carrier deck.

8. CARRIER DECK ALIGNMENT FILTER

As pointed out in Section 5, the application of a Kalman filter to the carrier deck alignment problem must be preceded by an estimation of the wander azimuth angle. Because of complete lack of knowledge of the azimuth orientation of the platform, the initial wander azimuth angle is uniformly distributed between $-\pi$ and π. An extension of the error model presented in Section 3 to include the unknown wander angle results in a complex nonlinear model involving sines and cosines of the wander angle. Thus a direct application of a Kalman filter is not possible.

Because the Kalman filter requires a linear error model, it is necessary to estimate the large-angle azimuth offset by some other means. Since the linear range of the filter ($\sin\theta \simeq \theta$) extends to 10^0, a gross estimate can be obtained using a minimum variance estimator[11] comprising a three-state model (azimuth offset, two platform tilts) which relates commensurate aircraft and SINS velocities. The scheme is invoked after pendulous leveling has occurred and the gyros have taken control of the platform. For convenience the vertical platform axis is torqued so that the coarse azimuth offset remains constant. Once the estimate of azimuth offset is obtained, the alignment process is linear, and the standard form of Kalman filter is used to perform the fine alignment.

One of the major considerations in the application of a Kalman filter is the number of states to be modeled. As the number of states is increased the performance of the system will improve. However, a point of diminishing returns is soon reached, since the inclusion of relatively insignificant error processes improve the performance slightly but increase the computer requirements enormously. The choice of the number of states for the application under consideration is actually based on a dual use of the filter. The filter is required not only to perform the carrier deck alignment but also to provide the suboptimal error control for Doppler-inertial navigation. The size of the filter is therefore dictated by the mode requiring the most number of states.

Consideration of the significant error processes leads to a choice of 13 states. For the carrier deck alignment process they are

Position errors	2
Velocity errors	2
Platform tilts	2
Azimuth error	1
Level gyro drifts	2
Horizontal lever arms	2
Reference velocity error	2

The model for the first seven errors is obtained directly from Figure 2. The level gyro drifts and reference velocity errors are modeled as first-order correlated noise processes. Only the horizontal lever arm components are modeled, since during alignment the aircraft is always on the carrier deck and the vertical lever arm (relative to the body axes) is known. While the aircraft is stationary on the flight deck the body-axis lever arm components are time-invariant, so that the model chosen to represent them is simply a constant. This type of model is the same as a first-order correlated noise process with infinite correlation time.

Although level gyro bias is included in the carrier align filter, practically no gyro biasing will occur. Because of the high quality of the gyros required and the short time of alignment, the other errors of the system will dominate and not permit biasing to occur. However they are included for two reasons. First, they are required for Doppler-inertial navigation. Second, they can be used for periodic gyro biasing to remove the long-term (order of months) drifts. This periodic gyro biasing would be employed at scheduled maintenance time. Under these conditions sufficient time is available to reduce the errors of the system to such an extent that the biasing will occur.

Periods of taxiing are handled as follows: while the aircraft is stationary on the carrier deck its wheel brakes are on. If the aircraft taxis the wheel brakes are released. Thus a simple switch indicating the state of the brakes provides the computer with a discrete. Using this discrete the computer will temporarily suspend alignment whenever the brakes are released and the system reverts to free-inertial navigation. During this time the lever arm estimates are altered by integrating the difference between the aircraft computed velocity and the SINS telemetered velocity compensated for lever arm effects. This process will be in error because complete alignment has not occurred. Therefore during the taxiing the filter covariance matrix is propagated to reflect the fact that the errors in the lever arm estimates are changing. Once the brake is reapplied the system reverts to the alignment process.

The observation process consists of comparing the aircraft computed velocity with the SINS-telemetered velocity compensated for lever arm effects. The errors associated with this process are the errors in aircraft-computed velocity and azimuth, the errors in the SINS velocity, the errors in the lever arm estimates, and the errors in the SINS pitch, roll, heading and their rates. Since the filter states do not model the errors in the SINS pitch, roll, heading, and their rates, these errors are approximated as a white noise in the observation process.

The task of filter synthesis and performance estimation requires an extensive digital computer simulation. The basic computational modules of the simulation program consist of

(i) the filter,

(ii) the real-world system error model,

(iii) dynamics.

The filter module is a replica of the filter to be programmed on the aircraft's digital computer. As such it will be manipulated to obtain a compromised suboptimal performance. The basic objective is to obtain adequate control of the real world system error model and at the same time minimize the computational complexity. The real-world system error model module consists of the error models of all the various subsystems involved in the alignment process, i.e. the aircraft inertial system, the SINS, the telemetering system, etc. If Monte Carlo techniques are employed it is not necessary that these error models be linear. The dynamics module simulates all the various motions of the ship and the aircraft. For the carrier deck alignment problem it is important to include the dynamics of the ship's course and its pitching, rolling, and yawing[12]. Additionally it is required to simulate the aircraft's motion on the carrier deck, while it is catapulting and after catapulting.

The simulation process for filter synthesis and performance estimation consists of controlling the real-world system error model with the gains generated by the filter under test for a particular set of dynamics. The literature contains many examples of such studies for problems similar to the one under consideration here[13-14].

9. CONCLUSION

This chapter has considered the application of Kalman filtering to the problem of alignment of carrier aircraft inertial navigation systems. Considerable detail has been presented to establish the constraints under which the filter is to be designed.

In view of its usage in the solution of the problem, the local-level wander azimuth mechanization and error model were presented. For carrier deck alignment the interplay between the operational, environmental, and control constraints were used to define the conditions under which the filter must perform.

Catapult and in-flight alignment were examined as alternative or supplementary techniques. Problems similar to those encountered in carrier deck alignment were found, and the same or alternative solutions were discussed.

Finally a brief discussion of the filter for carrier deck alignment was given.

REFERENCES

1. Brockstein, A.J. Kouba, J.T. — *Derivation of Free Inertial, General Wander Azimuth Mechanization Equations.* Litton Systems, Guidance and Control Systems Division, Publication 7973, September, 1968.

2. Pinson, J.C. — *Inertial Guidance for Cruise Vehicles.* pp.113-187 of "Guidance and Control of Aerospace Vehicles", C.T.Leondes (Ed.), McGraw-Hill, New York, 1963.

3. — *Final Engineering Report, Transfer Align Demonstration.* US Navy Bureau of Weapons Contract NOw 62-0993-f, prepared by Litton Systems, Guidance and Control Division, Reference No. 2113-357, June, 1963.

4. Miller, B. — *Faster Carrier Inertial Alignment Sought.* Aviation Week, McGraw-Hill, 21 November, 1966.

5. Powers, H.B. Jr Henderson, V.D. — *Analytic Thermal Compensation of Fast Reaction Inertial Navigation Systems.* Presented at The National Aerospace Electronics Conference, Dayton, Ohio, 16-18, May, 1966.

6. McMurran, M. — *Fast-Reaction Alignment of Inertial Navigators.* Space/Aeronautics, July, 1967.

7. Lindner, R.I. — *Sensitivity Study in α Estimation and α Updating for a Wander Azimuth System.* Litton Systems, Guidance and Control Division, Publication 5956, 15 February, 1967.

8. Doty, R.L. Nease, R.F. — *Initial Conditions and Alignment.* pp.176 and 209 "Inertial Guidance", G.R.Pitman, (Ed.), Wiley, New York, 1962.

9. Kouba, J.T. — *At Sea SCRAMALIGN(1).* Litton Systems, Guidance and Control Division, Office Correspondence JTK-6812-174, 9 December, 1968.

10. Brockstein, A.J. — *Minimum Variance Solution to the At-Sea Coarse-Alignment Problem,* Litton Systems, Guidance and Control Division, Publication 8007, March 1969.

11. Liebelt, P.B. — *An Introduction to Optimal Estimation.* Addison-Wesley, Cambridge, Mass., 1967, p.164.

12. Brockstein, A.J. — *Design of a Model to Simulate Significant Aircraft Carrier (CVA-59) Motions,* Litton Systems, Guidance and Control Systems Division, Publication 7793, June, 1968.

13. Hodson, A.E. — *Optimal Leveling and Gyrocompassing of Inertial Systems.* Institute of Electrical and Electronic Engineers, Transactions on Aerospace and Navigational Electronics, March, 1963.

14. Jurenka, F.D. Leondes, C.T. — *Optimum Alignment of an Inertial Autonavigator.* Institute of Electrical and Electronic Engineers, Transactions on Aerospace and Electronic Systems, Vol.AES-3, No.6, November, 1967.

15. Danik, E. — *Integrated Inertial Velocity-Aided and Position-Aided Aircraft Navigation.* Presented at the 24th Annual Meeting of the Institute of Navigation, 21 June, 1968.

16. Sutherland, A.A. Jr — *The Kalman Filter in Transfer Alignment of Inertial Guidance Systems.* Journal of Spacecraft, Vol.5, No.10, October, 1968.

TABLE I

Definition of Variables for Figure 1

x, y, z	Navigation system coordinate axes
A_x, A_y	Accelerometer outputs
V_x, V_y, V_z	Velocity components relative to Earth
ρ_x, ρ_y, ρ_z	Platform angular rate components relative to Earth
Ω_x, Ω_y, Ω_z	Earth rate components
h	Geodetic altitude
a	Mean equatorial radius of the Earth
f	Flattening of the Earth
Φ	Geodetic latitude
λ	Longitude
α	Azimuth wander angle
C_{ij}	Direction cosines defining latitude, longitude and azimuth wander angle. Specifically $C_{xz} = \cos\alpha\cos\lambda - \sin\alpha\sin\Phi\sin\lambda$ $C_{xy} = -\sin\alpha\cos\lambda - \cos\alpha\sin\Phi\sin\lambda$ $C_{xz} = \cos\Phi\sin\lambda$ $C_{yx} = \sin\alpha\cos\Phi$ $C_{yy} = \cos\alpha\cos\Phi$ $C_{yz} = \sin\Phi$

TABLE II

Wander Angle Rate and Vertical Platform Angular Rate for Local-Level Mechanizations

Mechanization		$\dot{\alpha}$	ρ_z
North-pointing ($\alpha = 0$) (no polar capability)		0	$\frac{C_{Yy}\rho_y + C_{Yx}\rho_x}{1 - C_{Yz}^2} C_{Yz}$
Wander Azimuth (polar capability)	Free-azimuth	$-\dot{\lambda}C_{Yz}$	0
	Foucault	$-(\dot{\lambda}+\Omega)C_{Yz}$	$-\Omega C_{Yz}$
	Unipolar Upper sign Northern Hemisphere Lower sign Southern Hemisphere	$w\dot{\lambda}$	$-\frac{C_{Yy}\rho_y + C_{Yx}\rho_x}{(C_{Yz} \pm 1)}$

TABLE III

Definition of Variables for Inertial Error Model

$R_x = 1/a\,[1 - h/a - f(1 - 3C_{Yz}^2 - C_{Yy}^2)]$	approximation to the radius of curvature in the x , z plane
$R_y = 1/a\,[1 - h/a - f(1 - 3C_{Yy}^2 - C_{Yx}^2)]$	approximation to the radius of curvature in the y , z plane
A_z'	vertical acceleration minus gravity
g	gravity
∇_x, ∇_y	accelerometer errors
$\epsilon_x, \epsilon_y, \epsilon_z$	gyro errors
$\underline{\omega} = \underline{\rho} + \underline{\Omega}$	inertial angular rate

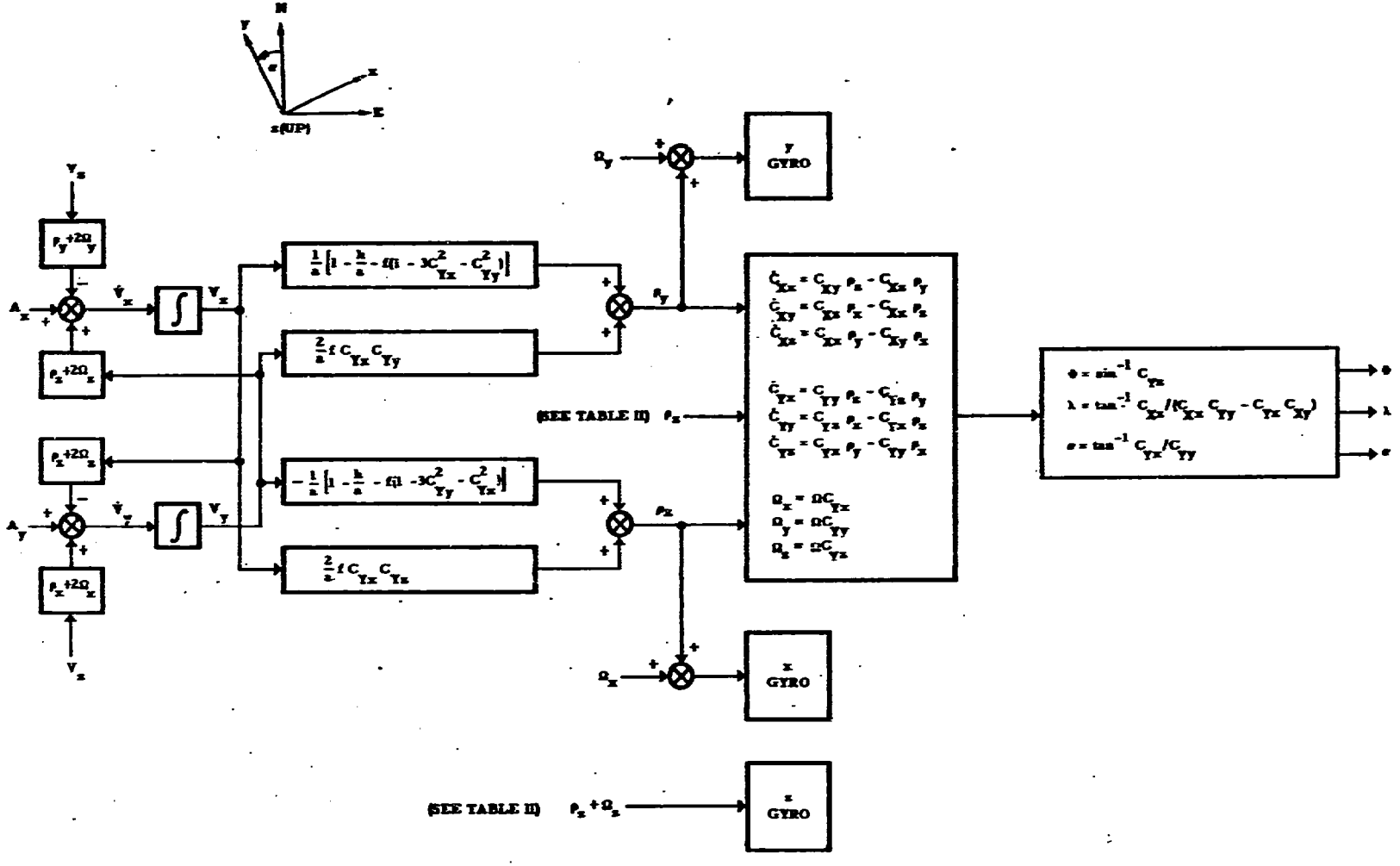
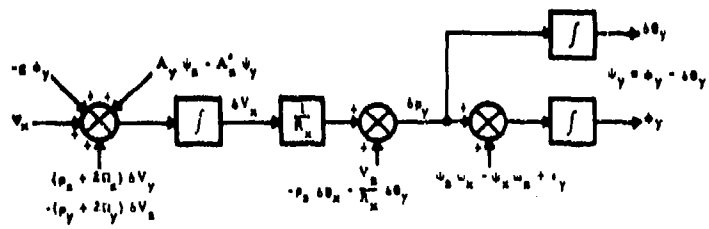
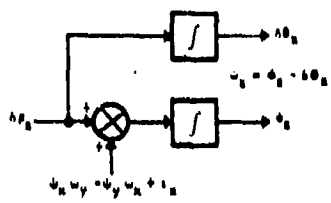


Fig.1 Generalized mechanization block diagram for local-level free-inertial navigation systems





MECHANIZATION		$\Delta\rho_s$
NORTH-POINTING (NO POLAR CAPABILITY)		$\text{Tan}\,\phi\Delta\rho_y - \rho_x \sec^2\phi\Delta\theta_y$
WANDER AZIMUTH (POLAR CAPABILITY)	FREE-AZIMUTH	$\rho_y\Delta\theta_x - \rho_x\Delta\theta_y$
	FOUCAULT	$\omega_y\Delta\theta_x - \omega_x\Delta\theta_y$
	UNIPOLAR UPPER SIGN NORTHERN HEMISPHERE LOWER SIGN SOUTHERN HEMISPHERE	$\frac{1}{C_{Ys}\Delta t}\left[\Delta(\rho_y\Delta\theta_x - \rho_x\Delta\theta_y) - (C_{yy}\Delta\rho_y + C_{yx}\Delta\rho_x)\right]$

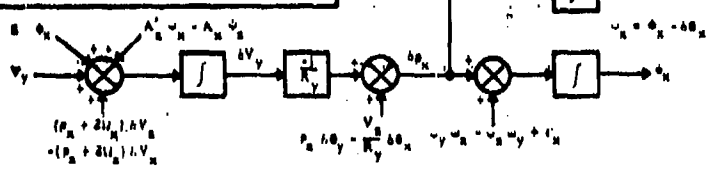
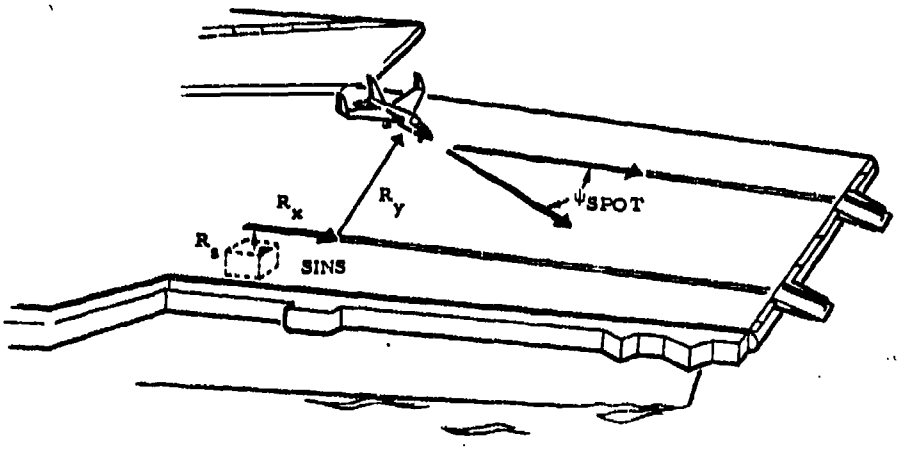


Fig.2 Generalised error block diagram for local-level free-inertial navigation systems

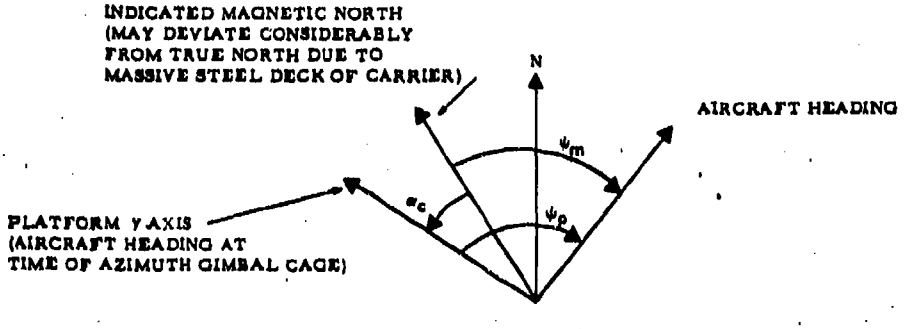


A. GEOMETRY

- NO RESTRICTION ON POSITION OF AIRCRAFT ON FLIGHT DECK
- NO DATA OR POWER CABLE
- NO MANUAL INSERTION OF LEVER ARM
- NO MANUAL INSERTION OF SPOTTING ANGLE
- AIRCRAFT MAY BE MOVED AFTER ALIGNMENT HAS BEEN INITIATED
- CAPABILITY OF ALIGNING IN FIVE MINUTES

B. DESIRED OPERATIONAL CHARACTERISTICS

Fig.3 Carrier deck alignment geometry



α_c - COMPUTER WANDER ANGLE

ψ_p - PLATFORM HEADING (AZIMUTH GIMBAL SYNCHRO READOUT)

ψ_m - INDICATED MAGNETIC NORTH

$\alpha_c = \psi_p - \psi_m$

Fig.4 Wander angle slaving to magnetic north

CHAPTER 16

SECTION II - NAVIGATION AT SEA USING THE INVARIANTS FORM OF KALMAN FILTERING

by

M.W.Sanbrook* and N.Halamandaris†

*Litton Systems Inc., Woodland Hills, California, USA

†Satellite Positioning Corporation
16033 Ventura Blod
Encino, California, USA

NOTATION

Σ, σ	covariance matrices
$\hat{X}$	estimates of state variables
x	state variable matrix
F	system dynamics matrix
G	system noise matrix
w	internal system noise
H	observation matrix
z	observation
v	measurement noise
Q	covariance of system noise
t	time
δ_j^k	Kronecker delta
A_{ij}	position-dependent covariance matrix elements
B_{ij}	position/non-position covariance matrix elements
S_{ij}	non-position-dependent covariance matrix elements
σ_{ij}	elements of standard covariance matrix
$\hat{x}_k$	transposed state variable matrix
U_P, U_S	matrices used in covariance matrix update
σ_N, σ_ϕ, σ_β^*	standard deviation of measurement noise
G_β, G_s	transposed observation matrix
K_T	Kalman gains
Superscript −	indicates before fix
Superscript +	indicates after fix
Subscripts x, y, z,	platform axes
ϕ_T	true latitude
λ_T	true longitude
ϕ_a	azimuth angle error
$\delta\phi$	latitude error
$\delta\lambda$	longitude error
V_{px}	platform x-velocity
V_{py}	platform y-velocity
Ω	earth rotation vector
Ω_N, Ω_L	north component of earth rate
Ω_v	vertical component of earth rate

Ω_x	platform x-axis component of earth rate
Ω_y	platform y-axis component of earth rate
Ω_z	platform vertical-axis component of earth rate ($= -\Omega_y$)
$\vec{\phi}$	platform-tilt vector
ϕ_x	tilt about platform x-axis
ϕ_y	tilt about platform y-axis
s	Laplace operator
ϵ_x, ϵ_E	X-gyro drift rate
ϵ_y, E_N	Y-gyro drift rate
$\delta\Omega_x$, $\delta\Omega_y$	error in x and y components of earth rate
EN77	drift rate of azimuth wander angle
$\vec{\omega}$	platform spatial rate vector
ω_{zr}	z-axis spatial rate
ω_{xr}	x-axis spatial rate
ω_{yr}	y-axis spatial rate
$\vec{\rho}$	craft rate vector (V_T/R)
ρ_x	x-axis craft rate ($-V_{py}/R$)
ρ_y	y-axis craft rate ($+V_{px}/R$)
ρ, ρ_{ij}	correlation coefficient
R_0	mean earth radius
V_x, δV_E	platform x-axis velocity error
V_y, δV_N	platform y-axis velocity error
CKS	accelerometer noise constant
g, G	gravity vectors
A_x, A_y	x- and y-axis acceleration components
DV_x, DV_y	incremental x- and y-axis platform velocities
δC_x, δC_y	x- and y-axis Coriolis errors
CKD	gyro-noise-model constant
DD	negative reciprocal of gyro correlation time
CKA	Doppler noise model constant
BB	negative reciprocal of Doppler noise model correlation time
ω_n	white noise
ω_i	white-noise components on state variables
θ, θ^*	observations
h	altitude
δV_{pL}	Doppler lateral-velocity error

δV_{DR}	Doppler range-velocity error
V_g, V_T	ground speed
δ_T, θ_D	total Doppler-drift-angle error
ψ_p	platform heading (craft to X platform axis)
δV_x, δV_y	velocity difference of inertial-Doppler units
δV_G	error in ground speed (including conversion)
SF	Doppler scale factor
δV_f	fluctuation error
δV_{Dx}	x-axis Doppler error
δV_{Dy}	y-axis Doppler error
A_z	vertical acceleration
μ_i, σ_i	accelerometer non-orthogonalities ($i = 1,3$)
∇_x, ∇_y	accelerometer bias errors
SF_{Ax}	x-accelerometer scale factor
SF_{Ay}	y-accelerometer scale factor
ν	reciprocal of accelerometer correlation time
TMX	x-gyro torquing signal from filter
TMY	y-gyro torquing signal from filter
ξ	reciprocal of gyro correlation time
γ_1^2	amplitude of gyro noise drift
β_1^2	amplitude of accelerometer noise
SF_{Tx}	x-gyro scale-factor error
SF_{Ty}	y-gyro scale-factor error
SF_{Tz}	z-gyro scale-factor error
R_x, y, z	gyro torquing rates
$k_1 - k_6$	gyro non-orthogonalities
C_i	acceleration-sensitive drift coefficients for gyros
C_{ik}	acceleration-squared drift coefficients for gyros
τ_g	gyro-drift correlation time
η	integration algorithm constant
μ	integration algorithm constant

CHAPTER 16

SECTION II - NAVIGATION AT SEA USING THE INVARIANTS FORM OF KALMAN FILTERING

H. W. Bambrook and H. Halamandaris

1. INTRODUCTION

Mechanization concepts employing optimal control techniques have been applied to guidance and navigation problems over the past five years, the object being to produce a system which exhibits a minimum error variance under the constraint of time-optimality. The latter is attained by separating the signal from the noise in the aid sensor data and maximally utilizing this data.

Inertial measurement units (IMU) exhibit error characteristics which grow with time. A mechanization utilizing aid sensors to damp and/or "bound" the inertial-unit-error characteristics is desired. By mixing a Doppler radar which possesses good long-term stability and an inertial measurement unit which has good short-term stability, the overall system accuracy is greatly improved over that attainable from either sensor independently.

The inertial system is the primary sensor and the Doppler serves as the aid sensor. Observations made by the Doppler are contaminated by the sea current drifts, since the Doppler measurements are relative to the water mass. The application discussed in this section specifically addresses an airborne application; however, the mathematical mechanization as configured is directly applicable to a shipboard inertial unit damped by an EM log or Doppler sonar. In either the airborne or shipboard case, the largest error contributor in the Doppler observation is the sea current drift. The hardware biases characteristic of modern airborne and shipboard sonar Dopplers are relatively small compared to the sea noises.

This section contains a brief summary of Kalman filtering theory which, through a transformation of variables, is extended to the invariants form of Kalman filtering. A north-seeking model is discussed herein for illustrative purposes.

Practical considerations play an important role in configuring a Kalman filter; for instance, to mechanize an optimal filter for an inertial navigation system requires an expanded memory airborne (or shipborne) computer. Due to weight, volume, cost, power reliability and maintainability constraints, it is not, in general, desirable to implement a fully-optimized system. Therefore, suboptimal filter mechanizations are generally influenced by the above-mentioned constraints. In keeping with this thinking, the invariants technique minimizes the number of differential equations to be programmed in the covariance matrix. Also scaling requirements, for the covariance matrix, are not as stringent because of the bounded characteristics of subsets of the covariance matrix. Further, the theoretical accuracy realizable from the invariants form of the Kalman filter is as good as the classical Kalman filter. However, the word-length requirements can be relaxed somewhat for the invariants technique because it is less sensitive to truncation and round-off effects than the classical Kalman filter.

To illustrate the potential accuracy improvement attainable from the implementation of the optimal filter, Section 5 shows the result of a simulation for which the filter was run as an (open-loop) estimator rather than a (closed-loop) controller. The degree of predictability can be seen from the illustration in Section 5. The conclusion is that the invariants form of Kalman filtering can indeed predict the states of the system.

2. KALMAN FILTER THEORY FOR THE INVARIANTS TECHNIQUE

2.1 General

This subsection contains the summary of Kalman's equation and the application of the Kalman filter based on the Litton invariants technique. The general method is that employed by Kalman. Certain modifications have been made to accommodate constraints imposed by the suboptimal characteristics of the mechanized filter. Also certain refinements, such as exponential torquing of the gyros and Doppler, have been incorporated to simulate the effect of exponentially-correlated noise sources. The following discussion presents the general philosophy of the Kalman filter.

The inertial navigation element (IMU plus computer) provides the complete navigation function; however, by using redundant data, its performance can be improved. The primary aid sensor considered (the Doppler radar) provides velocity information in the form of ground speed and Doppler drift angle. The Doppler radar also requires a heading reference such that combination of these sensors can also perform the navigation function, although it is somewhat degraded. An alternative use of the Doppler velocity is to damp the inertial system. The aid sensors also provide intermittent position data.

Thus it is seen that the various other avionic subsystems provide a redundant measurement of position and velocity. Optimal usage of this redundant information requires a detailed examination of the relative errors of the various sensors in addition to the availability of the data.

There are various methods of utilizing the redundant information, their quality depending heavily on the accounting for the many types, particular magnitudes and spectra of error sources. Of prime concern are classical and Kalman mechanizations. Since all mechanizations are imperfect, the problem is to extract signal from noise. The classical mechanization approach is to let the systems do much of the filtering by utilizing feedback.

This chapter considers Kalman filtering for airborne alignment and navigation. It is well known that such a filter technique results in a minimum variance error for the assumed-system-error model.

The error propagation of the inertial unit and Doppler error are both strongly flight-path-dependent, hence time-varying. Since the flight path is not fixed for all possible missions, it is clear that there should be some onboard accounting for flight path in the filtering, as referred to later, to obtain best system performance. Where intermittent or occasional information is supplied, such as position data, the Kalman filter is especially suited because of its capability of making large corrections throughout the inertial system. Provided the system error model is sufficiently realistic, improved system performance is obtained by employing Kalman filtering techniques. The mechanization advantages of classical feedback are obtainable in Kalman filtering. With a digital computer, it is important to note that the selection of filtering methods is reduced to only two considerations:

(i) System performance

(ii) Computer requirements.

This implies a sophisticated model because there are many important error sources and types of error propagation. Any increase of sophistication of the error model typically implies increased programming requirements. Utilization of results of investigations of computation techniques closely related to the deeper aspects of the Kalman filter approach are evidently not only warranted, but should be given priority.

A frequently considered aspect of the use of any filtering technique is the dependence on correctness of the hypothesized error model. If a sophisticated error propagation model is used, then unrealism can stem only from inaccuracy of statistical estimates of error source magnitudes, e.g. drift-rate characteristics of the particular gyro. If the mechanization of the Kalman filter is a feedback system, as is the proposed system, then the advantages of the relative insensitivity of that type of system to unexpectedly large errors are attained. In all simulations of Kalman filters at Litton, it should be mentioned that error sources have particular characteristics in each run, randomly chosen from a distribution with the assumed characteristics.

The principal difficulty in applying Kalman filtering to airborne systems was the large computer capacity required to mechanize the myriad equations. Litton has managed, by transformation of variables, sequential computation, and approximations, to simplify Kalman's general mechanization equations.

The transformation of variables is one in which the majority of the elements of the system covariance matrix are invariant to the position aid sensor data. This transformation of variables is presented in more detail in a later section.

2.2 Classical Kalman-Filter-Equation Summary and Mathematical Preliminaries

The mathematical development of the classical Kalman-filter equations is formulated in myriad recent publications, including Kalman's complete documentation[1]; hence, these equations will not be derived here but merely summarized and described in general terms.

Table I summarizes the Kalman/Litton equations that would typically be mechanized in an aided inertial navigation system for alignment and navigation. The notation employed is that used by Kalman. Underlined capital letters signify matrices, underlined lower-case letters signify vector quantities (column matrices). A prime signifies the transpose of, the "hat" sign signifies best estimate of, the + sign signifies the pseudo-inverse of, and E means the expected value of. The classical Kalman filter is modified slightly in that the best estimates of state variables $\underline{x}$ are updated with information including the N^{th} increment of time.

In setting up a Litton mechanization, the first step is to select a set of state variables. These will normally be any and every quantity which we wish to update and sufficient states to adequately describe the dynamics of the system.

The next step is to represent the system dynamics by a set of first-order linear differential equations of the form of Equation (2.1). These exist as a set of such equations by defining the state variables as error quantities. For example, the state variables are not defined as "position" and "velocity", but as "error in position" and "error in velocity". The standard linearized error equations (neglecting second-and-higher-order terms in error quantities) will then provide equations of the form of Equation (2.1). Inspection of these equations yields the $\underline{F}$ and $\underline{G}$ matrices, where $\underline{F}$ is referenced as the system dynamics matrix.

The $\underline{G}\underline{w}$ terms represent internally generated noise. Gyro drift noise is an example of internal noise represented by elements of $\underline{G}\underline{w}$.

TABLE 1

Summary of Kalman Filtering Equations

Defining Equations

State variables: $\underline{x}$ $\quad \frac{d\underline{x}}{dt} = \underline{F}\underline{x} + \underline{G}\underline{w}$ (2.1)

Observed quantities: $\underline{z}$ $\quad \underline{z} = \underline{H}\underline{x} + \underline{y}$ (2.2)

Initialize with best available estimates of state variables $\underline{\hat{x}}$ and covariance matrix $\underline{\Sigma}$ (error in estimates)

Between updating points:

(a) Propagate state variable estimates by $\frac{d\underline{\hat{x}}}{dt} = \underline{F}\underline{\hat{x}}$ (2.3)

(b) Propagate covariance by $\frac{d\underline{\Sigma}}{dt} = \underline{F}\underline{\Sigma} + \underline{\Sigma}\underline{F}' + \underline{G}\underline{Q}\underline{G}'$, (2.4)

where $\underline{\Sigma} = \text{cov}[\underline{\hat{x}}] = E[\underline{\hat{x}}\underline{\hat{x}}']$

$\underline{Q} = \text{cov}[\underline{w}] = E[\underline{w}\underline{w}']$

$\underline{R} = \text{cov}[\underline{y}] = E[\underline{y}\underline{y}']$

Note: To use Equations (3.4) through (3.7) we assume $E[\underline{w}\underline{y}'] = 0$, i.e., $\underline{w}$ and $\underline{y}$ are uncorrelated.

(a) Correct state variables by $\Delta\underline{\hat{x}} = \underline{K}(\underline{z} - \underline{H}\underline{\hat{x}})$ (2.5)

$\underline{K} = \underline{\Sigma}\underline{H}'[\underline{H}\underline{\Sigma}\underline{H}' + \underline{R}]^{\dagger} = \frac{\underline{\Sigma}\underline{H}'}{D}$, (2.6)

where $D = \underline{H}\underline{\Sigma}\underline{H}' + \underline{R}$ is a scalar quantity if updates are based on only one measurement at a time.

(b) Correct covariance matrix by $\Delta\underline{\Sigma} = -\underline{K}\underline{H}\underline{\Sigma} = -D\underline{K}\underline{K}'$. (2.7)

The observed or measured quantities used to update the state variables must also be specified. These are position checkpoints and Doppler velocity measurements. These "observables" must be expressed as a linear function of the state variables, as in Equation (2.2). The linearized error equations expressing errors in observations as functions of errors in the state variables provide Equation (2.2) and the observation matrix $\underline{H}$.

The $\underline{y}$ vector represents measurement errors. In most practical applications, the measurement errors $\underline{y}$ are uncorrelated with (statistically independent of) the internal noise $\underline{w}$. To simplify the remainder of the equations, assume that $E[\underline{w}\underline{y}'] = 0$.

Equations (2.3) and (2.4) are the mechanization equations for real time computation of the estimated values of the state variables and the expected variance in these estimates. Equations (2.3) are essentially the navigation equations which are already mechanized in the inertial navigation system. Equation (2.4) is new to the mechanization. The covariance $\underline{\Sigma}$ represents a running estimate of the accuracy or confidence in the system estimates of each state variable. In effect, the system CEP is continually computed, implying that the covariance matrix is a running error analysis.

Equations (2.3) and (2.4) are differential equations which are solved by integrating from some known starting point. Initialization is a familiar problem for Equation (2.3) and corresponds to aligning the platform, and setting in initial velocity and positions into the computer. To initialize Equation (2.4), an estimate of platform alignment, gyro calibration and computer initialization accuracy must be made. Such estimates are based on error analyses and actual test data. These estimates are not critical to the process if they are not grossly underestimated. Studies show that it is better to overestimate the error for the self-correlated terms (variances) rather than to underestimate, whereas it is better to underestimate the cross-correlated terms (covariances). Upon initializing Equations (2.3) and (2.4) the computer continuously tracks the state variables and updates the covariance matrix by integrating Equations (3.3) and (3.4) in real time.

When aid information is obtained from separate measurements or observations, the present estimates of the state variables and covariance matrix are upuated according to Equations (2.5) and (2.7). Equation (2.5) states that a comparison of the measured value $\underline{z}$ with the present estimate $\underline{H}\underline{\hat{x}}$ be made, and this difference signal multiplied by a feedback gain $\underline{K}$.

One of the simplifications instituted by Litton is to serially update after a measurement. Even if data from two or more measurements are available simultaneously, the data can still be processed serially as several single measurements and the system sequentially updated. This procedure reduces each of the $\underline{H}$ matrices to a single row matrix. This in turn makes the quantity $\underline{H}\Sigma\underline{H}' + \underline{R}$ a scalar quantity, and the inverse of a scalar D is simply its reciprocal D^{-1}.

Once the filter gain $\underline{K}$ and the scalar D from Equation (2.6) are computed, the corrections to the elements of the covariance matrix become almost trivial, as seen in Equation (2.7). Following the update corrections, the systems computations revert back to Equations (2.3) and (2.4) in order to propagate the system estimates forward in real time until the next system updating occurs.

In addition to the above, the development of the correction invariant formulas requires a cursory knowledge of tensor notation, specifically, the concept of the summation convention. In tensors, to denote different quantities, superscript indices are employed. In the development which follows, subscripts are used throughout except on the Kronecker delta terms.

In writing homogeneous linear and quadratic functions summed over some index, the following types of formulas, respectively, result:

$$\sum_{j=1}^{n} a_j x_j = a_1 x_1 + a_2 x_2 + a_3 x_3 + \cdots$$

and

$$\sum_{k=1}^{n} \sum_{j=1}^{n} a_{kj} x_k x_j = a_{11}x_1^2 + a_{22}x_2^2 + \cdots + a_{nn}x_n^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + \cdots + a_{1n}x_1x_n + a_{21}x_2x_1 + a_{23}x_2x_3 + \cdots + a_{2n}x_2x_n + \cdots + a_{n1}x_nx_1 + \cdots$$

These two equations are systems of the first order and systems of the second order, respectively. A system of order zero has no indices and thus is invariant.

In the course of the derivations that follow, the summation signs (Σ) would be used quite frequently. To present a more compact appearance for the formulas, the Σ sign will be dropped and is implied whenever an index is repeated in the expression, e.g.,

$$\sum_{j=1}^{n} a_j x_j \triangleq a_j x_j$$

and, since j is repeated, the summation over j is implied on the right-hand side of the expression above. The convention applies in similar manner to the double summation, since the indices j and k are both repeated. To facilitate the analysis, a dummy index is frequently employed. An index not repeated is a free index and spans any dimension assigned.

The Kronecker delta (δ_j^k) is a mixed tensor of the second order (rank two), because it employs both subscripts and superscripts and has the unique property that

$$\delta_j^k = \begin{cases} 0 & \text{if} \quad j = k \\ 1 & \text{if} \quad j = k \end{cases}$$

2.3 General Implications of Correction Invariants

The error propagation of the inertial system can be shown to be computed by differential equations in the covariance matrix, which, by a special transformation of variables, has the form

$$\begin{bmatrix} \underline{A} & \underline{B} \\ \underline{B} & \underline{S} \end{bmatrix}$$

where $\underline{B}, \underline{S}$ are invariant under Kalman corrections with respect to a chosen subset of observed state variables and $\underline{A}$ is the inverse of the matrix of that chosen subset of variables which are, in part, directly observed by aid sensors.

The $\underline{A}$ submatrix, which comprises the small portion of non-invariants of the computation set, is functionally the inverse of a partition of the error matrix and, as such, has diagonal terms that vary proportionally to system accuracy which, in turn, lies between zero and a sharply defined upper band. In contrast, the direct error quantities are relatively unbounded. Useful information is best mapped in terms of accuracy measure in statistics for combination with other information. (Thus, defining accuracy $A \approx 1/\sigma_p^2$, where σ_p^2 is the variance of the position data, it follows that the accuracy obtainable by many observations is the sum of A_i.)

Defining Kalman's covariance matrix as σ, with subscripts x and y denoting the position variables, the A submatrix is functionally

$$A_{ps} = \sigma_{ps}^{-1} = \frac{\begin{vmatrix} \sigma_{yy} & -\sigma_{xy} \\ -\sigma_{xy} & \sigma_{xx} \end{vmatrix}}{\sigma_{yy}\sigma_{xx} - \sigma_{xy}^2}$$

The R invariants, which appear in the Kalman weighting for inertial system trim-up, possess the functional behavior of ratios of related covariances. The ratios have relatively limited range compared to related correlations. Functionally, the R submatrix is

$$R_{rs} = \sum_{k=1}^{k} \sigma_{rk} A_{ks} ,$$

where $r > k$, $s \leq k$, and r and s refer to the members of the R submatrix.

The S invariants (comprising the larger part of modified matrix equations), which correspond to correlations of non-observables, are the most insensitive to the aid data profile. This offers a potential memory reduction by using approximations which do not materially compromise performance optimality. The S submatrix elements have the same units as the corresponding members of the classical covariance matrix (σ) and functionally are

$$S_{rs} = \sigma_{rs} - \sum_{k=1}^{k} R_{rk}\sigma_{ks} ,$$

where $r, s > k$.

The choice of correction invariants with respect to position observations leads to error propagation differential equations in which latitude error effects in level and vertical components of earth rate and Coriolis computation are largely eliminated from the R and S submatrix elements. This presents a memory saving and removes problems of computation of very small terms whose integrated magnitude over 10 hours may, nevertheless, be substantial.

The correction invariants tend to retain the simple cross-channel characteristics of the pure inertial system, otherwise disturbed by the aiding process. These characteristics are principally the result of azimuth error and, usually to a lesser degree, latitude error entering earth rate estimates. With computation in invariant form, it has been demonstrated that a reduction in differential equation programming is feasible, the obviated equations being replaced with a few closed form azimuthal correlation relations which enable azimuth correction based on accelerative effects (important at high latitudes). The computation techniques proposed for the invariants are presented in the subsequent paragraphs.

The A submatrix is restricted to contain only position variables; the R submatrix contains the covariance of one position variable with one non-position variable. The S submatrix contain the covariance of all non-position variables with non-position variables.

2.4 Algebraic Form of Correction Invariants

An observation involving a given subset of system error ϵ_k, $k = 1, \ldots, K$ of form

$$\underline{z} = \sum_{k=1}^{K} \epsilon_k G_k + N$$

is utilized for Kalman corrections (analogous to Kalman's expression, $\underline{z} = \underline{H}\underline{x} + \underline{v}$) of each system variable, where it may be shown that the change upon correction implies that

$$\Delta\Sigma_{rs} = -\frac{U_r^- U_s^-}{\sigma_N^2 + G_k U_k} = -U_s^- U_r^+ / \sigma_N^2 , \qquad (2.8)$$

where σ_N^2 is the observation error variance, r and s refer to variable labels, the + and − superscripts refer to after and before system correction, and

$$U_r = \sum_{k=1}^{k} \sigma_{rk} G_k .$$

Furthermore, A_{ps} is defined

$$A_{ps} = \sigma_{ps}^{-1} . \qquad (2.9)$$

It may be proven, from the form of the Kalman correction formulas presented above, that

$$R_{rs} = \sum_{k=1}^{N} \sigma_{rk} A_{ks} , \qquad \begin{matrix} r > K^{\dagger} \\ s \leq K \end{matrix}$$

$$S_{rs} = \sigma_{rs} - \sum_{k=1}^{N} R_{rk} \sigma_{ks} , \qquad s, r > K$$

undergo changes under system correction $\Delta R_{rs} = 0$, $\Delta S_{rs} = 0$; that is, R_{rs} , S_{rs} are correction invariants for the chosen invariant set (position). Note that the invariants are not functions of the observation weights (G_k) nor observation accuracy. In particular, it is seen that invuriants for any position observation components are the same. It can be shown that

$$\Delta A_{ps} = G_p G_s / \sigma_N^2 , \qquad p , s \leq K . \tag{2.10}$$

The error propagation of $\underline{A}$, $\underline{R}$, $\underline{S}$ may be expressed directly by differential equations in terms of $\underline{A}$, $\underline{R}$, $\underline{S}$ alone. The Kalman system corrections have weight

$$K_r = \frac{U_r^+}{\sigma_N^2} = \frac{1}{\sigma_N^2} \sum_s R_{rs} U_s^* , \tag{2.11}$$

where $U_s^* = \Sigma_p G_p A_{sp}^{-1}$. For the case where the correction invariants are defined with respect to position-type observations,

$$\sigma_{sp} = A_{sp}^{-1} = \frac{\begin{bmatrix} A_{22} & -A_{12} \\ -A_{12} & A_{11} \end{bmatrix}}{A_{11}A_{22} - A_{12}^2} .$$

This implies that, upon receipt of a position fix, only the A matrix (3 elements) need be updated; the R and S submatrices (39 elements) are invariant, and, hence, do not change. An abbreviated proof of A , R , S properties follows:

(a) The correction formulas for A_{ps} defined with respect to a selected observation type are obtained from the ordinary Kalman formula for that observation type after taking the first difference of the identity,

$$A_{pk}\sigma_{kr} = \sigma_{pk}^{-1}\sigma_{kr} = \delta_p^r , \qquad p , r , k , \leq K .$$

Thus

$$\left.\begin{aligned} \Delta A_{pk}\sigma_{kr}^{+} + A_{pk}^{-}\Delta\sigma_{kr} &= 0 \\ \Delta A_{pk}\sigma_{kr}^{+} - A_{pk}^{-}U_k^{-}U_r^{+}/\sigma_\beta^2 &= 0 , \end{aligned}\right\} \tag{2.12}$$

where a summation convention is used on paired indices. Multiplying by $(\sigma_{rs}^{-1})^{+}$ and noting that

$$A_{pr}U_r = \sigma_{pr}^{-1}U_r = \sigma_{pr}\sigma_{rs}G_s = G_p , \qquad s \leq K ,$$

obtain

$$\Delta A_{ps} - A_{pk}^{-}U_k^{-}U_r^{+}A_{rs}^{+}/\sigma_\beta^2 = 0 , \tag{2.13}$$

whence the correction formula

$$\Delta A_{ps} = \frac{G_p G_s}{\sigma_\beta^2} , \qquad p , s \leq K . \tag{2.14}$$

(b) The invariance of R_{rs} defined with respect to a selected observation type is shown, taking the first difference of the defining equation for R ,

$$\Delta R_{rs} = \Delta\sigma_{rk}A_{ks}^{+} + \sigma_{rk}^{-}\Delta A_{ks} \qquad \begin{matrix} r > K \\ k , s \leq K . \end{matrix} \tag{2.15}$$

Substituting the ordinary Kalman correction and the ΔA_{ps} expressions

$$\Delta R_{rs} = \frac{-U_r^{-}U_k^{+}}{\sigma_\beta^2} A_{ks}^{+} + \sigma_{rk}^{-}\frac{G_k U_s}{\sigma_\beta^2} , \tag{2.16}$$

and using Equation (2.12) and the definition of U_r .

---

† These restrictions are redundant; e.g., all S invariants are zero where position is involved, and $R_{xx} = 1$, $R_{yy} = 1$ and $R_{xy} = 0$.

$$\Delta R_{rs} = \frac{-U_r^- G_s}{\sigma_\theta^2} + \sigma_{rk}^- \frac{G_k G_s}{\sigma_\theta^2} = \frac{G_s}{\sigma_\theta^2}\left[-U_r^- + \sigma_{rk}^- G_k\right] = 0 .$$

Hence R_{rs} is invariant.

(c) The invariance of S_{rs} defined with respect to a selected observation type is shown by taking the first difference of the defining equation, using the relation $\Delta R_{rk} = 0$,

$$\Delta S_{rs} = \Delta\sigma_{rs} - R_{rk}\Delta\sigma_{sk} , \qquad \begin{matrix} r , s > K \\ k \leq K , \end{matrix} \tag{2.17}$$

and then substituting the ordinary Kalman expressions to obtain

$$\Delta S_{rs} = \frac{-U_s^- U_r^+}{\sigma_\theta^2} + R_{rk}\frac{U_s^- U_k^+}{\sigma_\theta^2}$$

$$= \frac{U_s^-}{\sigma_\theta^2}\left[-U_r^+ + R_{rk}^+ U_k^+\right]$$

using $R_{rs}^+ = R_{rs}^-$.

Substituting the defining expression for R_{rk} ,

$$R_{rk}^+ U_k^+ = \sigma_{rp}^+ A_{pk} U_k^+ = \sigma_{rp}^+ G_p = U_r^+ .$$

Hence $\Delta S_{rs} = 0$; that is, S_{rs} is correction invariant relative to the associated type of observation.

2.5 Correction Formulas for Disparate Observations

The correction invariant set for observations θ (involving, for example, position errors) are not invariant under corrections with different kinds of observations (such as involving velocity or azimuth errors). The following is a derivation of correction formulas for disparate observations, which are shown to be closely similiar to ordinary Kalman formulas but in the variables of the correction invariant set. The algebraic derivation differs for unassociated observations as a result. In Equation (2.8) the terms $U_s \rightarrow U_s$ where†

$$\begin{aligned} U_r &= \sigma_{rs} G_s , \quad s \leq K \\ U_r &= \sigma_{rq} G_q , \quad q > K . \end{aligned} \tag{2.18}$$

That is, q is not in the set of variables for which A_{ps} is defined. Consequently, $U_r A_{rs} \neq G_r$ and Equation (2.14) does not hold. Indeed, for the unassociated observations $U_r A_{ts} = \sigma_{rq} G_q A_{rs} = G_q R_{qs} = R_{\theta s}$, where R is the correction invarant set of the computation process and denoting $R_{\theta s} = G_q R_{qs}$ for θ disparate type of observation. The Equation (2.10) is replaced with

$$\Delta A_{ps} = \frac{R_{\theta * p} R_{\theta * s}}{\sigma_\theta^2} , \tag{2.19}$$

which is a formula computationally similar to the ordinary Kalman formula except for sign of the correction. The correction R_{rs} for unassociated observation type is obtained from the R-matrix defining Equation (3.13) and

$$\Delta\sigma_{rk}^- = \frac{-U_r^- U_k^+}{\sigma_\theta^2} . \tag{2.20}$$

Further manipulation yields the correction formula for R_{rs} with unassociated observation:

$$\Delta R_s = \frac{-R_{\theta * s}^+ S_{r\theta *}^-}{\sigma_{\theta *}^2} \qquad \begin{matrix} r > K \\ s \leq K . \end{matrix} \tag{2.21}$$

The implied computation is directly analogous to the ordinary Kalman correction formulas.

The corrections of S_{rs} for unassociated types is obtained in the same manner.

---

† Retaining use of summation convention on paired indices.

The correction formula for unassociated types of observations is

$$\Delta s_{rm} = - \frac{s^*_{rc} * s^*_{mc} *}{\sigma^2_c} \quad \text{where} \quad r, m > k .$$

The implied computations are again directly analogous to the ordinary Kalman correction formulas.

2.6 Computational Features of Correction Invariants

The Kalman filtering computations relating to error growth, when executed in correction invariant variables, have a number of markedly different characteristics from the direct form with impact on computer cost, including effects of the following:

(i) Obviation of the bulk of the error-matrix update computations following inertial system corrections with aid data, this being the consequence of the invariance properties. This feature makes a wired incremental-computation program of digital differential analyzer program attractive because of substantial reduction of memory requirements associated with instruction memory, the variable memory, however, being increased to a lesser degree.

(ii) Reduced word-length requirements and/or decreased computation error sensitivity.

(a) The bulk of the variables in the correction invariants computations have readily determined variation ranges which change little for disparate flight and aid data profiles. In contrast, variances which have the range of the square of errors which themselves have wide range are unwidely computationally and can require monitoring to prevent overflow under varying operational conditions.

(b) The A matrix, which comprises the small portion of non-invariants of the computation set, is functionally the inverse of a partition of the error matrix and as such has diagonal terms which vary proportionately to system accuracy, which lies between zero and a sharply defined upper band. In contrast, the direct error quantities are relatively unbounded. Useful information is best mapped in terms of accuracy measure in statistics for combination with other information. (Thus, defining accuracy $A = 1/\sigma^2$, where σ^2 is the variance, it follows that the accuracy obtainable by many observations is the sum of accuracy.)

(c) The R invariants, which appear in the Kalman weighting for inertial system trim-up, have functional behavior of ratios of related variances. The ratios have relatively limited range compared to related correlations.

(d) The S invariants (comprising the larger part of modified matrix equations), which correspond to correlations of non-observables, are the most insensitive to aid data profile, offering a potential of extensive memory reduction by approximations which do not materially compromise performance optimality.

(iii) Memory saving by reduced computation program relating to cross-channel inertial variable correlations.

The correction invariants tend to retain the simple cross-channel characteristics of the pure inertial system, otherwise disturbed by the aiding process, which are principally the result of azimuth error (and latitude error entering earth rate estimates to a usually smaller degree). With computation in invariants form, it has been demonstrated that a 20-40% reduction in *differential equations programming* is feasible, the obviated equations being replaced with a few closed form azimuthal correlation relations which enable capability for azimuth correction based on accelerative effects (important at high latitudes).

3. SYSTEM-ERROR MODELS

The basic error sources considered are characteristic of the inertial sensors (gyros and accelerometers) and misalignments of these sensors. Random errors which are described by statistical autocorrelation functions were considered for the sensors. These error models are described in more detail in the subsequent sections. A north-seeking coordinate frame is assumed with the axes orientation shown in Figure 1. (The inertial-system error-model block diagram is shown later in Figure 5.)

3.1 Inertial Sensor-Error Models

3.1.1 General

The inertial sensor-error models are contained in the general IMU error model; however, a brief description is presented here to clarify the models. In the description of the various sensors, the noise function considered is exponentially correlated with an autocorrelation function of the form

$$R(\tau) = \beta^2 e^{-\nu|\tau|} ,$$

where β^2 is the noise-amplitude function and ν is the reciprocal of the correlation time characteristics of the noise. Non-white noise driving functions of this nature are generated by driving a shaping filter with white noise having an autocorrelation function

$$R'(\tau) = 2\nu\beta^2\delta(\tau) .$$

The shaping filter is shown in Figure 2.

No thermal models are considered for the inertial instruments.

3.1.2 Accelerometers

Errors which are characteristic of the accelerometers in this analysis are bias, random, scale factor and non-orthogonalities. Bias is defined as day-to-day or long-term repeatability and is assumed constant over any mission. Scale-factor error refers to calibration errors and long-term stability of the accelerometer sensing mechanism (including electronics). This error is assumed constant for the flight duration. Random is a short-term change due to shifts in the accelerometer bias and scale factor errors. Non-orthogonalities are composed of two basic errors – mechanical misalignments in mounting the accelerometer and mechanical misalignments internal to the accelerometer.

3.1.3 Gyros

The gyro terminology employed is in general consistent with the standard gyro terminology prepared by the Gyro and Accelerometer Panel of the Electronic Parts Committee, Aerospace Industries Association. The error sources considered are acceleration-insensitive bias and random drifts, acceleration-sensitive drifts (mass unbalance and compliance or anisoelasticity), torquer scale-factor errors and mechanical misalignments.

The bias error is assumed constant over any mission, since it is defined as the day-to-day or turn-on repeatability. The random drift is described by an exponential autocorrelation function and does change over any given mission. Torquer scale-factor errors are due primarily to electronics and calibration errors and remain constant over the mission. Mechanical misalignments are composed of gyro-mounting errors and internal gyro non-orthogonalities.

The differential equation describing the gyro-noise model to be used in the Kalman filter is:

$$\dot{\epsilon} = DD \times \epsilon + CKD \times \delta_s^{\tau} ,$$

where

$$\delta_s^{\tau} = \begin{cases} 1 & \text{if } \tau = s \\ 0 & \text{if } \tau \neq s \end{cases}$$

$$DD = \frac{1}{\tau_g}$$

τ_g = correlation time of gyro noise

$$CKD^2 = \frac{\epsilon^2 \times DD}{2}$$

ϵ = gyro random-drift-rate amplitude .

3.1.4 Doppler-Error Model

The Doppler-error model simulated is composed of several error types, which are shown in Figure 3. The errors considered are bore-sight, scale factor, high frequency electronics noise, low-frequency terrain (overwater) noise, conversion biases and flexure errors. The basic differential equation describing the Doppler noise model is

$$\delta\dot{V}_D = BB \times \delta V_D + CKA \times \delta_s^{\tau}$$

$$\delta_s^{\tau} = \begin{cases} 1 & \text{if } \tau = s \\ 0 & \text{if } \tau \neq s \end{cases}$$

$$BB = \frac{-1}{\tau_D}$$

τ_D = Doppler correlation time.

Also, shown in Figure 4 is the resolution of the Doppler velocity through platform heading and drift angle into platform axis components. The symbols δV_x and δV_y represent the differences between platform velocities (V_{px}, V_{py}) and the respective Doppler velocities.

3.2 IMU Error Model

The IMU error model is shown in Figure 5 with each axis decoupled pictorially but functionally coupled. The level channels are completely shown in the figure with only the vertical gyro. The vertical loop is not presented because a conventional third-order baro-inertial loop is assumed; hence, the vertical-velocity errors are small.

Functions describing each of the errors are presented and discussed below, and terminology is consistent with that employed in most guidance texts and that specified by the Gyro and Accelerometer Panel of the Electronic Parts Committee, Aerospace Industries Association.

Coriolis Errors

$$\delta C_X = -\delta(\rho_Z + 2\Omega_Z)V_{PY} - (\rho_Z + 2\Omega_Z)V_Y ,$$

$$\delta C_Y = \delta(\rho_Z + 2\Omega_Z)V_{PX} + (\rho_Z + 2\Omega_Z)V_X .$$

Cross-Coupling Acceleration Errors

$$f_1(\vec{A})_X = -\phi_Y A_Z + \phi_Z A_Y$$

$$f_1(\vec{A})_Y = +\phi_X A_Z - \phi_Z A_X .$$

Accelerometer Non-orthogonalities

$$f_2(\vec{A})_X = -\mu_2 A_Y + \sigma_2 A_Z$$

$$f_2(\vec{A})_Y = \mu_1 A_X - \sigma_1 A_Z .$$

Accelerometer Noise and Bias

$$\beta_X^2 e^{-\nu|\tau|} + \nabla_X$$

$$\beta_Y^2 e^{-\nu|\tau|} + \nabla_Y .$$

Accelerometer Scale-Factor Error

SF_{AX} = X accelerometer scale-factor error

SF_{AY} = Y accelerometer scale-factor error

A_X , A_Y are true accelerations.

Computational-Error Rates

$\delta\rho_X = -V_Y/R$

$\delta\Omega_Y = +\Omega \sin\phi_T \delta\phi_T$

TFN = Gyro torquing term from filter

$\delta\rho_Y = V_X/R$

$$\delta\rho_Z = \tan\phi_T \frac{V_X}{R} + \frac{V_{XT}}{R} \sec^2\phi_T \delta\phi_T$$

$$\delta\Omega_Y = \Omega \cos\phi \, \delta\phi_T = \delta\Omega_Z .$$

Gyro Noise and Bias

$$\gamma_X^2 e^{-\ell|\tau|} + \epsilon_X$$

$$\gamma_Y^2 e^{-\ell|\tau|} + \epsilon_Y$$

$$\gamma_Z^2 e^{-\ell|\tau|} + \epsilon_Z .$$

Gyro-Scale-Factor Error

SF_{YX} = scale-factor error in X gyro

SF_{YY} = scale-factor error in Y gyro

SF_{YZ} = scale-factor error in Z gyro .

Gyro Cross-Coupling Rates

$$\varepsilon_1(\vec{\omega})_X = -\omega_{YT}\phi_Z + \omega_{ZT}\phi_Y$$

$$\varepsilon_1(\vec{\omega})_Y = -\omega_{ZT}\phi_X + \omega_{XT}\phi_Z$$

$$\varepsilon_1(\vec{\omega})_Z = -\omega_{XT}\phi_Y + \omega_{YT}\phi_X .$$

Gyro Non-orthogonalities

$$h_1(\vec{R})_X = k_4 R_Y - k_5 R_Z$$

$$h_1(\vec{R})_Y = k_1 R_Z - k_3 R_X$$

$$h_1(\vec{R})_Z = k_5 R_X - k_6 R_Y .$$

Gyro Acceleration-Sensitive Errors

$$f_3(\vec{A})_X = - A_Y C_3 + A_X C_4 - A_X A_Y C_{22}$$

$$f_3(\vec{A})_Y = A_Y C_1 - A_X C_2 + A_X A_Y C_{22}$$

$$f_3(\vec{A})_Z = - A_X C_5 + A_Z C_4 + A_X A_Z C_{22} .$$

3.3 Environmental Models

Observations made by the airborne Doppler sensor are relative to the sea, which contaminates the observation with three basic error components. These components fall in three categories: white noise, short-term and long-term correlated noise with autocorrelation functions of the form described in Section 3.1 (Figure 2). The white-noise component results from white caps. The short-term correlated sea current can be characterized by correlation distances of 18-20 nautical miles with nominal amplitudes of 0.4 to 1.0 knots. The long-term correlated noise has nominal amplitudes of 3 knots with correlation distances on the order of 500 nautical miles.

4. STATE VARIABLE-SELECTION-AND-COVARIANCE-MATRIX PROPAGATION

4.1 State Variable-Selection-and-Defining Equations

Due to physical constraints, such as computer program size and processing time, a fully optimized Kalman filter mechanization is virtually impossible for an airborne computer. Hence, several trade-offs must be made in selecting a suboptimal filter which will offer the advantages available from an optimal filter and, yet minimize the drawbacks. The selection of the state variables which adequately describe the system is the most prominent area for trade studies with computer restrictions in mind. The state variables should include all the quantities required to describe the condition of the system and, when possible, its significant error contributors, since these error contributors can be reduced with the aid of the optimal filter. The degree of optimality is dependent on several factors such as accuracy in the description of the error model, processing techniques and selection of the state variables. The states of the inertial system which improve the performance accuracy when accurately modeled are: position, velocity, platform tilts, and platform heading. System performance accuracy is improved considerably, also, if the principal contributors to the errors in each of the above states can be modeled. The inertial system as presently modeled for the alignment and navigation modes has 41 dominant-error sources not including the effects due to aid sensors, computer or environment. Because many of these errors contribute only second-order effects (even in a conventionally-damped system), it is meaningless to include these as state variables. This reduces the number of significant error sources to basically the bias and random error components of the gyros and accelerometers.

Prior to the final selection of the states, one must also examine the aid sensors available to the system. If the errors induced by utilization of the aid sensors are significant as compared to the state variable candidates mentioned, then the aid-sensor contributions must also be considered as possible candidates. For the present application, the Doppler performance specifications are assumed sufficiently large, such that the Doppler error is one of the primary error contributors to the system performance accuracy and must be considered along with the gyros and accelerometers.

In performing the final system-trade studies for state variable selection, the inertial sensor and aid-sensor-error sources were weighed against each other. Although the level-accelerometer bias and random errors are primary contributors in system accuracy, the level accelerometers were eliminated as state variables since the reference velocity error was so large in comparison to the inertial system error attributable to the accelerometers. In fact, in the present mechanization, the inertial system is used to trim the Doppler. As a result of the trade studies, the level-gyro drifts and two Doppler-velocity states were added.

Several other errors were considered as possible state variables. These were vertical gyro drift, gyro torquer scale factors, accelerometer scale factors, and vertical accelerometer bias. In the vertical channel, the accelerometer is not compensated by the filtering technique; however, the vertical-channel error is minimized by the baro-inertial-loop mechanization. The azimuth-gyro drift is a large contributor to the system position error; however, to explicitly model this as a state did not improve system accuracy to the extent to warrant mechanizing an additional state. Instead, the azimuth gyro is modeled as a "pseudo-state" by incorporating its approximate effects as a noise on the platform heading error equation. Similar rationale was employed in eliminating the level accelerometer errors as states. In the interest of maintaining the computational-load minimal and in the trade-off of computer capacity required versus improvement in navigation accuracy, the remainder of the error sources were eliminated.

Formally, Litton has selected eleven state variables to model the system:

ϕ , $\delta\lambda \cos\phi_T$ — latitude and east position errors, respectively

δV_N , δV_E — platform north and east axis velocity errors, respectively (V_x,V_y)

ϕ_E , ϕ_N , ϕ_z — angular errors in platform tilt (ϕ_E and ϕ_N) and azimuth angle, respectively

ϵ_E , ϵ_N — platform level-gyro-drift rates, east and north (ϵ_x and ϵ_y) .

The state-variable-defining equations corresponding to these states for the north-seeking coordinate system are:

$$\delta\dot{\phi}_T = \delta V_N / R_M$$

$$\delta\dot{\lambda}\cos\phi_T = \frac{\delta V_E}{R_N} + \frac{V_E}{R_N}\tan\phi_T\,\delta\phi - \frac{V_N}{R_M}\tan\phi\ (\delta\lambda\cos\phi_T)$$

$$\delta\dot{V}_N = \frac{g}{R}\phi_E - \frac{A_E}{R_N}\phi_z - (\rho_y + 2\Omega_y)\frac{\delta V_E}{R_N} - \left(\frac{V_N}{R_N}\tan\phi\right)\frac{\delta V_E}{R_N} - \left(2\Omega\frac{V_E}{R_N}\cos\phi_T + \frac{V_E^2}{R_N^2}\sec^2\phi_T\right)\delta\phi + CK6 \times \delta\bar{V}_N$$

$$\delta\dot{V}_E = \frac{A_N}{R_M}\phi_z - \frac{g}{R}\phi_N + (\rho_y + 2\Omega_y)\frac{\delta V_N}{R_M} + \frac{V_N}{R_N}\tan\phi_T\frac{\delta V_E}{R_N} + \frac{V_N}{R_M}\left(\frac{V_E}{R_N}\sec^2\phi_T + 2\Omega\cos\phi_T\right)\delta\phi + CK6 \times \delta\bar{V}_E$$

$$\dot{\phi}_N = \epsilon_N + \omega_z\phi_E - \omega_y\phi_z - \Omega_y\delta\phi + \frac{\delta V_E}{R_N}$$

$$\dot{\phi}_E = \epsilon_E + \omega_y\phi_N - \omega_x\phi_z - \frac{\delta V_N}{R_M}$$

$$\dot{\phi}_z = \omega_x\phi_E - \omega_x\phi_N + \Omega_L\,\delta\phi + \frac{V_E}{R_N}\sec^2\phi_T\delta\phi + \tan\phi_T\frac{\delta V_E}{R_N} + EN77 \times \delta\bar{\phi}_z$$

$$\dot{dD}_x^b = -\dot{\theta}D_y$$

$$\dot{dD}_y^b = \dot{\theta}D_x$$

$$\frac{dD_x^s}{dt} = w_{Dx} - K_1 D_x^s = CKA \times \delta\bar{D}_x + BB \times D_x^s$$

$$\frac{dD_y^s}{dt} = w_{Dy} - K_1 D_y^s + CKA \times \delta\bar{D}_y + BB \times D_y^s$$

$$\epsilon_y = \epsilon_N = DD \times \epsilon_N + CKD \times \delta\bar{\epsilon}_N$$

$$\epsilon_x = \epsilon_E = DD \times \epsilon_E + CKD \times \delta\bar{\epsilon}_E .$$

where

t = time

ω_x , ω_y , ω_v = platform spatial rates

g = gravity

Ω_L = level component of earth rate ($= \Omega\cos\phi_T$)

Ω_E , Ω_V = vertical component of earth rate ($\approx \Omega \sin \phi_T$)

V_N , V_E = platform component of craft velocity

A_N , A_E = north and east components of platform acceleration, respectively.

(Note, this can also be computed by $\frac{\int A_N dt}{\Delta t} = \frac{V_N}{\Delta t}$, for small Δt .)

R_N , R_M = radii from center of earth.

Superscripts H and S refer to hardware and sea components, respectively, and the w_i terms are the elements of the internally generated noise matrix. The last four defining equations define the Doppler and gyro states. Note that a first order lag simulation model is used. From the above equations the $\underline{F}$, $\underline{G}$ and $\underline{w}$ matrices referred to in Kalman's works can be written.

The $\underline{X}$, $\underline{F}$, $\underline{G}$ and $\underline{w}$ matrices are shown in Table II. The state variable matrix is comprised of the eleven states and $\underline{F}$, $\underline{G}$ and $\underline{w}$ matrices can be written by inspection of the state variable defining equations.

TABLE II

$\underline{X}$, $\underline{F}$, $\underline{G}$, and $\underline{w}$ Matrices

	$\delta\phi_T$	$\delta\lambda \cos\phi t$	$\delta V_N/R_M$	$\delta V_E/R_N$	ϕ_N	ϕ_E	ϕ_Z	D_x	D_y	ϵ_y	ϵ_x		
$\delta\dot{\phi}_T$	0	0	1	0	0	0	0	0	0	0	0	$\delta\vartheta$	0
$\delta\dot{\lambda}\cos\phi_T$	$\frac{V_E}{R_N}\tan\phi_T$	$\frac{-V_N}{R_M}\tan\phi_T$	0	1	0	0	0	0	0	0	0	$\delta\lambda\cos\phi_T$	0
$\delta\dot{V}_N$	$-2\Omega\frac{V_E}{R_N}\cos\phi_T$ $-\frac{V_E^2}{R_N}\sec^2\phi_T$	0	0	$-(\rho_v+2\Omega_v)$ $-\frac{V_E}{R_N}\tan\phi_T$	0	g/R	$-\frac{A_E}{R_N}$	0	0	0	0	$\frac{\delta V_N}{R_M}$	CK6 x $\delta^T_{V_x}$
$\delta\dot{V}_E$	$\frac{V_N}{R_M}\frac{V_E}{R_N}\sec^2\phi_T$ $+2\Omega\cos\phi_T$	0	$(\rho_v-2\Omega_v)$	$\frac{V_N}{R_M}\tan\phi_T$	-g/R	0	$\frac{A_N}{R_M}$	0	0	0	0	$\frac{\delta V_E}{R_N}$	CK6 x $\delta^T_{V_y}$
$\dot{\phi}_N$	$-\Omega_v$	0	0	1	0	$-\omega_v$	ω_x	0	0	0	1	ϕ_N	0
$\dot{\phi}_E$	0	0	-1	0	ω_x	0	$-\omega_y$	0	0	1	0	ϕ_E	0
$\dot{\phi}_Z$	$\Omega_L+\frac{V_E}{R_N}\sec^2\phi_T$	0	0	$\tan\phi_T$	$-\omega_x$	ω_v	0	0	0	0	0	ϕ_Z	EN 77 x $\delta^T_{\phi_Z}$
$\dot{D}_x$	0	0	0	0	0	0	0	BB	$-\theta$	0	0	D_x	CKA x $\delta^T_{D_x}$
$\dot{D}_y$	0	0	0	0	0	0	0	θ	BB	0	0	D_y	CKA x $\delta^T_{D_y}$
$\dot{\epsilon}_x$	0	0	0	0	0	0	0	0	0	DD	0	ϵ_x	CKD x $\delta^T_{\epsilon_x}$
$\dot{\epsilon}_y$	0	0	0	0	0	0	0	0	0	0	DD	ϵ_y	CKD x $\delta^T_{\epsilon_y}$

$|\dot{x}|$ $|F|x$ $|G|w$

Note that Doppler-error components are chosen rather than ground-speed and drift-angle errors. This was done so that the cross-channel-correlation computation formulas could be used to full advantage. Also, X- and Y-axis position errors are state variables but system latitude and longitude are updated in the inertial corrections.

4.2 Covariance Matrix Computation

The covariance matrix and Kalman gain computations form the bulk of the computer program to be mechanized. This section addresses the computations required to propagate the covariance matrix. As previously stated, the covariance matrix is a running estimate of the system performance and hence the ability to predict and correct the system is totally dependent on the exactness of this matrix. Regarding one aspect in programming the covariance matrix, the technique employed in propagating the matrix over time and computational exactness must be investigated.

Litton has chosen to use direct integration of the covariance matrix which was derived on a continuous basis. In keeping with the accuracy requirements, a higher-order-integration algorithm was found necessary. Litton is currently using a product of variables-type algorithm. This algorithm is presented in more detail in the following section.

In the interest of minimizing computer requirements and maintaining good performance accuracy, trade-offs performed on the various elements of the covariance matrix resulted in segmenting the matrix into two major computational parts. These major subdivisions were based on the rate of change of the covariance matrix elements. As a result, the covariance matrix is composed of "fast" loops and "slow" loops. The significance of these loops are that the fast loop components vary quite rapidly and possess a wide dynamic range compared to the slow loop computations. Hence, direct integration of the slow loops at the same rate as the fast loops would result in a more coarse estimation in that round-off effects would be more pronounced. Therefore, to maintain computational precision, the slow loop processing interval is approximately 80 times the fast loop processing interval.

Figure 6 illustrates the notational convention and the members of the covariance matrix that are computed. Where blanks appear, the matrix elements are computed by their closed form approximations if these elements are needed. The notation used is a capital letter A, R or S to indicate that the element is a member of that submatrix. The subscripts on these elements define the state variables which are statistically correlated by that function, e.g., $S_{\dot{x}\phi_y}$ is the S-matrix covariance of x-axis velocity ($\dot{x}$) and y-axis platform tilt (ϕ_y).

The covariance-matrix elements which are left blank in Figure 6 are not computed explicitly. If they are needed in propagating other elements, they are computed from closed form, cross-channel correlation formulas†. This allows computing of only 42 of the 66 elements of the covariance matrix.

4.3 Integration Algorithm (for Filter)

The programmed integration algorithm is one designed to correctly integrate a product of two variables with constant rates of change.

The algorithm should compute, from $t = 0$ to $t = \Delta t$,

$$\Delta J = \int_0^{\Delta t} yx\,dt = \int_0^{\Delta t} (y_0 + \dot{y}_0 t)(x_0 + \dot{x}_0 t)\,dt$$

$$= [y_0 x_0 + (\dot{y}_0 x_0 + \dot{x}_0 y_0)(\Delta t/2) + \dot{y}_0 \dot{x}_0 (\Delta t^2/3)]\,\Delta t .$$

Consider an algorithm which utilizes the computer value for the preceding increment in the computation of the present increment. Assume the algorithm is correct for variables with constant rate, so that the computed value of the preceding increment is equal to

$$\Delta J_{old} = [y_0 x_0 - (\dot{y}_0 \dot{x}_0 + \dot{x}_0 y_0)(\Delta t/2) + \dot{y}_0 \dot{x}_0 (\Delta t^2/3)]\Delta t .$$

Let the algorithm have the form

$$\Delta J = \eta \Delta J_{old} + (1-\eta)(y_{old} + \mu \Delta y)(x_{old} + \mu \Delta x)\,\Delta t ,$$

where η, $(1-\eta)$ weights are implied by the necessity of computing correctly for zero rate variables expressing ΔJ in terms of y_0, x_0, $\dot{y}_0$ $\dot{x}_0$.

$$\Delta J = \eta[y_0 x_0 - (\dot{y}_0 x_0 + \dot{x}_0 y_0)(\Delta t/2) + \dot{y}_0 \dot{x}_0 (\Delta t^2/3)]\Delta t + (1-\eta)(y_0 + \mu \dot{y}_0 \Delta t)(x_0 + \mu \dot{x}_0 \Delta t)\,\Delta t$$

$$= \left[y_0 x_0 + (\mu(1-\eta) - \eta/2)(\dot{y}_0 x_0 + \dot{x}_0 y_0) + \left\{\frac{\eta}{3} + (1-\eta)\mu^2\right\}\dot{y}_0 \dot{x}_0 \Delta t^2\right]\Delta t$$

$$= [y_0 x_0 + C_1(\dot{y}_0 x_0 + \dot{x}_0 y_0) + C_2 \dot{y}_0 \dot{x}_0 \Delta t^2]\Delta t .$$

---

† The closed form relationships utilize invariant products.

For correctness, the linear and quadratic coefficients C_1 and C_2 must be 1/2 and 1/3, respectively. Hence take

$$\mu(1-\eta) - \eta/2 = 1/2$$

$$\eta/3 + (1-\eta)\mu^2 = 1/3 ,$$

which gives

$$\mu = 1/\sqrt{3} = 0.577$$

$$\eta = \frac{2-\sqrt{3}}{2+\sqrt{3}} = 0.0718.$$

An approximate solution

$$\eta = 2^{-4} , \qquad \mu = (2^{-1} + 2^{-4})$$

implies

$$C_1 = \frac{1}{2}\left(1-2^{-7}\right), \qquad C_2 = \frac{1}{3}\left[1 - \frac{3}{64}\left(1 + \frac{1}{64}\right)\right] ,$$

which might be adequate if desired.

The general form of error-propagation equations, including the invariant form, involves the summation of disparate terms to obtain each integral increment of a basic variable during a program cycle.

The computation scheme as stated implies the storage of 42 quantities each for old values of integral increments, for interpolated values, and for integrated state variables. This totals 126 words of storage for the integration algorithm. This is 84 words more than the simplest natural algorithm.

Consider algorithm computational requirements for initialization and inertial system correction where the old estimates of the integral increments *may not be available*. One approach is to add a cycle of the increment computation where there is no accumulation of the integral, and thereafter return to the normal program for the first iteration, using $\mu = 1/4$, $\eta = 2/3$. System corrections with the Doppler radar are programmed as a quasi-continuous process in which their effects in error estimation are part of the steady-state integration process. The position fixes, however, require an initialization provision to account for step changes in the derivatives of A and R matrix variables. This may be accomplished with the pseudo-cycle computation with $\mu = 1/4$, $\eta = 2/3$ or, if it is found that initialization after dead reckoning with Doppler can be tolerated without a pseudo-cycle, the pseudo-cycle can be obviated altogether provided the steps in a pseudo-old-increment of A , R for position fix are computed. These increments are computed by formulas derived by incrementing the difference equations for error growth of Δ , R , for increments of A on fixing. The additional computations involve 16 multiplications.

4.4 Implementation of Invariants-Scaling Technique

In implementing a scaling algorithm, "real-world" constraints dictate several of the limiting conditions mechanized. For instance, separate scales are needed for various mission phases; however, to fix a distinct set of scale factors obviously imposes system constraints and these constraints lead to the development of an adaptive rescaling module. The rescaling module would be exercised at the beginning of each filter-updating cycle. The successful implementation of this rescaling technique relies on the regular behavior of the elements of the covariance matrix. A large change could occur in these elements if, for instance, a position observation is made. If the incremental change in these elements is too large to be accommodated in one filter cycle, a "segmented" fix can be easily mechanized so as to avoid register overflow or constrain maximum information by left-shifting if the quantity being updated is rapidly decreasing.

In implementing the rescaling module, the philosophy is to test the diagonal elements of the covariance matrix and to constrain the scaled quantities to lie within a specified range of the register by controlling the scale factors. With the chosen constraints mechanized, the rescaling loop shifts left or right by one bit if the rescaling tests fail.

4.4.1 *Initialization Scaling*

For the scaling of the program parameters, all of the scale factors are formed from basically six factors, SF(1) through SF(6). The numerical value in parentheses refers to the type of parameter being scaled, e.g.,

SF (1) position

SF (2) velocity

SF (3) tilt

SF (4) gyro drift

SF (5) Doppler

SF (6) azimuth .

At time zero, values for each of the scale factors are selected.

The elements of the covariance matrix can be scaled by implementing the equivalent to the following FORTRAN DO loop.

```
DO 44 I = 1, 6
DO 44 J = 1, 6
SFP(I,J) = 1/SF (I)/SF (J)
 X(I,J) = SFP (I,J)* X(I,J)
44Y(I,J) = SFP (I,J)* Y(I,J) .
```

Scaling of special functions at time zero is performed separately.

All parameters must have an assigned scaling before their insertion into the program; however, this is not a drawback, since the chosen parameters in the filter program can be *rescaled* at will before the actual filter initialization by implementing the Litton technique for adaptive rescaling, e.g. the quantities SF (J) , J = 1, 6 are rescaling factors for changes in the covariance matrix elements X(I,J) , Y(I,J) and for the quantities DX(I,J) and DY(I,J) , where DX(I,J) and DY(I,J) are the propagation increments to covariance matrix elements X(I,J) and Y(I,J) , respectively. Note: indexing is used in programming the covariance matrix in order to minimize computer instructions. This does result in some processing time penalty.

4.4.2 *Rescaling Module*

The Litton adaptive rescaling module is to be tested at the beginning of each complete pass through the filter. The philosophy is to scale both channels identically; therefore, the maximum values of the corresponding diagonal elements (designated by AMM(J)) of the covariance matrix must be determined. Scaling constraints on the elements are set so that all state variables lie between 1/16 and 1/4 of full scale except for azimuth, which is scaled between 1/4 and 1/2, since division by azimuth is necessary in the closed form cross-channel correlations. With these constraints mechanized, the remainder of the scaling loop is necessary to scale the quantities indicated by shifting either left or right one bit, depending on the shifting parameter. Figure 7 is a flow chart of the rescaling program.

5. SIMULATION RESULTS

This section contains simulation results for a test mission. The assumptions are that the system was ground-aligned using conventional techniques and the filter was initialized at the instant the vehicle began moving. In order to illustrate the potential accuracy improvement by using a filter, the system and filter were run "open-loop", i.e., the filter was run as an estimator, not a controller, such that normalized rather than dependent errors could be compared.

In the system-performance curves (Figures 8 through 11) the net resulting error is the difference between the estimated and actual system errors. The curves presented in Figures 8 through 11 are based on a profile which contains several velocity fluctuations; starting and stopping; during the profile; 45° latitude; and error sources similar to an LN-15 type platform (the system errors are larger than that expected from an actual LN-15 for a similar application; however, for the simulated mission, the full knowledge of error-source correlations was not employed).

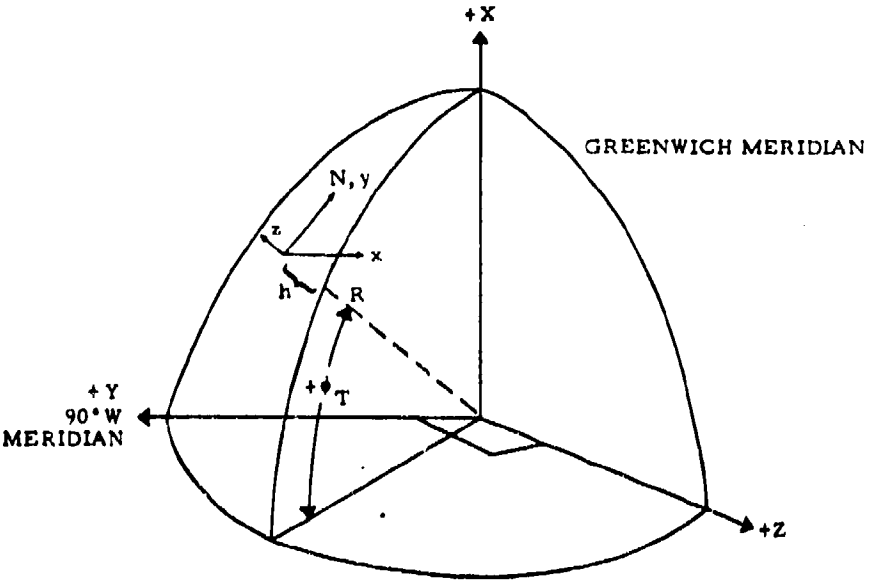
REFERENCES

1. Kalman, R. E. — *New Methods and Results in Linear Prediction and Filtering Theory*, RIAS Technical Report 61-1.

2. Bambrook, H. et al. — *LORAN D/Inertial Navigation System Integration*. Litton Publication AQ22009M65 (Air Force contract No. AF-33(657)-14309), January 1966.

Φ_T = LATITUDE

λ_T = LONGITUDE



X AXIS COINCIDENT WITH TERRESTRIAL POLE
Y AXIS IN THE EQUATORIAL PLANE AT THE 90°W MERIDIAN
Z AXIS IN THE EQUATORIAL PLANE AT GREENWICH
x, y, z – PLATFORM COORDINATE AXES

Fig. 1 Earth and navigation system coordinate frames

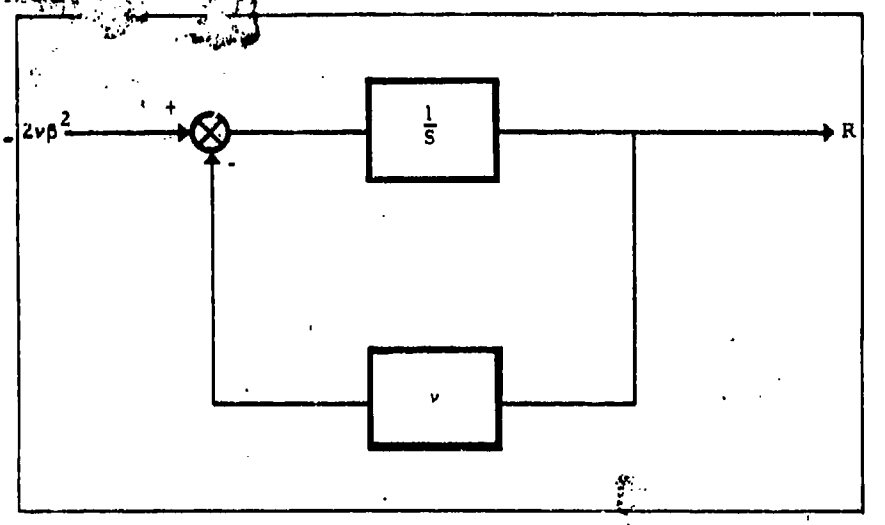


Fig. 2 Exponential noise-shaping filter

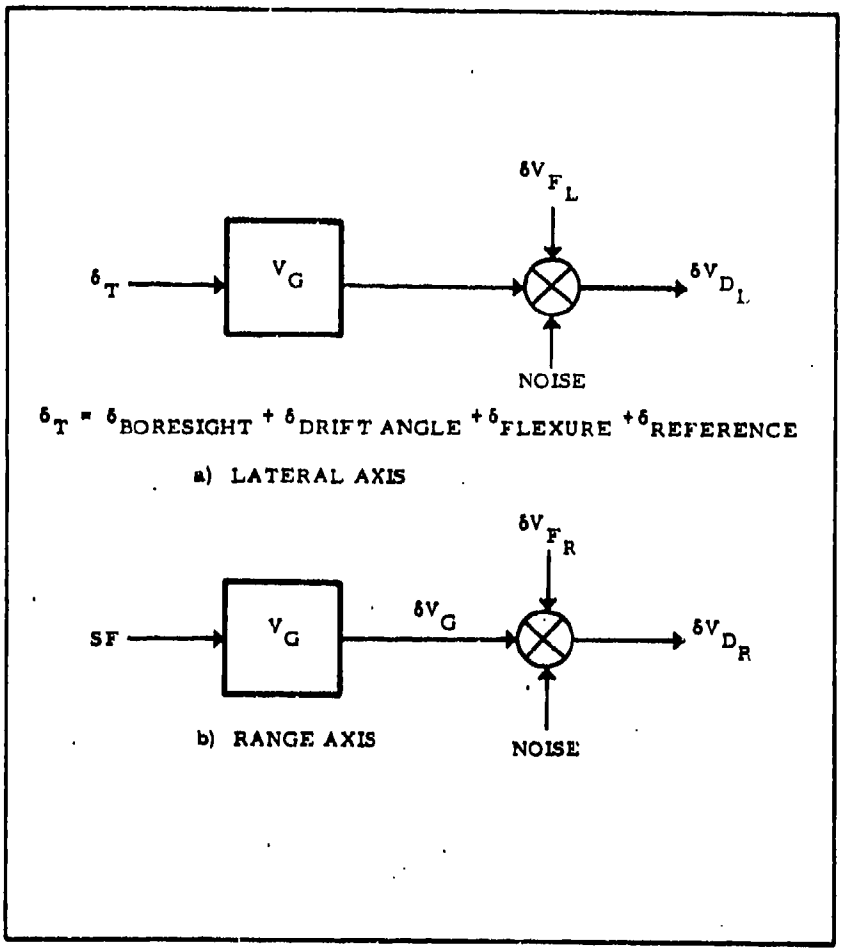


Fig. 3 Doppler-error model

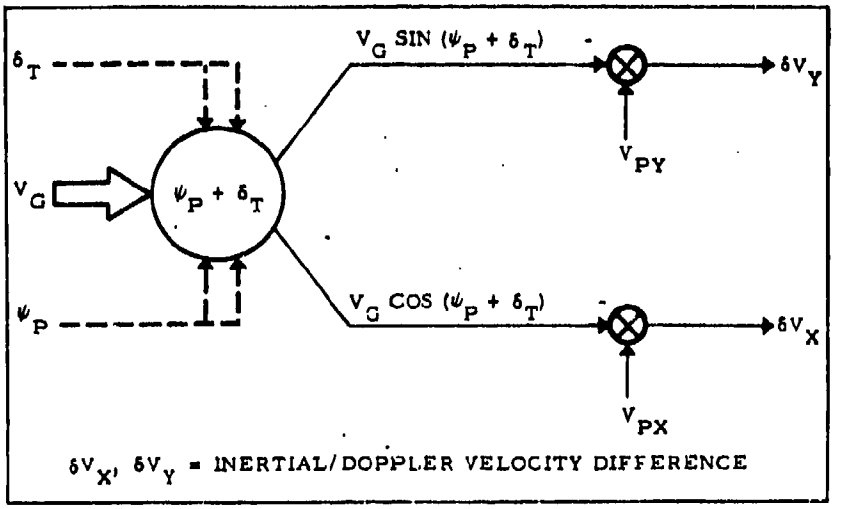


Fig. 4 Velocity-difference-signal formation

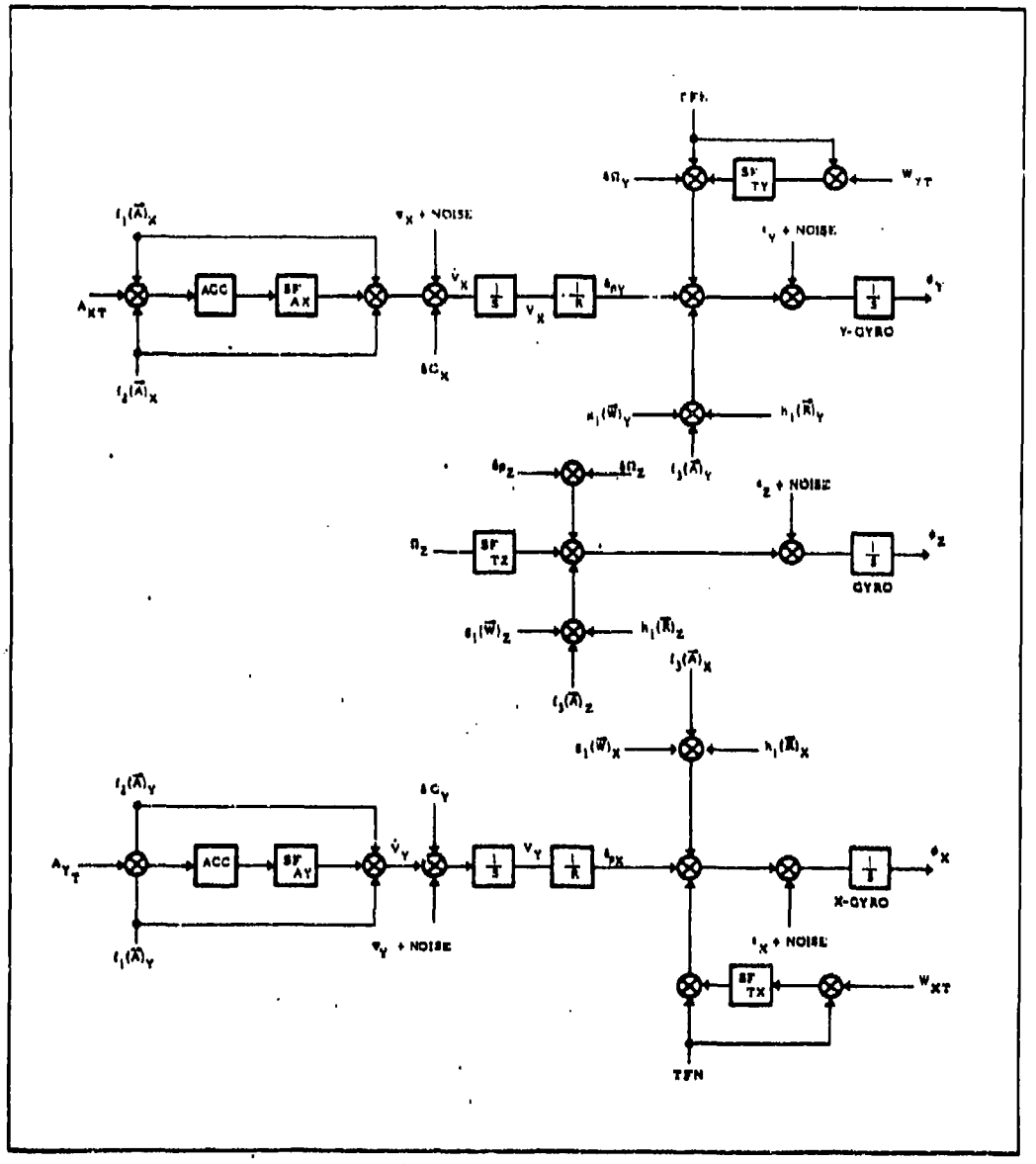
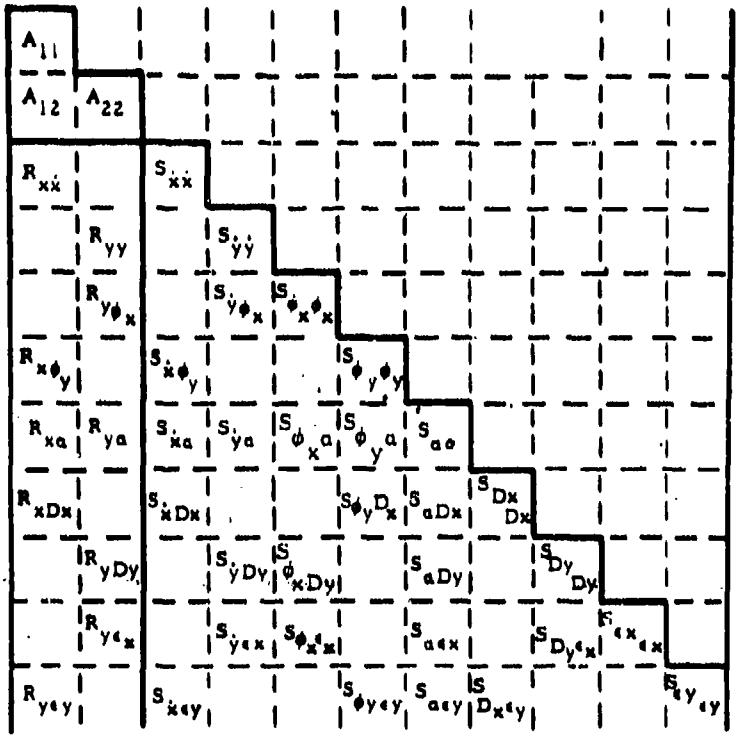


Fig.5 IMU error model block diagram



A-ELEMENTS: POSITION ELEMENTS WITH POSITION ELEMENTS

R-ELEMENTS: NON-POSITION ELEMENTS WITH POSITION ELEMENTS

S-ELEMENTS: NON-POSITION ELEMENTS WITH NON-POSITION ELEMENTS

Fig. 6 Notation convention used in partitioned covariance matrix

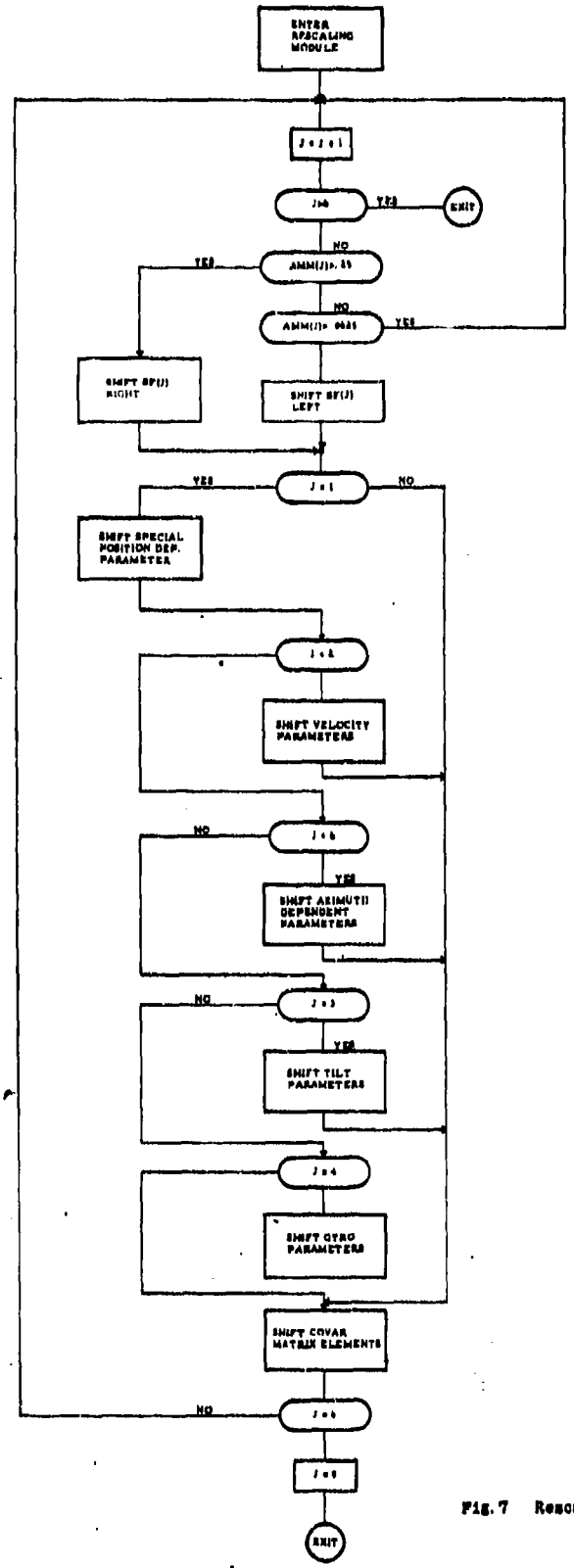


Fig. 7 Rescaling module flow diagram

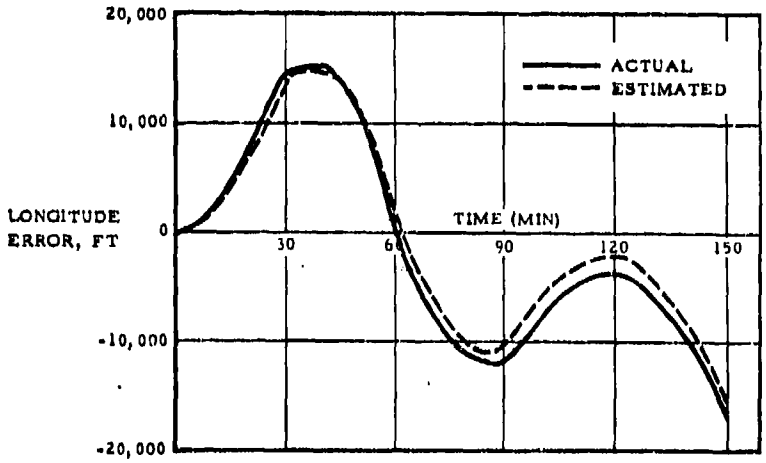


Fig.8 Longitude error versus time

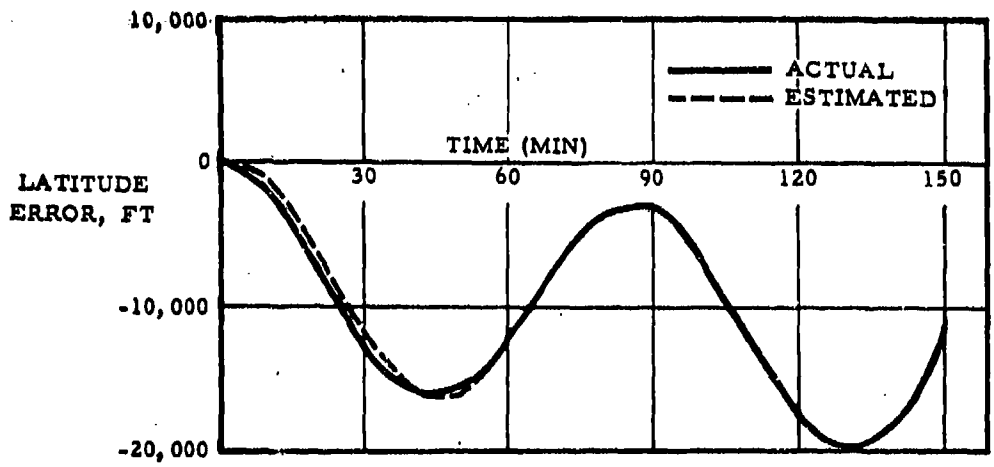


Fig.9 Latitude error versus time

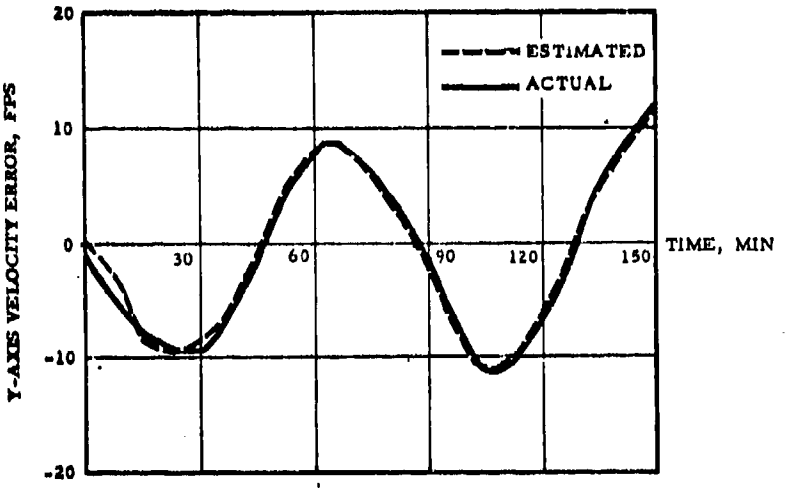


Fig. 10 X-axis velocity error versus time

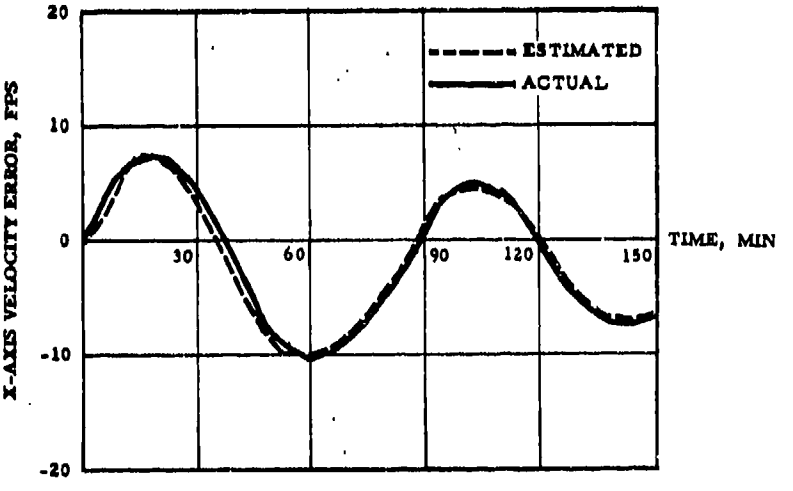


Fig. 11 Y-axis velocity error versus time

CHAPTER 17 - MARINE APPLICATIONS OF KALMAN FILTERING

by

J. Holdsworth and J. Stolz

Advanced Marine Technical Division
Litton Industries
9920 W. Jefferson Blvd
Culver City, California, USA

CHAPTER 17 - MARINE APPLICATIONS OF KALMAN FILTERING

J. Holdsworth and J. Stolz

1. INTRODUCTION AND DESCRIPTIONS OF THE PROBLEM

This chapter is concerned with the application of Kalman filtering techniques to estimating the equations of motion of submerged objects in the ocean. Applications can be made to monitor the movement of submarines as well as the movement of schools of large fish or submarine mammals. Another interesting application is the optimal control of cabled communication, wire-guided submarine devices. Here the estimation of the remote control device's track would be useful to evaluate if a device has sufficiently searched a particular topographical area of interest. Also of interest is the situation where a single command and control ship monitors and controls the operation of some activity in the ocean area of its immediate vicinity. The Kalman filter can be used to track the activity of all objects in a predetermined area. Thus the Kalman filter can be used to guide the movement of scuba divers in undersea exploration or in the harvesting of food on the ocean floor.

The ocean is the medium in which the Kalman filter's estimation process has to operate. Communications between two objects separated by several thousand yards is obtained by sound. The speed of sound in water is nominally 4800 ft/sec and it is primarily a function of water temperature, pressure, and salinity and is given by the empirical formula

$$C = 4422 + 11.25t - 0.045t^2 + 0.0182d + 4.3\,(s-34)\,,$$

where

C = velocity of sound (ft/sec)

d = water depth (ft)

t = temperature of water at d ($^{\circ}$F)

s = salinity (parts per thousand).

The propagation path of sound in water is quite different from the paths of electromagnetic radiations used by radar tracking systems. The propagation path of sound is quite non-linear and its general nature can be described by ray theory[1]. The different modes of propagation are various combinations of direct path, surface channel, bottom bounce and convergence zone. The direct path mode is where the sound path that connects two objects is not reflected by any acoustical boundary. The surface channel mode is where the sound path connecting two objects is reflected at least once by the physical boundary between the water and the atmosphere. The bottom bounce mode is where the sound path connecting two objects is reflected at least once by the boundary between the water and the ocean floor. The convergence zone mode is where the sound path connecting two objects is refracted due to the sound ray "vertexing". This occurs when the ray path attains a horizontal tangency at a certain ocean depth[1].

Thus, the total transit time of sound from one object to another is a function of the propagation path and the velocity of sound in the water. The different propagation modes are illustrated in Figure 1.

The underwater systems that either communicate by sound or monitor sound sources are called sonars. The sonar system of interest to us is the echo-ranging sonars, which operate by monitoring sound reflected by an underwater object that was originally an input to the water by the receiving ship. The measurements taken by the echo-ranging sonar are range, bearing and sometimes range-rate. The range information (R) is calculated by

$$R = \frac{t}{2} V_s\,, \tag{1.1}$$

where t is the total travel time of sound and V_s is the average speed of sound in water. The bearing information (B) is obtained by measuring the direction of the reflecting underwater object. The range-rate ($\dot{R}$) information is obtained from the relationship

$$f = 0.69\, f_s \dot{R}\,, \tag{1.2}$$

where f_s is the frequency of the transmitting source (kc/s), f (c/s) is the observed change in frequency when the sound signal is returned and 0.69 is a Doppler constant for echo transmission (c/kt kc). The measured observations ($B, R, \dot{R}$) are corrupted by several types of noise sources. The structure of the noise is very complex[1] and will not be dealt with here. The accuracies to which the observations B, R, $\dot{R}$ can be measured are a function of beam-width, pulse-length, frequency discrimination, and signal-to-noise ratio measuring characteristics of each sonar. For the purposes of this analysis, the additive noise on bearing information can be thought of as a stationary zero mean, gaussian process where the standard deviation of the error is about 0.5 degrees. The range information can also be thought of as a gaussian process where the variance changes linearly with relative range (1%) (distance between one object and another). The time difference between consecutive samples is long enough

such that autocorrelation of bearing and ranges is insignificant. Also the cross correlation between ranges and bearings is insignificant. The accuracies of the range-rate information are about 6 knots.

2. APPLICATION OF KALMAN FILTERING TO THE MARINE TRACKING PROBLEM

2.1 General

We shall now briefly describe the basic equations which are required to perform the tracking process. The tracking process is basically a Kalman filter operating on the observations of range, bearing, and Doppler. The above observational set corresponds to the active sonar mode of operation. The equations for the passive sonar mode are different and will not be considered here. The observations have first to be corrected to give true bearing, horizontal range and range rate. This involves pre-processing the observations to account for the sound propagation mode. The observation pre-processing can be accounted for directly in the Kalman filter itself, but this would greatly complicate the formulation. Our treatment will be brief and descriptive instead of detailed and rigorous. Derivations of the equations from a minimum mean square prediction error criterion are well known[2] and are detailed in previous sections of this publication.

2.2 Kalman Filter Equations

The basic component of the estimation process is a state vector whose value at some time t_k is given by

$$\bar{X}(k) = \begin{bmatrix} \dot{X} \\ X(k) \\ \dot{Y} \\ Y(k) \end{bmatrix} . \tag{2.1}$$

The components of the target's state vector are the components of the relative position with respect to our tracking vessel and the absolute components of the velocity of the target. The state equation which describes the dynamic development of the target state vector in time may be written

$$\bar{X}(k) = \phi(k)\bar{X}(k-1) , \tag{2.2}$$

where, for our case, the state transition matrix is given by

$$\phi(k) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \Delta t_k & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \Delta t_k & 1 \end{pmatrix} . \tag{2.3}$$

The quantity Δt_k is the time between receipt of the $(k-1)^{th}$ and the k^{th} piece of data (report). If the data rate is constant, then $\Delta t_k = \Delta t$ is constant and the transition matrix $\phi(k)$ described by Equation (2.3) does not depend upon time. The value of Δt is determined *a priori* by a range setting that is well beyond the range of the target. The actual processing of the data will be divided into two parts. The first involves range ($\tilde{R}_k$) and bearing ($\tilde{B}_k$) only. The second involves, in addition to range ($\tilde{R}_k$) and bearing ($\tilde{B}_k$), range-rate ($\tilde{\dot{R}}$) as well. The data vector Z_k' for the first set of observations may be written as follows:

$$Z_k' = \begin{pmatrix} \tilde{R}_k \sin \tilde{B}_k \\ \tilde{R}_k \cos \tilde{B}_k \end{pmatrix} = M'(k)\bar{X}(k) + \epsilon_k' . \tag{2.4}$$

In Equation (2.4), $M'(k)$ is called the measurement matrix, which relates the observed data vector Z_k' to the target state vector $\bar{X}(k)$. In our case the measurement matrix does not depend upon the time t_k and may be explicitly written out as

$$M'(k) = M' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} . \tag{2.5}$$

The term ϵ_k' denotes the 2×1 noise vector due to the noise perturbations on the measured bearing angle $\tilde{B}_k$ and range $\tilde{R}_k$. Assuming that the measurement noise on the range and bearing is uncorrelated and zero mean, it follows that the covariance matrix of our noise vector is approximately given by

$$W_k' = E[\epsilon_k' \epsilon_k'^T] = \left(\begin{array}{c|c} \sin^2\tilde{B}_k\sigma_R^2 + \tilde{R}_k^2 \cos^2\tilde{B}_k\sigma_B^2 & \sin \tilde{B}_k \cos \tilde{B}_k (\sigma_R^2 - \tilde{R}_k^2\sigma_B^2) \\ \hline \sin \tilde{B}_k \cos \tilde{B}_k (\sigma_R^2 - \tilde{R}_k^2\sigma_B^2) & \cos^2\tilde{B}_k\sigma_R^2 + \tilde{R}_k^2 \sin^2\tilde{B}_k\sigma_B^2 \end{array}\right) . \tag{2.6}$$

The target state vector $\bar{X}(k)$ is unknown to us and must be estimated from the observed data vector Z_k' and whatever knowledge we have of the measurement noise added to the observed ranges and bearings. Let $\hat{X}(k-1)$ denote the estimate of the target state vector based upon whatever *a priori* knowledge we have, together with all the observed data obtained up to and including time t_{k-1}. At time t_k a new data vector Z_k' is observed.

In the Kalman filter model, the new updated estimate of the target state vector is recursively determined according to the following equation:

$$\hat{X}(k) = \phi(k)\hat{X}(k-1) + K(k)[Z'_k - M'\phi(k)\hat{X}(k-1)] \quad (2.7)$$

In Equation (2.7), the term $\phi(k)\hat{X}(k-1)$ is the updated prediction of the target state from the dynamic Equation (2.2), based just on the data available at stage $k-1$ and before. Likewise the term

$$M'\phi(k)\hat{X}(k-1) \quad (2.8)$$

is our prediction of the measurement vector Z'_k based on the $(k-1)^{th}$ stage data, so that

$$Z'_k - M'\phi(k)\hat{X}(k-1) \quad (2.9)$$

is the observed discrepancy between the measured data vector at stage k and the predicted data vector based upon prior information.

$K(k)$ in Equation (2.7) is the Kalman gain matrix which weights the observed measurement discrepancy. It is chosen to minimize the mean square error in the estimates of the components of the target state vector. If we let $P(k)$ denote the covariance matrix of the estimate of the target state vector at time t_k, i.e.,

$$P(k) = E[(\hat{X}(k) - \bar{X}(k))(\hat{X}(k) - \bar{X}(k))^T], \quad (2.10)$$

where $E[\,]$ denotes the expected value operation, then it may be shown[8] that the Kalman gain matrix may be written as

$$K(k) = \phi(k)P(k-1)\phi^T(k)M'^T[M'\phi(k)P(k-1)\phi^T(k)M'^T + W'_k]^{-1}. \quad (2.11)$$

Beginning with an initial covariance matrix $P(0)$ of an estimated state vector, it may be shown that the development of the target state covariance matrix, a function of time is obtained from the following recursive equation:

$$P(k) = \phi(k)P(k-1)\phi^T(k) - K(k)M'\phi(k)P(k-1)\phi^T(k). \quad (2.12)$$

When Doppler data is available the procedure, when combined with the foregoing process, is as follows. The computational equations are quite similar for the second set of observations. The state transition matrix $\phi(k)$ becomes

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.13)$$

and the data vector may be written as

$$Z''_k = \begin{pmatrix} \tilde{R}_k \sin \tilde{B}_k + \tilde{R}_k \cos \tilde{B}_k \tilde{B} \\ \tilde{R}_k \cos \tilde{B}_k - \tilde{R}_k \sin \tilde{B}_k \tilde{B}_k \end{pmatrix} = M''(k)\bar{X}(k) + \epsilon''_k. \quad (2.14)$$

In Equation (2.14), $M''(k)$ is given by

$$M''(k) = M'' = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (2.15)$$

and, similar to Equation (2.4), the term ϵ''_k in Equation (2.14) is a 2×1 noise vector due to the noise perturbations on the measured bearing angle $\tilde{B}_k$, range $\tilde{R}_k$, range-rate $\tilde{\dot{R}}_k$ and the support observation $\tilde{\dot{B}}_k$. Since $\tilde{\dot{B}}_k$ is not an explicit observation, it can be synthesized by either of two methods. The term $\tilde{\dot{B}}_k$ can be calculated by

$$\tilde{\dot{B}}_k = (\tilde{B}_k - \tilde{B}_{k-1})/\Delta t \quad (2.16)$$

or by

$$\tilde{\dot{B}}_k = \frac{Y(k)\dot{X}_r - X(k)\dot{Y}_r}{Y(k)^2 + X^2(k)}. \quad (2.17)$$

The components $Y(k)$ and $X(k)$ in Equation (2.17) are obtained from Equation (2.1). The terms $\dot{X}_r$ and $\dot{Y}_r$ are obtained by

$$\left.\begin{aligned} \dot{X}_r &= \dot{X} - \dot{X}_0 \\ \dot{Y}_r &= \dot{Y} - \dot{Y}_0 . \end{aligned}\right\} \quad (2.18)$$

where $\dot{X}_0$ is the motion of the observation vessel in the X-direction and $\dot{Y}_0$ is the motion of the observation vessel in the Y-direction. Also, the range-rate information in the data vector given by Equation (2.14) is relative range-rate, which was obtained by

$$\tilde{\dot{R}}_k = \tilde{\dot{R}}_T - \dot{R}_0 . \quad (2.19)$$

The term $\tilde{\dot{R}}_T$ is the actual observation obtained from the sonar. The term $\dot{R}_0$ is obtained from the observational vessel's sensors. The covariance matrix of the noise vector given by Equation (2.14) may be written as

$$W_k'' = E[\epsilon_k'' \epsilon_k''^T]$$

$$\left(\begin{array}{c|c} \sigma_R^2 \sin^2\bar{B}_k + \sigma_\theta^2 \cos^2\bar{B}_k + \bar{R}_k^2 \cos^2\bar{B}_k\sigma_B^2 + & \cos\bar{B}_k \sin\bar{B}_k[\sigma_R^2 - \bar{R}_k^2\sigma_\theta^2 - \bar{R}_k^2\sigma_\theta^2 + \\ + (\bar{R}_k \cos\bar{B}_k - R_k \sin\bar{R}_k\bar{B}_k)^2\sigma_B^2 & + \sigma_\theta^2(\bar{R}_k^2\bar{B}_k^2 - \bar{R}_k^2)] + \sigma_\theta^2\bar{R}_k\bar{R}_k (\sin^2\bar{B}_k - \cos^2\bar{B}_k) \\ \hline \cos\bar{B}_k \sin\bar{R}_k[\sigma_R^2 - \bar{R}_k^2\sigma_\theta^2 - \bar{R}_k^2\sigma_\theta^2 + & \sigma_R^2 \cos^2\bar{B}_k + \sigma_\theta^2 \sin^2\bar{B}_k\bar{B}_k' + \bar{R}_k^2 \sin^2\bar{B}_k\sigma_B^2 + \\ + \sigma_\theta^2(\bar{R}_k^2\bar{B}_k^2 - \bar{R}_k^2)] + \sigma_\theta^2\bar{R}_k\bar{R}_k (\sin^2\bar{B}_k - \cos^2\bar{B}_k) & + (\bar{R}_k \sin\bar{B}_k + \bar{R}_k \cos\bar{B}_k\bar{B}_k)^2\sigma_\theta^2 \end{array}\right) \quad (2.20)$$

The same process is repeated for Equations (2.7), (2.11) and (2.12), replacing Z_k', M', and W_k' by Z_k'', M'' and M_k''.

The beginning of the Kalman filtering process is given by

$$P(0) = \begin{pmatrix} 5000 & 0 & 0 & 0 \\ 0 & 5000 & 0 & 0 \\ 0 & 0 & 5000 & 0 \\ 0 & 0 & 0 & 5000 \end{pmatrix}, \quad (2.21)$$

where the quantity 5000 is somewhat arbitrary, the important thing being that the diagonal elements of $P(0)$ be large so that the filter will forget the initial values as more data arrive, and

$$X(0) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (2.22)$$

The process can be carried out iteratively for each new set of observational reports.

The variables σ_R, σ_θ, σ_ϕ in Equations (2.6) and (2.20) are determined *a priori* and can be readjusted adaptively if desired. The variable σ_ϕ in Equation (2.20) is calculated by

$$\sigma_\phi = \sqrt{2}\sigma_\theta/\Delta T . \quad (2.23)$$

3. THE MANEUVERING TARGET PROBLEM

3.1 General

In the typical tracking problem the target (submarine object of interest) will not, in general, maintain a policy of a constant course and speed. If the target is a submarine, it will usually perform a sequence of short straight-line maneuvers. If the target is a cabled communication, wire-guided submarine device, the movements are constrained by the topography of the ocean floor. Since the mission usually requires that the target operates in close proximity to ocean floor the motion could be highly nonlinear. In this case, the classical estimation of the target's motion will be corrupted by the nonlinear movement of the target. Thus the problem indicates a need to model directly the uncertainties of the target's motion in the estimation process. The Kalman filter can be adapted to reflect explicitly the uncertainty of the target's state estimates at any time due to the random nonlinearity of the target. This has led to the development of an uncertainty gain matrix U_g.

3.2 The Derivation of the Uncertainty Gain Matrix (U_g)

The state vector of a target with random maneuvering capability is given by

$$\bar{X} = \bar{X}_s + \bar{R}_z . \quad (3.1)$$

where $\bar{X}$, $\bar{X}_s$, $\bar{R}_z$ are 4 x 1 column vectors.

In Equation (3.1), $\bar{X}_s$ is given by

$$\bar{X}_s = \begin{bmatrix} \dot{X} \\ X \\ \dot{Y} \\ Y \end{bmatrix} \quad (3.2)$$

and N_k is given by

$$N_k = \begin{bmatrix} \epsilon_{\dot{x}} \\ \epsilon_x \\ \epsilon_{\dot{y}} \\ \epsilon_y \end{bmatrix} \tag{3.3}$$

In Equation (3.2) the components represent the true motion of the target for the non-maneuvering case and the expected motion for the random maneuvering case. In Equation (3.3) the components represent the deviations about the expected state vector for the maneuvering case. Thus the error covariance matrix P is given by

$$E((\hat{X}-X)(\hat{X}-X)^T) = E((\hat{X}-X_B)(\hat{X}X_B)^T) + E(N_k N_k^T) \tag{3.4}$$

$$P = P' + U_Q . \tag{3.5}$$

In the non-maneuvering case, the uncertainty gain matrix (U_Q) is zero in Equation (3.5) and $P = P'$.

The components are derived as follows:

$$\dot{x} = S \sin C \tag{3.6}$$

$$\epsilon_{\dot{x}} = \epsilon_s \sin C + \epsilon_c S \cos C \tag{3.7}$$

$$\dot{y} = S \cos C \tag{3.8}$$

$$\epsilon_{\dot{y}} = \epsilon_s \cos C - \epsilon_c S \sin C \tag{3.9}$$

$$\epsilon_x = \epsilon_{\dot{x}} \Delta t \tag{3.10}$$

$$\epsilon_y = \epsilon_{\dot{y}} \Delta t \tag{3.11}$$

where S = target speed, C = target course and Δt = times between consecutive data reports.

All the components of N_k are assumed to be uncorrelated, as well as the ϵ quantities. Thus U_Q is a diagonal matrix obtained by calculating the variances of Equations (3.7), (3.9), (3.10) and (3.11).

$$U_Q = \begin{bmatrix} \sigma_{\dot{x}}^2 & 0 & 0 & 0 \\ 0 & \sigma_x^2 & 0 & 0 \\ 0 & 0 & \sigma_{\dot{y}}^2 & 0 \\ 0 & 0 & 0 & \sigma_y^2 \end{bmatrix} \tag{3.12}$$

The components of Equation (3.12) are

$$\sigma_{\dot{x}}^2 = \sigma_s^2 \sin^2 C + \sigma_c^2 S^2 \cos^2 C \tag{3.13}$$

$$\sigma_x^2 = \Delta t^2 \sigma_{\dot{x}}^2$$

$$\sigma_{\dot{y}}^2 = \sigma_s^2 \cos^2 C + \sigma_c^2 S^2 \sin^2 C \tag{3.14}$$

$$\sigma_y^2 = \Delta t^2 \sigma_{\dot{y}}^2 .$$

In Equations (3.13), (3.14), σ_s is the speed change capability in the time interval Δt and σ_c is the course change capability in the time interval Δt. The variables σ_s and σ_c are obtained by

$$\sigma_s = \dot{S} \Delta t \tag{3.15}$$

$$\sigma_c = \dot{C} \Delta t , \tag{3.16}$$

where $\dot{S}$ in Equation (3.15) and $\dot{C}$ in Equation (3.16) are input quantities that describe the target's speed and course change capability.

4. ADAPTIVE DETERMINATION OF MEASUREMENT NOISE CHARACTERISTICS

Thus far in our treatment we have assumed that we possessed total statistical knowledge of the noise which is assumed to additively corrupt our range and bearing measurements. As can be seen from several of the foregoing equations, for example, Equation (3.6), the values of the bearing and range error measurement noise variances are necessary to calculate the gain matrix. A perfect knowledge of the defining parameters of the noise statistics would permit us to utilize these quantities in the most efficient manner in the design of the filter and to construct a theoretically optimal filter.

Unfortunately one rarely has complete knowledge of the statistical structure of the measurement noise. Moreover, in marine applications the actual noise is likely to possess a very complicated structure so that the assignment of fixed numerical values to such quantities as measurement variances is likely to be presumptuous at best. This practice of assuming values for measurement noise variances and building them into the filter could result in the construction of the best filter for the wrong problem.

Intuitively, what one would like is to design a filter which had a minimal dependance upon numerical values of parameters which were known with imperfect confidence. Some knowledge of measurement noise variances is necessary, however, in order to allow the filter to apply the appropriate weighting to the input data. This suggests the possibility of designing the filter so that it can estimate, or at least monitor, the value of the measurement noise variances to ensure that the proper weighting was being given to the data. If this could be done, it would release the filter designer from the necessity of supplying numerical values for measurement noise variances where the validity of these values was quite uncertain.

One manner in which this adaptive variance estimation might be performed is as follows. Let us denote

$$\bar{X}^*(k) = \phi(k)\hat{X}(k-1) = \begin{pmatrix} \dot{X}^* \\ X^*(k) \\ \dot{Y}^* \\ Y^*(k) \end{pmatrix}. \tag{4.1}$$

Then based upon our $(k-1)^{th}$ stage estimate, we may *predict* the value of the bearing to be observed at time t_k by writing

$$\beta_k^* = \tan^{-1}\frac{X^*(k)}{Y^*(k)}. \tag{4.2}$$

At stage n, the bearing error variance may be adaptively estimated from the equation

$$\hat{\sigma}^2_{\beta_n} = \frac{1}{n}\sum_{i=1}^{n}(\beta_i^* - \beta_i)^2; \tag{4.3}$$

where β_i is the actual measured value of the bearing at time t_i. For real time processing Equation (4.3) may be written in the following very simple recursive form

$$\hat{\sigma}^2_{\beta_n} = \left(1 - \frac{1}{n}\right)\hat{\sigma}^2_{\beta_{n-1}} + \frac{1}{n}(\beta_n^* - \beta_n)^2. \tag{4.4}$$

Similarly, if we define our predicted range R_i^* to be given by

$$R_i^* = \sqrt{(X^*(i)^2 + Y^*(i)^2)}, \tag{4.5}$$

then equations analogous to Equations (4.3) and (4.4) may be written to adaptively estimate the variance on the range measurement errors.

One of the noise sources might conceivably be non-stationary. In particular, in many situations it seems that the range measurement error variance is itself dependent upon range. In this event we could account for the non-stationarity in the noise variance by limiting the "memory" of our adaptive range variance estimator to the most recent m pieces of data. In this case the "smoothed" estimated range variance at time t_n would be given by

$$\hat{\sigma}^2_{R_{n/m}} = \frac{1}{m}\sum_{i=n-m+1}^{n}(R_i^* - \bar{R}_i)^2. \tag{4.6}$$

The trick in using Equation (4.6) is to choose m to be sufficiently large so that sampling fluctuations are smoothed out in the computation, but sufficiently small so that the actual variation of the true variance with the range is reflected. Equation (4.6) can be easily manipulated and written in the following recursive form for real time computation.

$$\hat{\sigma}^2_{R_{n/m}} = \hat{\sigma}^2_{R_{(n-1)/m}} + \frac{1}{m}\left[(R_n^* - \bar{R}_n)^2 - (R_{n-1-m}^* - \bar{R}_{n-1-m})^2\right]. \tag{4.7}$$

We have applied the adaptive techniques just described to the case of a linearly moving and a maneuvering target. The resulting root mean square speed errors, based upon 10 Monte Carlo runs are shown in Figure 3. This shows that the results of the adaptive variance estimation gave considerable improvement over the non-adaptive estimates, where, for the non-adaptive case, the input measurement standard deviations were in error by a factor of 3. Errors in knowledge of input measurement variances can often be this large so that we feel that this figure represents a realistic situation and that efforts to extend such adaptive techniques should be pursued.

The potentiality for an adaptive measurement noise variance estimation capability to be included within the filter is especially promising for tracking situations where the tracking extends over an appreciable period of time. For very short term tracking engagements the times involved are likely to be too short for the adaptive variance estimates to stabilize. Thus, for short tracking encounters, one will probably have to be content with utilization of whatever *a priori* knowledge he has of the measurement noise variances in the data weighting. However, for long time period tracking, this technique looks very promising.

5. EXAMPLES

We will consider four typical steering patterns for submarines: constant course and speed (Fig.3), fast 90° turn (Fig.4), slow 90° turn (Fig 5), and a sequence of maneuvers that will maintain an average speed in a preferred direction (Fig.6). The constant course and speed geometry will test how well the Kalman filter can track a cooperative submarine. For this case the Kalman filter will have a zero uncertainty gain function and this will demonstrate its heavy smoothing capability. The fast 90° turn maneuver will test two important factors. The first factor is the maneuver recognition time and the second is the tracker recovery time. The maneuver recognition time is the time from the beginning of a sharp manouver to the time the Kalman filter is definitely correcting to the new heading. The tracker recovery time is the time from the termination of the sharp maneuver to the time that the Kalman filter is estimating the new course and speed within some predetermined error bounds. The slow 90° turn manouver will test the Kalman filter's steady-state response to a constant rate course change The sequence of fast 90° maneuvers will test the Kalman filter's ability to track a rapidly maneuvering submarine. The Kalman filter will never be able to estimate, to a high degree of accuracy, the target course and speed, but it will be able to given an indication of the agility of the tracker for different uncertainty gain functions. In this case the position errors and velocity errors become the important evaluation factors. For the experiments we will consider a North-East coordinate system, where North is the Y-axis and East is the X-axis. We will also have the tracking platform (our ship) standing still at $X = 0$ and $Y = 0$.

Geometry I. Constant Course and Speed

The geometry is illustrated in Figure 3. Initially the target starts at X = 2000 yards and Y = 10,000 yards. The target will maintain a constant speed of 25 knots on a course of 135 degrees.

Geometry II. Fast 90-Degree Turn

The geometry is illustrated in Figure 4. Initially the target starts at X = 2000 yards and Y = 10,000 yards on a course of 180 degrees maintaining a speed of 25 knots. After 204 seconds the submarine performs a 45-second maneuver to a new course of 90 degrees and a speed of 15 knots. At the time of the maneuver the target was approximately 7000 yards from our own ship.

Geometry III. Slow 90-Degree Turn

The geometry is illustrated in Figure 5. Initially the target starts at X = 2000 yards and Y = 10,000 yards. on a course of 180 degrees maintaining a speed of 25 knots. After 10 seconds the submarine performs a 450-second maneuver to a new course of 90 degrees at the same speed of 25 knots.

Geometry IV. Sequence of Fast 90-Degree Turns

The geometry is illustrated in Figure 6. Initially the target starts at X = 3000 yards and Y = 10,000 yards on a course of 180 degrees maintaining a speed of 25 knots. The submarine will perform a sequence of 90-degree turns so that it will still have a preferred direction of 180 degrees. The speed in the preferred direction was to be 90% of the actual speed of the submarine. The time duration of each maneuver was 80 seconds, so that the final range will be approximately 4000 yards.

6. ANALYSIS OF RESULTS

The purpose of this section is to present a description of the graphical results. All the computer runs were made with 10 different Monte Carlo runs to obtain the root mean square (r.m.s.) error statistics associated with each geometry. The results presented are for r.m.s. course, speed, velocity and position errors. The error in course is found by calculating the r.m.s. difference between the estimated and true course. The error in speed is found by taking the r.m.s. difference between the true and estimated speed. The velocity error is the r.m.s. magnitude of the velocity error vector. The position error is the r.m.s. magnitude of the position error vectors.

The result of the adaptive estimation process is presented by Figure 4. In the weighting of the observational data it was assumed (guessed) that the noise sigma on the range information was one percent of range and one degree on the bearing observations. Actually the noise sigma was three percent of range and three degrees on the bearing observation. Two effects are immediately apparent. First the adaptive noise estimation process is worthwhile in that it has smaller associated error statistics. Second it gives a significantly more stable estimate of the target's state vector. Both techniques used the same bad guess sigmas values for the first nine data reports. From the tenth report on, the adaptive Kalman filter used its own estimate of the noise sigmas, while the other continued on with the bad guess.

The results of the constant course and speed target are presented in Figure 3.

This geometry was run with the Kalman filter processing range and bearing information (KF) and with the Kalman filter with Doppler processing range, bearing and range-rate information (KFWD). The only significant difference is the smaller initial errors of the KFWD tracking system.

The results of the fast 90-degree turn target are presented in Figure 4. As in the previous case, both tracking systems (KF, KFWD) were employed. Due to the rapid closing rate of the target, the initial convergence of the KFWD tracker is very significant. Also the use of Doppler information greatly aided the response of the Kalman filter during the maneuver that occurred at 204 seconds into the run. The time difference between the two trackers to return to the new track is significant.

The results of the slow 90-degrees turn target are presented in Figure 5. The geometry indicates the difference between the steady-state response of the KF and KFWD trackers. The steady-state response to the constant turn rate (0.2 degrees/second) of the KF tracker was 2.5 knots speed error, 9.0 degrees course error and 80 yards position error. The steady-state response of the KFWD tracker was 1.7 knots speed error, 4.5 degrees course error and 70 yards position error.

The results of the sequence of the fast 90-degree maneuvering target are presented in Figure 6. This geometry indicates the Kalman filter's response to a sequence of consecutive 90-degree turns of the target. It is apparent that the KFWD tracker benefits from the Doppler information. The speed error of the KFWD tracker, throughout the entire sequence of maneuver, is one-half the KF tracker's speed errors.

ACKNOWLEDGEMENT

The authors are indebted to many who have facilitated the preparation of this chapter. In particular, the authors acknowledge the support given them by the Advanced Marine Technology Division of Litton Industries, and especially Mrs.Estelle Hasson, who typed and otherwise prepared the manuscript.

REFERENCES

1. Tucker, D.G.
Galey, B.K. — *Applied Underwater Acoustics*. Pergamon Press, Oxford, 1966.

2. Sorenson, H.W. — *Theoretical and Computational Considerations of Linear Estimation Theory in Navigational Guidance*. AC Electronics Division, General Motors, Memorandum-LA-3005, November 1965.

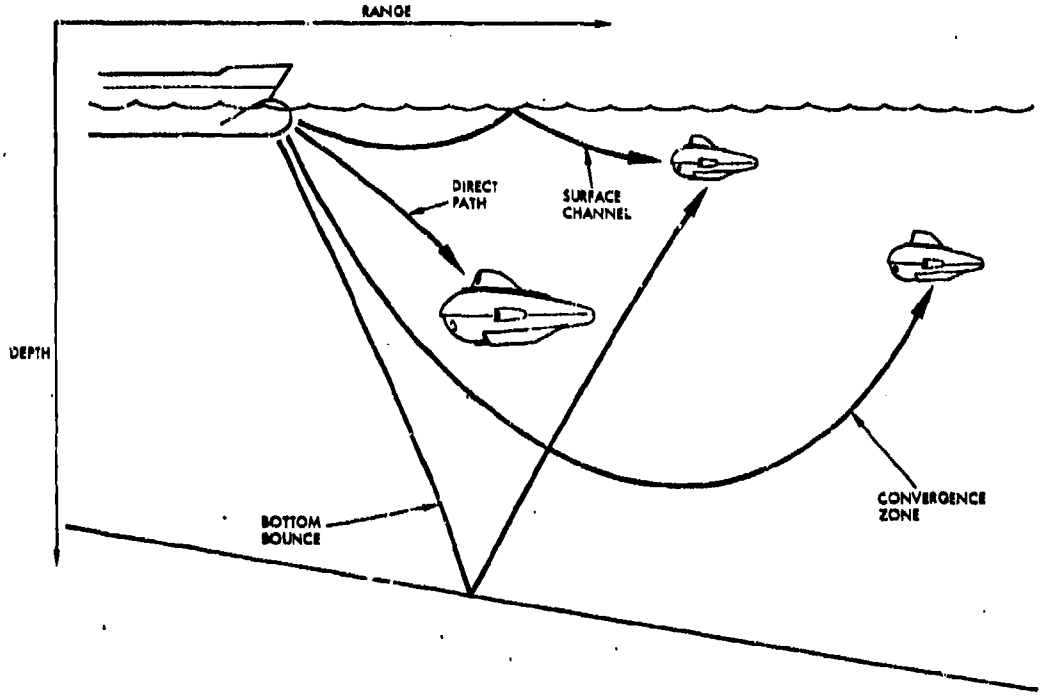


Fig. 1 Sound propagation modes

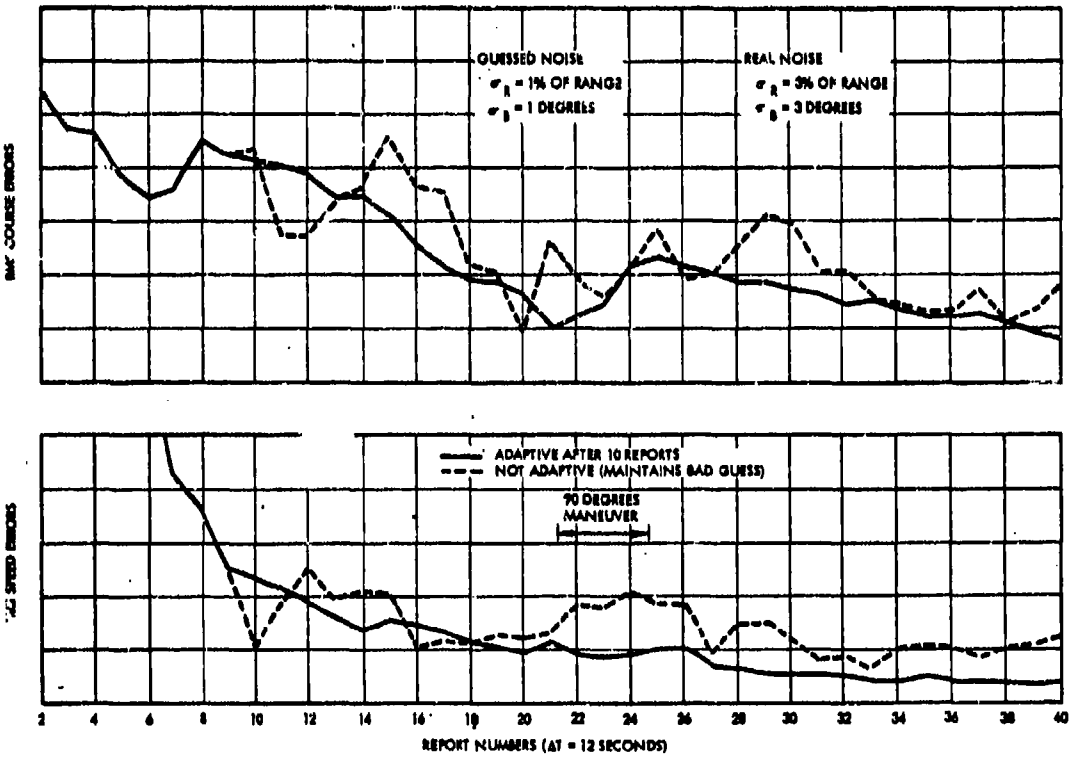


Fig. 2 Adaptive estimation process

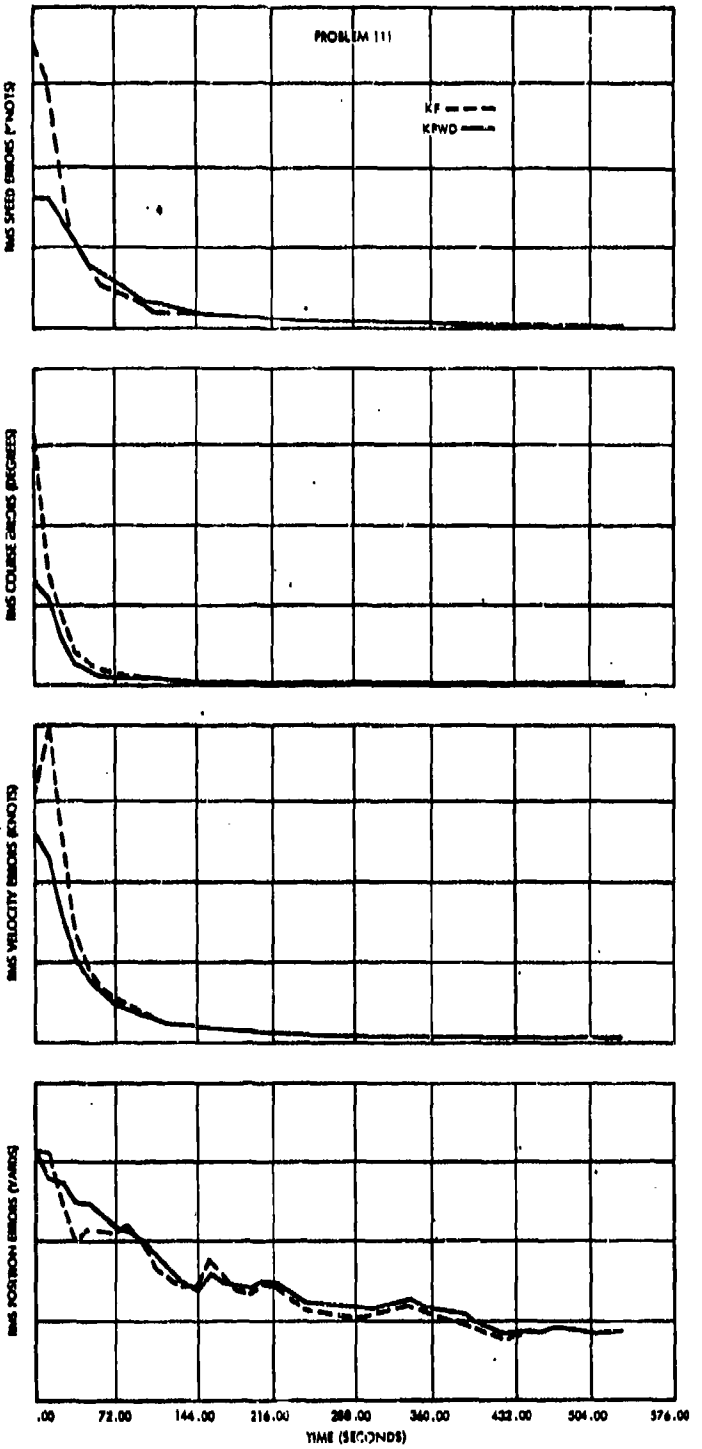


Fig.3 Constant course and speed geometry

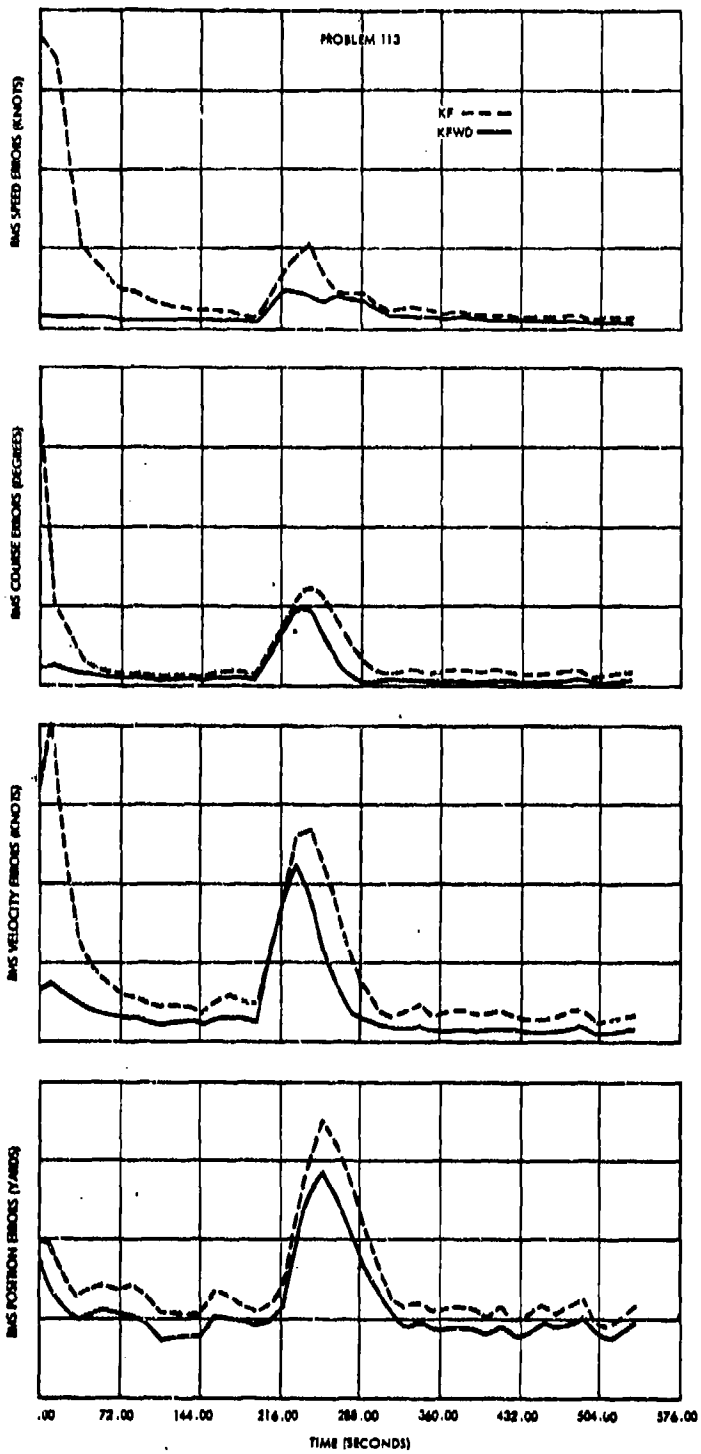


Fig.4 Fast 90-degree turn geometry

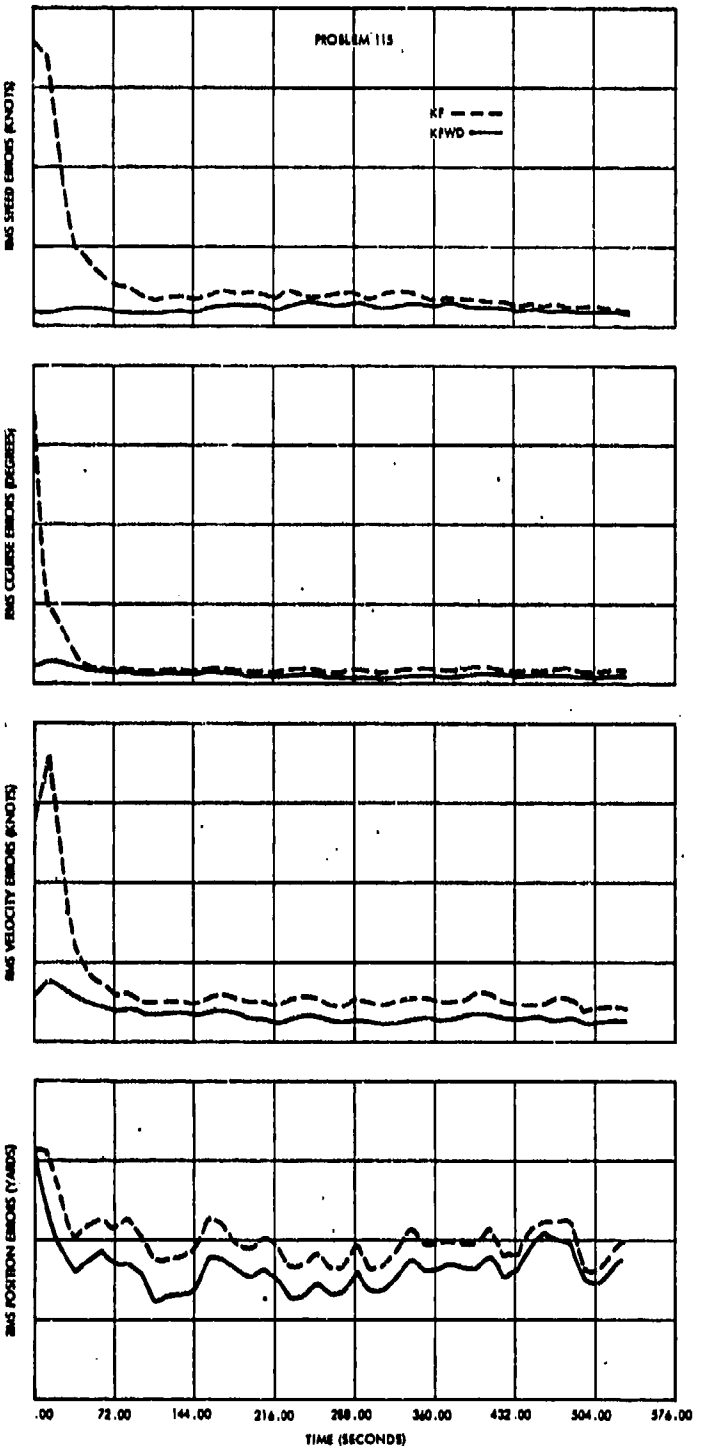


Fig. 8 Slow 90-degree turn geometry

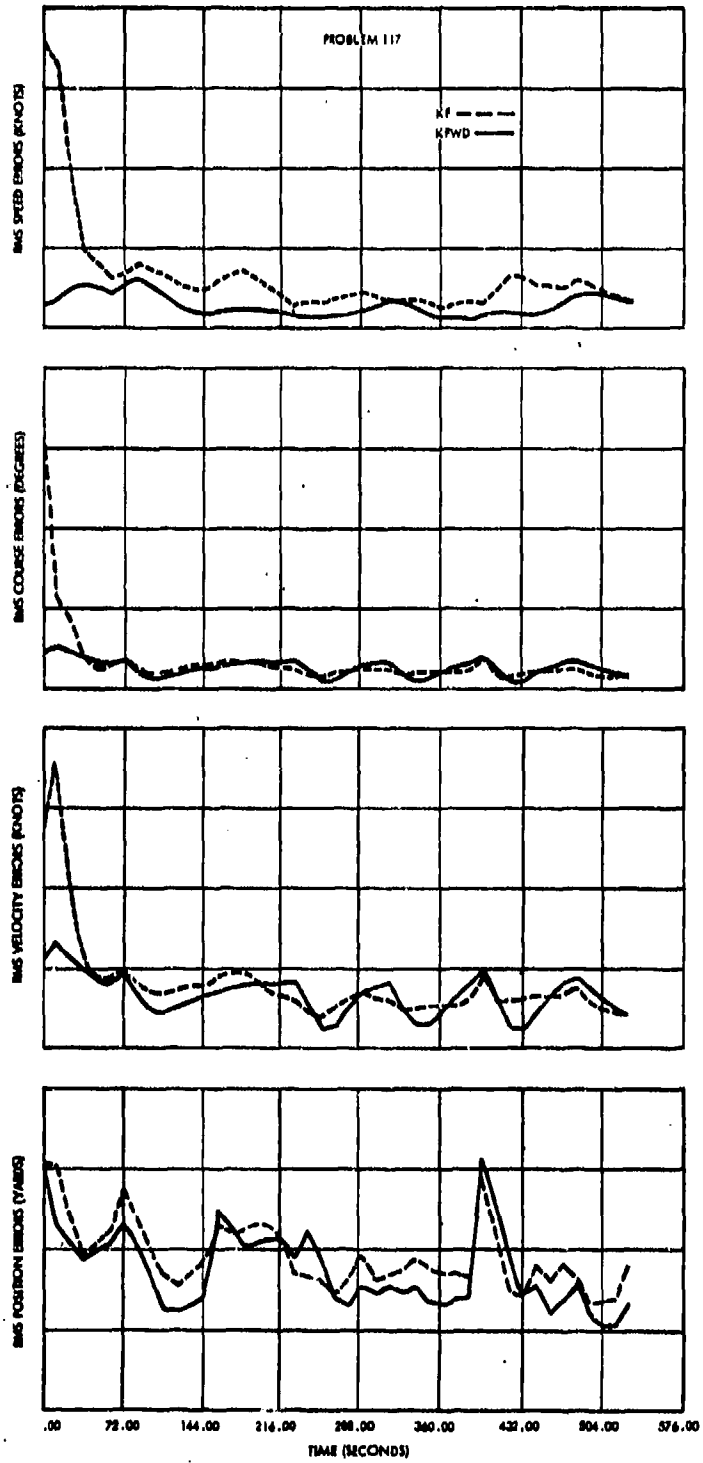


Fig. 8 Sequence of fast 90-degree turns

CHAPTER 18 - OPTIMAL USE OF REDUNDANT INFORMATION IN AN INERTIAL NAVIGATION SYSTEM

by

F.G.Unger and N.Ott

TELDIX, Heidelberg, Germany

NOTATION

Vectors: Lower case letters (Latin or Greek, with arrows)

Matrices: Capital letters (Latin or Greek, underlined)

Functions and Constants: Capital or lower case letters (Latin or Greek)

$\vec{a}$		thrust acceleration vector
$\vec{b}$		total acceleration vector
$\vec{r}$		position vector
$\vec{v}$		velocity vector
$\vec{\omega}$		rate vector of navigational coordinate system with respect to inertial space
$\vec{\omega}_E$		earth rate vector with respect to inertial space
$\vec{\rho}$	=	$\vec{\omega} - \vec{\omega}_E$, rate vector of the navigational coordinate frame with respect to an earth-fixed frame
$\vec{g}_a$		vector of mass attraction
$\vec{g}$	=	$\vec{g}_a - \vec{\omega}_E \times (\vec{\omega}_E \times \vec{r})$, gravity vector
g_0		value of the gravity vector at the surface of the earth
r_0		length of the radius of the spherical earth
$\vec{\nabla}$		accelerometer bias vector
$\vec{\varepsilon}$		gyro drift vector
$\vec{\phi}$		error angle vector between navigational and platform coordinate frame
$\vec{\psi}$		error angle vector between computer and platform coordinate frame
$\vec{\Delta r}$		position error vector of the inertial system
$\vec{\Delta v}$	=	$\dot{\vec{\Delta r}}$, velocity error vector of the inertial system
$\vec{\delta v}$		reference velocity error vector
$\vec{v}_n$		velocity vector computed in the inertial system
$\vec{v}_r$		reference velocity vector
$C_\nu (\nu = 1,2,3)$		control signals in updating mechanisations
$K_\nu (\nu = 1,2,3)$		gain constants
ω_s		Schuler frequency
T_s		Schuler period
t, τ		time variables
s		Laplace operator
$x(t)$		input signal of a linear system
$y(t)$		output signal of a linear system
σ^2		variance
$g(\tau)$		impulse response of a linear system
$G(s)$		transfer function of a linear system
τ_x		correlation time of the input signal

$\delta(\tau)$		Dirac δ-distribution
$\rho(\tau)$		autocorrelation function
$\phi(\omega)$		power spectral density
J	=	$\sqrt{(-1)}$, imaginary unit vector
$(\underline{M})$		direction cosine matrix
$\overrightarrow{x(t)}$, $\vec{x}_k$		state vector
$\vec{x}_k^1$		controllable part of the state vector
$\vec{x}_k^2$		uncontrollable part of the state vector
$\overrightarrow{y(t)}$, $\vec{y}_k$		measurement vector
$\underline{A(t)}$		state coefficient matrix
$\underline{M(t)}$, $\underline{M}_k$		measurement matrix
$\overrightarrow{n(t)}$, $\vec{h}_k$		system white noise vectors
$\vec{h}_k^1$, $\vec{h}_k^2$		white noise driving functions of the controllable and uncontrollable states
$\overrightarrow{q(t)}$, $\vec{q}_k$		measurement white noise vector
$\overrightarrow{u(t)}$, $\vec{u}_k$		control vector
$\hat{\vec{x}}_k^p$		*a posteriori* estimate of the state vector
$\hat{x}_k^{1p}$, $\hat{x}_k^{2p}$		the controllable and uncontrollable states of $\hat{\vec{x}}_k^p$
$\vec{x}_k^*$		*a priori* estimate of the state vector
$\vec{x}_k^{1*}$, $\vec{x}_k^{2*}$		controllable and uncontrollable states of $\vec{x}_k^*$
$\vec{y}_k^*$		*a priori* estimate of the measurement
$\underline{B}_k$		weighting matrix (gain matrix)
$\underline{P}_k^*$		covariance matrix of the *a priori* estimation error
$\underline{P}_k$		covariance matrix of the *a posteriori* estimation error
$\underline{\Phi_{t/t_n}}$, $\underline{\Phi_{t_n/t_{n-1}}}$, $\underline{\Phi_{k+1/k}}$		state transition matrix
$\underline{\Phi^{11}_{k+1/k}}$, $\underline{\Phi^{12}_{k+1/k}}$, $\underline{\Phi^{22}_{k+1/k}}$		appropriate parts of $\underline{\Phi_{k+1/k}}$
$\underline{I}$		identity matrix
$\underline{H}_k$		covariance matrix of the system white noise vector $\vec{h}_k$
$\underline{Q}_k$		covariance matrix of the measurement white noise vector $\vec{q}_k$
$\underline{\Gamma_{k+1/k}}$		distribution matrix of the control vector $\vec{u}_k$
$\underline{\Gamma^1_{k+1/k}}$		appropriate part of $\underline{\Gamma_{k+1/k}}$ for the controllable states
$w(t)$		a unity variance white noise source
$1/\beta$		correlation time
$f(t)$		output function of a shaping filter
r_{ww}		random walk parameter
ω		frequency in rad/sec
$E[\ldots]$		expectation operator
$\underline{N(t)}$, $\underline{N}_k$		covariance matrix of the white noise vector $\overrightarrow{n(t)}$
δ_{kL}		Kronecker symbol

CHAPTER 18 - OPTIMAL USE OF REDUNDANT INFORMATION IN AN INERTIAL NAVIGATION SYSTEM

F.O.Unger and N.Ott

1. INTRODUCTION

Accurate and reliable navigation is a very stringent requirement, especially for military missions. Desired outputs from a navigation system are position, velocity and orientation of the sensors and the vehicle with respect to a navigational coordinate system. A pure inertial navigation system can provide all this information, but the errors of an inertial system are not bounded, i.e. they increase with time. Therefore, with large mission times, the errors can become intolerably large. Other methods of navigation, like Doppler dead reckoning, radio, stellar etc., can provide only part of the information required, but with bounded errors in velocity or position. A navigation system that fulfils a wide variety of requirements for relatively long mission times will therefore be a combination of an inertial system with other navigational aids. Such a system can provide all the information required and, at the same time, control the system errors. But in order to use all the available information in the best possible way the statistics of the error sources must be known. In addition a heavy load is placed on the digital computer.

This chapter presents the different methods (i.e. conventional versus modern, or deterministic versus statistical) of aiding a pure inertial system with additional or redundant navigation information. But, before the basic problem of updating is treated, the available sensor outputs are categorised and a general block diagram for aided inertial navigation systems is derived. The pure inertial system in a Schuler-tuned mode is included in this block diagram as a special case.

2. BLOCK DIAGRAM OF AIDED INERTIAL NAVIGATION SYSTEMS

The central unit of a navigation system is the navigation computer. Its inputs are available data from sensors, such as inertial acceleration, ground speed, air speed, heading, position information, stellar angles, etc. These data may be continuous in analogue or digital form, or intermittent like ground fix points. The computer has to calculate the desired navigation outputs: position, velocity and orientation of the sensors and the vehicle with respect to a navigational coordinate system. In addition the computer must also compensate for sensor biases and calculate torquing commands for platform gyros or positional servos for stellar or radar equipment.

Figure 1 shows the possible inputs and required outputs of a navigation system.

The available sensor information, as well as the computed output data, can be classified into three levels:

Level 1: Acceleration
Level 2: Velocity
Level 3: Position and orientation.

Orientation belongs to the same level as position, because angular measurements lead to position information and both data complement each other. Furthermore, in a rotating coordinate frame orientation angles can be derived by integration of the angular rate vector $\vec{\omega}$, just as position is derived by integration of the velocity vector $\vec{v}$.

Figure 2 shows the three levels and the equivalent inputs.

Information for a higher level can be derived from a lower level by integration; but it is generally not possible to derive lower level information from higher levels by differentiation.

The information in each level is a vector in the navigation coordinate frame, except for orientation, where the information is in the form of three angles or the direction cosine matrix ($\underline{M}$) for possible coordinate transformations.

If the navigation frame is rotating with respect to inertial space, corrections for acceleration must be included. This can be recognized by writing the basic equations

$$\left(\frac{d^2\vec{r}}{dt^2}\right)_I = \vec{a} + \vec{g}_r \qquad (2.1)$$

$$\left(\frac{d\vec{r}}{dt}\right)_I = \vec{v} + \vec{\omega} \times \vec{r} \tag{2.2}$$

$$\left(\frac{d^2\vec{r}}{dt^2}\right)_I = \left(\frac{d\vec{v}}{dt}\right)_N + \dot{\vec{\omega}} \times \vec{r} + 2\vec{\omega} \times \vec{v} + \vec{\omega} \times (\vec{\omega} \times \vec{r}) , \tag{2.3}$$

where

$\vec{r}$ = position vector

$(d\vec{r}/dt)_I$ = time derivative of $\vec{r}$ with respect to inertial space

$(d\vec{r}/dt)_N = \vec{v}$ = derivative of $\vec{r}$ with respect to the navigation coordinate frame

$\vec{\omega}$ = rate of change of navigational coordinate system with respect to inertial space

$\vec{a}$ = thrust acceleration vector

$\vec{g}_m$ = vector of mass attraction

$\vec{b} = (d\vec{v}/dt)_N$ = total acceleration in navigation coordinates.

Substitution of Equation (2.3) into Equation (2.1) yields

$$\vec{b} = \vec{a} + \vec{g}_m - \dot{\vec{\omega}} \times \vec{r} - 2\vec{\omega} \times \vec{v} - \vec{\omega} \times (\vec{\omega} \times \vec{r}) , \tag{2.4}$$

Equation (2.4) shows that the following corrections are required to the acceleration level, in addition to measurable thrust acceleration:

(i) Acceleration due to mass attraction: $\vec{g}_m$

(ii) Coriolis acceleration: $2\vec{\omega} \times \vec{v}$

(iii) Tangential acceleration: $\dot{\vec{\omega}} \times \vec{r}$

(iv) Centripetal acceleration: $\vec{\omega} \times (\vec{\omega} \times \vec{r})$.

If the sensitive axes of the sensors do not coincide with the axes of the navigational coordinate system, the sensor outputs must be transformed with the orientation matrix ($\underline{M}$).

Figure 3 shows a general block diagram of navigation systems, including the corrections just mentioned for a rotating frame. Redundant data may enter the system at levels 2, 3a and 3b. These input data consist of the correct values and errors. The errors can be bias values or stochastic processes or a combination of both. Consequently the navigation information is erroneous too.

The following two problems must be solved in order to use the available redundant information in the best possible way:

(a) Determine the desired output information with minimum errors (minimum variance). This is a filtering or estimation problem.

(b) Control the system error states. This task is a system control problem.

Both tasks are accomplished with a minimum variance or Kalman filter. Task (a) requires valid mathematical models of the different error sources. Task (b) is accomplished by feedback loops (inside the computer) as indicated in Figure 2 by dotted lines.

Before these two tasks are discussed in more detail in Section 6, the error behaviour of pure inertial systems and the conventional methods of updating are considered.

3. DETERMINISTIC ERROR ANALYSIS AND ERROR BEHAVIOUR OF PURE INERTIAL SYSTEMS

In a pure inertial system thrust acceleration represents the only input formation measured at the first level. In most cases the input axes of the sensors are parallel to the axes of the navigational coordinate frame. If this is not the case, a coordinate transformation is necessary.

Figure 4 shows the block diagram of a pure inertial system used for navigation over the surface of the earth. In this case the velocity $\vec{v}$ of a vehicle is defined as the rate of change of the vector $\vec{r}$ with respect to the earth ($v = (d\vec{r}/dt)_E$). Therefore, instead of Equation (2.4), we now have the following relation for acceleration corrections:

$$\vec{b} = \left(\frac{d\vec{v}}{dt}\right)_N = \vec{a} + \vec{g}_m - \vec{\omega}_E \times (\vec{\omega}_E \times \vec{r}) - (2\vec{\omega} \times \vec{\rho}) \times \vec{v} ; \tag{3.1}$$

where

$\vec{\omega}_E$ = earth rate vector

$\vec{\rho} = \vec{\omega} - \vec{\omega}_E$ = rate vector of the navigational coordinate frame with respect to an earth-fixed frame.

As shown in Figure 4, a feedback loop exists owing to the gravity vector $\vec{g}$, i.e. the system is capable of oscillating. The oscillation occurs with the Schuler period

$$T_s = 2\pi\sqrt{\left(\frac{r_0}{g_0}\right)}.$$

The error behaviour of a pure inertial system in a Schuler-tuned mode will now be discussed[11].

Figure 5 shows the error block diagram for one horizontal channel of such a system. The errors considered are accelerometer bias, gyro drift, and azimuth misalignment.

The following notation is used in Figure 5:

$\Delta\dot{r}_x$ = computed velocity error in x-direction

Δr_x = computed position error in x-direction

∇_x = x-accelerometer bias

ϵ_y = fixed drift of the y-gyro

φ_y = y-component of the error angle vector between the navigational and platform coordinate frame.

ψ_z = z-component of the error angle vector between the platform and computer coordinate frame (azimuth misalignment).

The velocity and position errors, as derived from Figure 5, are

$$\Delta\dot{r}_x = \frac{s\nabla_x - g_0(\epsilon_y + \omega_x\psi_z)}{s^2 + \omega_s^2} \tag{3.2}$$

$$\Delta r_x = \frac{s\nabla_x - g_0(\epsilon_y + \omega_x\psi_z)}{s(s^2 + \omega_s^2)}. \tag{3.3}$$

The characteristic equation (denominator of Equation (3.2)) is that of an undamped oscillation with Schuler frequency. For constant sensor errors the steady-state value of the velocity error is

$$\Delta\dot{r}_x|_{s=0} = -\frac{g_0}{\omega_s^2}(\epsilon_y + \omega_x\psi_z). \tag{3.4}$$

The position error is constantly increasing with time,

$$\Delta r_x|_{s=0} \to \infty, \tag{3.5}$$

and it is unbounded owing to gyro drift and azimuth misalignment; accelerometer bias does not cause an unbounded position error.

4. DETERMINISTIC UPDATING METHODS

Although Schuler tuning is of great importance in navigation over the surface of the earth, it has been shown that for longer mission times the position errors can become intolerably large, especially because of gyro drift and azimuth error. To avoid this, the pure inertial system can be aided with additional or redundant navigation information from other sources. Available continuous or discrete data for the navigation computer may be of levels 2, 3a and 3b. No additional information on level 1 is available, because we do not know any other method of measuring acceleration except by inertial means.

Accordingly we distinguish different kinds of aided systems:

(i) Velocity aided systems (redundant data at level 2),

(ii) Position aided systems (redundant data at level 3a).

(iii) Stellar monitored systems (redundant data at level 3b).

(iv) Stellar monitored and velocity damped systems (redundant data at levels 2 and 3b).

The additional or redundant data can be intermittent and used to replace the computed data at certain time periods. If it is always available, it can be continuously compared with the computed information; the difference signal, multiplied with various gains K_ν, can be added to the accelerometer output signals or the gyro torquing commands either directly or over additional integrators. If the difference or error signal is introduced at the next lower level it is used to damp the oscillations of the system. If it is introduced at the same level inside the Schuler loop, then the frequency and amplitude can be changed. If additional integrators are used, sensor biases can be compensated.

Figure 6 shows the x-channel of a Schuler-tuned system and the mechanisation of the different updating methods.

These methods are now discussed in more detail, where Doppler velocity is available as redundant information. The Doppler velocity signal cannot very well be used to reset the inertial velocity signal directly because of its high noise level. The direct resetting method is not very useful and is only applicable if angular measurements or position fixes are available.

As Figure 6 indicates, the difference signal can be used to provide three different correction signals, C_1, C_2 and C_3, into the Schuler loop.

The first signal C_1, the difference signal multiplied with the gain factor K_1, is added to the accelerometer output. This provides damping of the Schuler oscillation. The second signal C_2, the difference signal multiplied with the gain factor K_2/r_0, affects the oscillation frequency by adding to it the gyro input. The third input C_3 provides compensation for some steady-state errors of velocity by introducing an additional integrator with gain K_3/r_0.

Assuming that both the computed inertial and the measured reference velocity are subjected to errors, the three corrections corresponding to the x-channel error block diagram can be written as follows:

$$C_{1x} = -K_1(v_{ox} - v_{rx}) = -K_1(\Delta\dot{r}_x - \delta v_x) \tag{4.1}$$

$$C_{2x} = \frac{K_2}{r_0}(v_{ox} - v_{rx}) = \frac{K_2}{r_0}(\Delta\dot{r}_x - \delta v_x) \tag{4.2}$$

$$C_{3x} = \frac{K_3}{r_0}\int(v_{ox} - v_{rx})\,dt = \frac{K_3}{r_0}\int(\Delta\dot{r}_x - \delta v_x)\,dt\,, \tag{4.3}$$

where $\Delta\dot{r}_x$ and δv_x represent the errors of the computed inertial and the measured Doppler reference velocity, respectively.

Figure 7 shows the x-channel error block diagram of the Doppler inertial system described.

The transfer functions of the velocity and position errors can be derived from Figure 7 as follows:

$$\Delta\dot{r}_x = \frac{s^2\nabla_x - sg_0(\epsilon_y + \omega_z\psi_z) + [\omega_s^2(K_2 s + K_3) + s^3K_1]\delta v_x}{s^3 + K_1 s^2 + (1+K_2)\omega_s^2 s + K_3\omega_s^2} \tag{4.4}$$

$$\Delta r_x = \frac{s^2\nabla_x - sg_0(\epsilon_y + \omega_z\psi_z) + [\omega_s^2(K_2 s + K_3) + s^3K_1]\delta v_x}{s[s^3 + K_1 s^2 + (1+K_2)\omega_s^2 s + K_3\omega_s^2]} \tag{4.5}$$

$$\Delta\dot{r}_x\big|_{s=0} = \delta v_x\,. \tag{4.6}$$

Equations (4.4) and (4.5) can be used to evaluate the velocity and position errors for different values of K_1, K_2 and K_3, or to obtain optimal values of these three gain factors. The steady-state velocity error now contains the reference velocity error δv_x only (Equation (4.6)). The fixed gyro drift and azimuth error are compensated by the additional integrator.

Figure 8 shows, as an example, the time functions of the velocity error due to gyro drift for the pure inertial system ($K_1 = K_2 = K_3 = 0$) and for the different methods of updating ($K_1 \neq 0$; K_1 and $K_2 \neq 0$ and $K_1, K_2, K_3 \neq 0$). In this example only the influence of the various gain factors K_ν on the error behaviour $\Delta\dot{r}_x$ are of interest, and not the values of these factors themselves. This means that the chosen values of the gain factors cannot be considered as optimal values. Furthermore, an exact definition of "optimal" has not been given so far, because only one error source is taken into account. The problem of optimal gains is discussed in Section 6.

5. STATISTICAL ERROR ANALYSIS AND ERROR BEHAVIOUR

So far constant errors (biases) have been discussed, but in many cases the errors are not constant, but rather stochastic processes characterised by their autocorrelation functions $\rho_\nu(\tau)$ or their spectral density functions $\phi_\nu(\omega)$.

For linear, non-time-varying systems the statistics of the output signal $y(t)$ follows directly from the time-independent statistics of the input signal $x(t)$. The mean value of the output signal is zero if the mean value of the input signal is zero. The average value of the square of the output signal, the time-dependent variance $\sigma^2_{y(t)}$, can be computed by the expression[3]

$$\sigma^2_{y(t)} = \int_0^t\int_0^t g(\tau_1)g(\tau_2)\rho_x(|\tau_1 - \tau_2|)\,d\tau_1 d\tau_2\,, \tag{5.1}$$

where

$$g(\tau) = \frac{1}{2\pi j}\int_{c-j\infty}^{c+j\infty} G(s)e^{s\tau}\,ds \ , \text{ system impulse response}$$

$G(s)$ = system transfer function

$\rho_x(|\tau|)$ = input signal autocorrelation function.

If the input signal is white noise, the autocorrelation function is given by

$$\rho_x(|\tau_1-\tau_2|) = A\delta(|\tau_1-\tau_2|) \ ; \quad A = \text{constant} \ . \tag{5.2}$$

Equation (5.2) substituted into Equation (5.1) yields

$$\sigma^2_{y(t)} = \int_0^t\int_0^t g(\tau_1)g(\tau_2)A\delta(|\tau_1-\tau_2|)\,d\tau_1 d\tau_2$$

$$= A\int_0^t g(\tau_1)\,d\tau_1 \int_0^t g(\tau_2)\,\delta(|\tau_1-\tau_2|)\,d\tau_2 \ .$$

The inner integral can be evaluated by using the distribution property of the δ-function

$$\int_0^t g(\tau_2)\delta(|\tau_1-\tau_2|)\,d\tau_2 = \left\{\begin{matrix} g(\tau_1) & \text{for } \tau_1 \text{ in } [0,t] \\ 0 & \text{for } \tau_1 \text{ not in } [0,t]\end{matrix}\right\}. \tag{5.3}$$

With Equation (5.3) the expression for $\sigma^2_{y(t)}$ becomes

$$\sigma^2_{y(t)} = A\int_0^t g^2(\tau_1)\,d\tau_1 \ . \tag{5.4}$$

If the input signal is bandwidth limited white noise with autocorrelation function

$$\rho_x(|\tau|) = \sigma_x^2 e^{-\frac{|\tau|}{\tau_x}} \ , \tag{5.5}$$

Equation (5.4) can also be used if τ_x is small compared with the system time constants. Equation (5.5) can then be approximated by

$$\rho_x(|\tau|) \cong 2\tau_x\sigma_x^2\delta(\tau) \tag{5.6}$$

and the output variance yields, in this case,

$$\sigma^2_{y(t)} = 2\tau_x\sigma_x^2\int_0^t g^2(\tau)\,d\tau \ . \tag{5.7}$$

Another method of determining the steady-state value of the variance $\sigma^2_{y(t)}$, if it is convergent, is given by[3]

$$\lim_{t\to+\infty}\sigma^2_{y(t)} = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\phi_y(\omega)\,d\omega \ . \tag{5.8}$$

The output spectral density function is given by

$$\phi_y(\omega) = |G(j\omega)|^2\phi_x(\omega) \ . \tag{5.9}$$

For bandwidth limited white noise, ϕ_x can be evaluated from the approximated autocorrelation function $\rho_x(|\tau|)$ of Equation (5.6),

$$\phi_x(\omega) = \int_{-\infty}^{+\infty}\rho_x(|\tau|)e^{-j\omega\tau}\,d\tau = 2\tau_x\sigma_x^2 \ . \tag{5.10}$$

Equations (5.9) and (5.10) substituted into Equation (5.8) yield

$$\lim_{t\to+\infty}\sigma^2_{y(t)} = (2\tau_x\sigma_x^2)\,\frac{1}{2\pi}\int_{-\infty}^{+\infty}|G(j\omega)|^2\,d\omega \ . \tag{5.11}$$

It should be noted that Equations (5.8) and (5.11) can only be used if $G(s)$ has no poles on the imaginary axis or in the right half-plane. A pure inertial system in a Schuler-tuned mode is undamped, i.e. it has poles on the imaginary axis. This means that the effect of noise inputs must be calculated according to Equations (5.1) or (5.7).

As an example the variance of the velocity and position errors due to a white noise Doppler input error δv_x for a damped system are now calculated.

Using Equations (4.4) and (4.5) the transfer functions for $K_2 = K_3 = 0$ are

$$G_{\Delta\dot{r}_x}(s) = \frac{K_1 s}{s^2 + K_1 s + \omega_s^2} \tag{5.12}$$

$$G_{\Delta r_x}(s) = \frac{K_1}{s^2 + K_1 s + \omega_s^2} . \tag{5.13}$$

At first Equation (5.11) is used:

$$\begin{aligned}
\sigma^2_{\Delta\dot{r}_x} &= 2\tau_{\delta v_x}\sigma^2_{\delta v_x}\frac{1}{2\pi}\int_{-\infty}^{+\infty}|G_{\Delta\dot{r}_x}(j\omega)|^2\,d\omega \\
&= 2\tau_{\delta v_x}\sigma^2_{\delta v_x}\frac{K_1^2}{\pi}\cdot\int_0^{\infty}\frac{\omega^2}{(\omega_s^2-\omega^2)^2 + K_1^2\omega^2}\,d\omega \\
&= 2\tau_{\delta v_x}\sigma^2_{\delta v_x}\frac{K_1\omega_s}{\sqrt{(4\omega_s^2 - K_1^2)}}\cos\tfrac{1}{2}\arctan\frac{K_1\sqrt{(4\omega_s^2-K_1^2)}}{2\omega_s^2 - K_1^2} \\
&= \tau_{\delta v_x}\sigma^2_{\delta v_x}K_1 .
\end{aligned} \tag{5.14}$$

$$\begin{aligned}
\sigma^2_{\Delta r_x} &= 2\tau_{\delta v_x}\sigma^2_{\delta v_x}\frac{K_1^2}{\pi}\int_0^{\infty}\frac{d\omega}{(\omega_s^2-\omega^2)^2 + K_1^2\omega^2} \\
&= \frac{\tau_{\delta v_x}\sigma^2_{\delta v_x}K_1}{\omega_s^2} .
\end{aligned} \tag{5.15}$$

Now Equation (5.7) is used to determine the time-dependent variances.

The system impulse responses become

$$g_{\Delta\dot{r}_x}(t) = K_1 e^{-\frac{1}{2}K_1 t}\left(\cos\omega_0 t - \frac{K_1}{2\omega_0}\sin\omega_0 t\right) = \dot{g}_{\Delta r_x}(t) \tag{5.16}$$

$$g_{\Delta r_x}(t) = \frac{K_1}{\omega_0}e^{-\frac{1}{2}K_1 t}\sin\omega_0 t , \tag{5.17}$$

where

$$\omega_0 = \tfrac{1}{2}\sqrt{(4\omega_s^2 - K_1^2)} . \tag{5.18}$$

The desired time-dependent variances are obtained as follows:

$$\begin{aligned}
\sigma^2_{\Delta\dot{r}_x}(t) &= 2\tau_{\delta v_x}\sigma^2_{\delta v_x}\int_0^t g^2_{\Delta\dot{r}_x}(\tau)\,d\tau \\
&= \frac{2\tau_{\delta v_x}\sigma^2_{\delta v_x}K_1^2}{\omega_0(4\omega_0^2+K_1^2)}\left[\frac{\omega_0(4\omega_0^2+K_1^2)}{2K_1} - e^{-K_1 t}\left(\frac{8\omega_0^2K_1^2 + 16\omega_0^4 + K_1^4}{8\omega_0 K_1} - \right.\right. \\
&\qquad \left.\left. - \frac{4\omega_0^2 + K_1^2}{4}\sin 2\omega_0 t - \frac{K_1^4 + 4\omega_0^2K_1^2}{8\omega_0 K_1}\cos 2\omega_0 t\right)\right]
\end{aligned} \tag{5.19}$$

$$\begin{aligned}
\sigma^2_{\Delta r_x}(t) &= 2\tau_{\delta v_x}\sigma^2_{\delta v_x}\int_0^t g^2_{\Delta r_x}(\tau)\,d\tau \\
&= \frac{2\tau_{\delta v_x}\sigma^2_{\delta v_x}K_1^2}{\omega_0(4\omega_0^2+K_1^2)}\left[\frac{2\omega_0}{K_1} - e^{-K_1 t}\left(\frac{4\omega_0^2 + K_1^2}{2\omega_0 K_1} + \sin 2\omega_0 t - \frac{K_1}{2\omega_0}\cos 2\omega_0 t\right)\right] .
\end{aligned} \tag{5.20}$$

If K_1 is small compared with ω_s, the following approximations for $\sigma^2_{\Delta\dot{r}_x}(t)$ and $\sigma^2_{\Delta r_x}(t)$ result:

$$\sigma^2_{\Delta\dot{r}_x}(t) \simeq \tau_{\delta v_x}\sigma_{\delta v_x}K_1(1 - e^{-K_1 t}) \tag{5.21}$$

$$\sigma^2_{\Delta r_x}(t) \simeq \frac{\tau_{\delta v_x}\sigma^2_{\delta v_x}K_1}{\omega_s^2}(1 - e^{-K_1 t}) . \tag{5.22}$$

Equations (5.21) and (5.22) show that the "filtering" of the Doppler noise through the inertial system depends largely on the gain factor K_1. Also, the time constant for the variance of the velocity and the position

error are inversely proportional to K_1. These facts give additional criteria for K_1. If K_2 and K_3 are also $\neq 0$, it is difficult to obtain a closed-form solution of Equation (5.7). In this case a computer program may be necessary to evaluate the integral and to obtain criteria for the values of K_2 and K_3.

6. STATISTICAL FILTERING AND UPDATING

The updating mechanism described in Section 4 used a deterministic approach. The advent of very powerful navigation computers made it possible to retrieve more of the information contained in the navigation measurements by applying statistical filtering techniques.

In this section the problem of optimal estimation and control of aided inertial system state variables is discussed. Although the problem of optimal estimation and filtering is actually a fairly old one, it has only recently yielded a decisive breakthrough by Kalman and others. The classical theory of Wiener and Kolmogoroff has, as an essential disadvantage, operation with the difficult formalism of integral equations. The Kalman discrete filter equations are purely algebraic; hence they are particularly well suited for processing on modern digital computers. Common to both theories is the fact that they are applicable only to linear systems and that they make use of the same criterion of being optimal, namely the smallest mean square error. Moreover, there exists an interesting connection between the two theories: if the Wiener-Hopf equation is extended to non-stationary processes and presented in matrix form, it is an exact analogy to the orthogonality theorem of the optimal Kalman filter equations[1]. This orthogonality theorem states that the mean square estimation error becomes a minimum if, and only if, the measurement error and the estimation error of the system state are orthogonal to each other.

Inertial navigation systems can be expressed mathematically by nonlinear differential equations. But the navigation errors are well describable by linear differential equations using only the first terms in the series expansion about the nominal values. The error sources, for example gyro drift, accelerometer bias etc., appear as forcing functions. If these error sources, as well as the error sources of the reference measurements, are considered as statistical variables with zero means and known second-order moments, then the modern Kalman theory of optimal filtering applies very well to aided inertial systems. In this way optimum use is made of the external measurements as additional information and of the known statistics of the error sources, to obtain the optimal system gain factors for every time instant and to keep the navigation errors as small as possible. It is important to recognise that, contrary to the conventional updating methods described in Section 4, the gain factors are not constant. They are recomputed in every Kalman cycle in order to use the redundant information in the optimal way within the system.

The Kalman filtering theory and the derivation of the equations is fully discussed elsewhere (References 8, 13, 4, 14, etc.) and can therefore be deleted.

The most important properties of a Kalman filter can be summarized as follows:

The filter estimates are all variables of the state vector in the least mean square error sense.

The estimation is based upon statistical data of all error sources and is completely carried out in the time domain.

The filter formulae satisfy minimum variance criteria for all problem parameters.

The formulae implemented are recursive. This means that the optimum estimate for the time being can be computed from the previous estimate and the current observation without recourse to earlier estimates or observations.

The recursive formulae are well suited to digital computers.

The basic method of using a linear optimal filter within an aided inertial system is shown in Figure 9.

The two upper blocks in Figure 9 represent a pure inertial system. The navigation equations are mechanised in the navigation computer in the conventional way. Compensation for sensor biases may be included in the navigation equations.

The information available in the navigation computer is compared with redundant external data. The differences are the measurement data used in the filter to improve the estimate of the system states.

These error estimates of the filter may be used either for updating the navigation information, and thus for correction of IMU errors (closed loop operation), or only for improvement of the error states (open loop operation)[2]. The block "controller" regulates this operation.

If an inertial navigation system is mechanised in the open loop operation, all errors and their external measurements can be represented by a set of two linear vector equations

$$\dot{\vec{x}}(t) = \underline{A(t)}\vec{x}(t) + \vec{n}(t) \tag{6.1}$$

$$\vec{y}(t) = \underline{M(t)}\vec{x}(t) + \vec{q}(t) . \tag{6.2}$$

Equation (6.1) is the system equation. It describes the propagation of all errors occurring in an aided navigation system. The elements of the $n \times n$ matrix $\underline{A(t)}$ are known functions of time. The components of the white noise vector $\overrightarrow{n(t)}$ represent the forcing functions. Equation (6.2) is the measurement equation. It relates the values of the error state vector to the measurement data. The vector $\overrightarrow{q(t)}$ represents the disturbing white noise of the external measurements.

It is an essential supposition of the Kalman filter theory that the vector differential equation (6.1) can be solved by the same method used for deterministic differential equations, namely the conventional Picard iteration[18]. The so-called transition matrix Φ_{t/t_0}, which plays a fundamental role in the whole theory, appears in this solution. Using this fact and dividing the time axis into equally spaced discrete intervals (the sampling times of the computer) the dynamics of the linear system under consideration can be described in the form

$$\overrightarrow{x_{K+1}} = \underline{\Phi_{K+1/K}}\overrightarrow{x_K} + \overrightarrow{h_K} \tag{6.3}$$

$$\overrightarrow{y_K} = \underline{M_K}\overrightarrow{x_K} + \overrightarrow{q_K} . \tag{6.4}$$

where

$\overrightarrow{x_K}$ = system error state vector at time t_K

$\underline{\Phi_{K+1/K}}$ = $n \times n$ state transition matrix, relating the state vectors $\overrightarrow{x_{K+1}}$ and $\overrightarrow{x_K}$ in the noise-free case

$\overrightarrow{h_K}$ = additive white noise sequence representing the forcing function in the dynamic system

(Note: If Equation (6.3) were presented in the continuous form, then the vector

$$\overrightarrow{h(t)} = \int_{t_0}^{t} \underline{\phi(t,\tau)}\,\overrightarrow{n(\tau)}\,d\tau ,$$

as the response of the system to the white noise input $\overrightarrow{n(t)}$, would be a correlated random process. But in the discrete form the values of the integrals

$$\overrightarrow{h_K} = \int_{t_K}^{t_{K+1}} \underline{\phi(t_{K+1},\tau)}\,\overrightarrow{n(\tau)}\,d\tau$$

for different K are independent of each other; therefore the vector $\overrightarrow{h_K}$ can be represented as a white noise sequence with the property

$$E[\overrightarrow{h_K}\overrightarrow{h_L}^T] = \underline{H_K}\delta_{KL}).$$

$\overrightarrow{y_K}$ = observation vector of measurement data at time t_K linearly related to $\overrightarrow{x_K}$

$\underline{M_K}$ = $m \times n$ measurement matrix at time t_K relating $\overrightarrow{y_K}$ to $\overrightarrow{x_K}$ in the noise-free case

$\overrightarrow{q_K}$ = additive white noise sequence corrupting measurement data at time t_K.

The following five basic equations represent the Kalman filter:

$$\hat{\overrightarrow{x_K}} = \overrightarrow{x_K^*} + \underline{B_K}(\overrightarrow{y_K} - \overrightarrow{y_K^*}) \tag{6.5}$$

$$\overrightarrow{x_{K+1}^*} = \underline{\Phi_{K+1/K}}\hat{\overrightarrow{x_K}} \tag{6.6}$$

$$\underline{B_K} = \underline{P_K^*}\underline{M_K^T}(\underline{M_K}\underline{P_K^*}\underline{M_K^T} + \underline{Q_K})^{-1} \tag{6.7}$$

$$\underline{P_K} = (\underline{I} - \underline{B_K}\underline{M_K})\underline{P_K^*} \tag{6.8}$$

$$\underline{P_{K+1}^*} = \underline{\Phi_{K+1/K}}\underline{P_K}\underline{\Phi_{K+1/K}^T} + \underline{H_K} . \tag{6.9}$$

where

$\overrightarrow{y_K^*} = \underline{M_K}\overrightarrow{x_K^*}$

$\overrightarrow{x_K^*}$ = *a priori* estimate of the system state vector predicted at time t_K before using redundant measurement data

$\hat{\overrightarrow{x_K}}$ = *a posteriori* estimate of the system state vector at time t_K after using redundant measured data

$\underline{P_K}$ = $n \times n$ covariance matrix of the *a posteriori* estimation error $\overrightarrow{x_K} - \hat{\overrightarrow{x_K}}$

$\underline{P_K^*}$ = $n \times n$ covariance matrix of the *a priori* estimation error $\overrightarrow{x_K} - \overrightarrow{x_K^*}$

$\underline{H_K}$ = $E[\overrightarrow{h_K}\overrightarrow{h_K}^T]$ = $n \times n$ covariance matrix of the system white noise vector

$\underline{Q_K}$ = $E[\overrightarrow{q_K}\overrightarrow{q_K}^T]$ = $m \times m$ covariance matrix of the measurement white noise vector

$\underline{B_K}$ = weighting (gain) matrix at time t_K.

These equations are processed on an iterative basis to determine consecutively the optimum estimate $\hat{\vec{x}}_K$ of the system state at time t_K. The mechanisation of an aided inertial navigation system in an open loop operation as just described is not satisfactory, especially for long missions. The outputs of the inertial navigation system could be corrected by subtracting the estimated errors; but, since some errors in the inertial system are not bounded, the linear error model of the navigator would become invalid. The requirement of small errors and linear error propagation can only be achieved by controlling the state vector in the closed loop operation. For such a controlled system some of the above equations must be modified slightly. Instead of Equations (6.1) and (6.3) the following equations are valid*:

$$\dot{\vec{x}}(t) = \underline{A(t)}\vec{x}(t) + \underline{D(t)}\vec{u}(t) + \vec{n}(t) \tag{6.10}$$

$$\vec{x}_{K+1} = \underline{\Phi_{K+1/K}}\vec{x}_K + \underline{\Gamma_{K+1/K}}\vec{u}_K + \vec{h}_K \,. \tag{6.11}$$

The terms $\underline{D(t)}\vec{u}(t)$ and $\underline{\Gamma_{K+1/K}}\vec{u}_K$ represent the above mentioned correction signals entering the system to be controlled; in the ideal case of a completely controllable system the matrices $\underline{D(t)}$ and $\underline{\Gamma_{K+1/K}}$ are of dimension $n \times n$.

The modification of the filter Equations (6.5) to (6.9) results in a slight change of Equation (6.6). This result is noticeable, since the original equations given for the uncontrolled system are derived in the literature under the important assumption that the mean values of the input forcing functions are zero, i.e. that always the condition is valid $E[\vec{n}(t)] = 0$ and $E[\vec{h}_K] = 0$, respectively. In the controlled system the condition of zero mean value inputs is not achieved, because now the forcing functions consist not only of the white noise vectors $\vec{n}(t)$ or $\vec{h}_K$, but also of control signals which are treated as known deterministic functions, based upon the computed estimate of the state vector. In most cases the condition of $E[\underline{D(t)}\vec{u}(t)] = 0$ or $E[\underline{\Gamma_{K+1/K}}\vec{u}_K] = 0$ is not valid for the control signals, because the estimator $\hat{\vec{x}}_K$ is unbiased and the initial value $E[\hat{\vec{x}}_0]$ does not need to be zero (Section D1 of Reference 13). Nevertheless it is obvious (and can be shown by minimising the mean square estimation error (Section C of Reference 13), that the modified Equation (6.6) now becomes

$$\vec{x}^*_{K+1} \quad \underline{\Phi_{K+1/K}}\hat{\vec{x}}_K \quad \underline{\Gamma_{K+1/K}}\vec{u}_K \,. \tag{6.12}$$

Equations (6.10) to (6.12) would be valid if all components of the state vector were controllable. Unfortunately the ideal case of complete controllability is not possible in an inertial navigation system. The controllable states are in most cases restricted to errors in position, velocity, and platform orientation. The effects of the other variables can also be removed by including compensation signals in the control signals, based upon the Kalman filter estimates, but nevertheless these variables are not controllable in a strict sense, because they are not accessible to direct measurements and corrections.

In an inertial navigation system the state variables are usually separated into two parts, the first one containing the controllable states, the other one the uncontrollable states. Thus the dynamics of such a system can be described in the following form:

$$\left[\begin{array}{c} \vec{x}^1_{K+1} \\ \hline \vec{x}^2_{K+1} \end{array}\right] = \left[\begin{array}{c|c} \underline{\Phi^{11}_{K+1/K}} & \underline{\Phi^{12}_{K+1/K}} \\ \hline \underline{0} & \underline{\Phi^{22}_{K+1/K}} \end{array}\right] \cdot \left[\begin{array}{c} \vec{x}^1_K \\ \hline \vec{x}^2_K \end{array}\right] + \left[\begin{array}{c} \underline{\Gamma^1_{K+1/K}} \\ \hline \underline{0} \end{array}\right] \vec{u}_K + \left[\begin{array}{c} \vec{h}^1_K \\ \hline \vec{h}^2_K \end{array}\right] . \tag{6.13}$$

The control signals for the time interval from t_K to t_{K+1} should be based upon the best estimate of the state vector $\vec{x}^*_K$ available at time t_K. But, since at this point in time a new measurement is made and the computation of the weighting matrix B_K by Equations (6.7) to (6.9) requires some time, the *a posteriori* estimate $\hat{\vec{x}}_K$ cannot be available at the time instant t_K. Therefore, the control signals $\underline{\Gamma^1_{K+1/K}}\vec{u}_K$ must be computed from the *a priori* estimate $\vec{x}^*_K$.

Considering the fact mentioned above and the so-called separation theorem[3,4,6] the following relation can be derived:

$$\underline{\Gamma^1_{K+1/K}}\vec{u}_K = [\underline{\Phi^{11}_{K+1/K}} \,|\, \underline{\Phi^{12}_{K+1/K}}] \cdot \left[\begin{array}{c} \vec{x}^{1*}_K \\ \hline \vec{x}^{2*}_K \end{array}\right] . \tag{6.14}$$

Substituting Equation (6.14) into Equation (6.12) the *a priori* and *a posteriori* estimates of the separated state vector in the Kalman equations are represented by the following equations:

$$\left.\begin{aligned} \vec{x}^{1*}_{K+1} &= [\underline{\Phi^{11}_{K+1/K}} \,|\, \underline{\Phi^{12}_{K+1/K}}] \, \underline{B_K}(\vec{y}_K - \vec{y}^*_K) \\ \vec{x}^{2*}_{K+1} &= \underline{\Phi_{K+1/K}}\hat{\vec{x}}^2_K \end{aligned}\right\} \tag{6.15}$$

---

* Note: It is permissible, in order to make the transition from Equation (6.10) to (6.11), to consider the vector $\vec{u}(t)$ as a constant $\vec{u}(t_K) = \vec{u}_K$ over the Kalman cycle interval. The term $\underline{\Gamma_{K+1/K}}\vec{u}_K$ is then given by the expression

$$\underline{\Gamma_{K+1/K}}\vec{u}_K = \int_{t_K}^{t_{K+1}} \underline{\Phi(t_{K+1},\tau)\, D(\tau)}\, d\tau \cdot \vec{u}(t_K) \,.$$

$$\begin{bmatrix} \vec{x}_K^{1+} \\ \hline \vec{x}_K^{2+} \end{bmatrix} = \begin{bmatrix} \vec{x}_K^{1-} \\ \hline \vec{x}_K^{2-} \end{bmatrix} + \underline{B}_K (\vec{y}_K - \vec{y}_K^{\circ}) \quad . \tag{6.16}$$

Some remarks are now made with regard to the state vector, the computation of the transition matrix $\underline{\Phi}$, and the computation of the matrix $\underline{H}_K$.

The state vector $\vec{x}_K$ must include all important errors and error sources in the system and the associated sensors which cannot be represented by white noise. It contains the navigation errors, i.e. errors in position, velocity and platform orientation, errors of the inertial sensors and their respective misalignment on the platform, and errors of the redundant sensors. The sensor errors may originate in the sensor itself, in the environment or in the measurement principle. It is important to note that all errors which can be represented by white noise must be contained in $\vec{h}_K$ and $\vec{q}_K$, while all the other errors must appear in the state vector $\vec{x}_K$. The time behaviour of the latter ones must be describable either by linear homogeneous or by linear inhomogeneous differential equations, so that they can be included in the system of Equations (6.1) or (6.10). Many of the uncontrollable state variables are supplements to the state vector, necessary when the respective error sources are not describable by white noise sequences.

The most important error models are Markov processes, random walks and constant random biases. The correlated random processes can be generated from white noise by linear systems (shaping filters).

The Markov process, whose power spectral density has the form

$$\phi(\omega) = \frac{2\sigma^2\beta}{\omega^2 + \beta^2} \tag{6.17}$$

can be described in the time domain by the linear differential equation

$$\dot{f}(t) + \beta f(t) = \sqrt{(2\sigma^2\beta)}\, w(t) \quad , \tag{6.18}$$

where $w(t)$, as the driving function, represents a unity variance white noise source.

A random walk process is obtained as the output of an integrator whose input is white noise. Its differential equation is of the form

$$\dot{f}(t) = \sqrt{(r_{rw})}\, w(t) \, . \tag{6.19}$$

A constant random bias can also be described by a linear differential equation, in fact by the most simple one which exists, namely,

$$\dot{f}(t) = 0 \, . \tag{6.20}$$

Its solution, supplied with a random initial condition, represents the desired constant bias error.

It can be seen that this practice of supplementing the state vector is a convenient method if the error source forcing functions do not represent uncorrelated, zero-mean random processes. But this procedure may lead to a considerable number of state variables. Since the computational extent grows excessively with the number of variables (approximately by the third power), one is often constrained to consider only the most important error sources; this means to "suboptimise" the system. Another method of suboptimising, which results in a procedure similar to that described above for the controllable and uncontrollable states, is the partitioning of the state vector by placing the variables to be eliminated in one subsystem and the remaining variables in one or more additional subsystems. By this means the correlation between the variables in different subsystems is, in fact, ignored and the estimate is somewhat deteriorated as compared to the optimal case, but this reduction of the matrix dimensions results in some savings in computing time for the approximate optimal weighting matrix[10]. In every case of designing such a suboptimal system, the performance deterioration due to elimination or partitioning of some state variables should be predetermined by digital computer simulations and compared with optimal results. For example, using the method of partitioning the state vector, the "best" partitioning depends also on the system conditions.

The computation of the transition matrix $\underline{\Phi}_{t/t_0}$ will now be discussed. It is known that $\underline{\Phi}_{t/t_0}$ can be obtained from the coefficient matrix $\underline{A}(t)$ by the differential equation[12,15]

$$\frac{d\underline{\Phi}_{t/t_0}}{dt} = \underline{A}(t)\,\underline{\Phi}_{t/t_0} \, . \tag{6.21}$$

where

$$\underline{\Phi}_{t_0/t_0} = \underline{I} \quad \text{(identity matrix)}$$

is the initial condition.

In the general case an explicit closed solution of this differential equation is not possible, but an infinite series can be derived, which converges uniformly in t for every matrix $\underline{A}(t)$. This solution is given as follows[15]:

$$\underline{\phi}_{t/t_0} = \underline{I} + \int_{t_0}^{t} \underline{A}(\tau_1)\,d\tau_1 + \int_{t_0}^{t} \underline{A}(\tau_1) \int_{t_0}^{\tau_1} \underline{A}(\tau_2)\,d\tau_2 d\tau_1 + \int_{t_0}^{t} \underline{A}(\tau_1) \int_{t_0}^{\tau_1} \underline{A}(\tau_2) \int_{t_0}^{\tau_2} \underline{A}(\tau_3)\,d\tau_3 d\tau_2 d\tau_1 + \ldots \quad (6.22)$$

In the case of constant coefficients, $\underline{A}$ = constant, Equation (6.22) can be transformed into the form

$$\underline{\phi}_{t/t_0} = \underline{I} + \underline{A}(t-t_0) + \underline{A}^2 \frac{(t-t_0)^2}{2!} + \ldots = e^{\underline{A}(t-t_0)}. \quad (6.23)$$

This result can also be derived directly from the differential Equation (6.21). However, in most cases the matrix $\underline{A}$ is not constant, at least some of the elements a_{ik} are functions of time. In these cases the computation of the terms in Equation (6.22) is laborious, and it is better to subdivide the whole interval $(t-t_0)$ and to consider $\underline{A}$ as a constant matrix during these partial intervals. The time-dependent elements of the matrix $\underline{A}$ are now approximated by a staircase function. By this means Equation (6.23) can be used to compute the partial transient matrices $\underline{\phi}_{t_p/t_{p-1}}$ and it can be shown[18] that the matrix $\underline{\phi}_{t/t_0}$ over the whole interval $(t-t_0)$ is simply the product of the partial matrices $\underline{\phi}_{t_p/t_{p-1}}$.

$$\underline{\phi}_{t/t_0} = \underline{\phi}_{t=t_p/t_{p-1}} \cdot \underline{\phi}_{t_{p-1}/t_{p-2}} \cdot \ldots \cdot \underline{\phi}_{t_2/t_1} \cdot \underline{\phi}_{t_1/t_0} \quad (6.24)$$

$$t = t_p > t_{p-1} > t_{p-2} > \ldots > t_1 > t_0$$

It is important to note that by this method the iteration time for the computation of the Kalman equations does not need to agree with the iteration time for computing the partial matrices $\underline{\phi}_{t_p/t_{p-1}}$. Thus essentially longer cycle times for the computation of the Kalman equations can be chosen, depending on the speed of the available computer and the times when additional redundant measurements are available. The matrix $\underline{\phi}_{K+1/K}$ over the whole Kalman cycle interval must in these cases be computed using Equation (6.24). It is of great importance to pay attention to the number of partial intervals into which the Kalman cycle interval is divided. If it is subdivided too finely, the linear term in the series expansion (6.23) is sufficient to compute the partial matrix $\underline{\phi}_{t_p/t_{p-1}}$, but then the computation of the total matrix $\underline{\phi}_{t/t_0}$ may consume a too great portion of the available computer time because of the great number of multiplications and additions contained in Equation (6.24). This is especially true, when the state vector contains many variables. It is often better, with regard to computing time and also with regard to accuracy, to subdivide the Kalman cycle for computation of the partial matrices $\underline{\phi}_{t_p/t_{p-1}}$ less finely and, instead, to take the quadratic term of Equation (6.23) into account. The matrix $\underline{\phi}_{t_p/t_{p-1}}$ for the time interval t_{p-1} to t_p can then be approximated by the equation

$$\underline{\phi}_{t_p/t_{p-1}} \approx \underline{I} + \underline{A}(t_{p-1})(t_p - t_{p-1}) + \underline{A}^2(t_{p-1}) \frac{(t_p - t_{p-1})^2}{2!}. \quad (6.25)$$

For every system it is necessary to determine which of the two methods is better in a practical case. Probably the first method (finer subdividing and consideration of only the linear term) is to be preferred in a real system, in which the matrix $\underline{A}(t)$ may change very quickly, whereas the second method is to be preferred for system performance error analysis, where comparatively slow changes of the matrix $\underline{A}(t)$ occur.

The computation of the matrix $\underline{H}_K$ can be derived from the vector $\vec{h}_K$ as follows:

$$\vec{h}_K = \int_{t_K}^{t_{K+1}} \underline{\phi}(t_{K+1},\tau)\,\vec{n}(\tau)\,d\tau \quad (6.26)$$

$$\underline{H}_K = E[\vec{h}_K \vec{h}_K^T] = E\left[\int_{t_K}^{t_{K+1}} \underline{\phi}(t_{K+1},\tau_1)\,\vec{n}(\tau_1)\,d\tau_1 \int_{t_K}^{t_{K+1}} \vec{n}(\tau_2)^T\,\underline{\phi}(t_{K+1},\tau_2)^T\,d\tau_2 \right].$$

Interchanging the order of averaging and integration, and defining the covariance matrix of the white noise vector $\vec{n}(t)$ as

$$E[\vec{n}(t_1)\vec{n}(t_2)^T] = \underline{N}(t_1)\,\delta(t_1 - t_2), \quad (6.27)$$

one can write further

$$\underline{H}_K = \int_{t_K}^{t_{K+1}} \underline{\phi}(t_{K+1},\tau_1)\,\underline{N}(\tau_1)\,d\tau_1 \int_{t_K}^{t_{K+1}} \delta(\tau_1 - \tau_2)\,\underline{\phi}(t_{K+1},\tau_2)^T\,d\tau_2$$

$$= \int_{t_K}^{t_{K+1}} \underline{\phi}(t_{K+1},\tau)\,\underline{N}(\tau)\,\underline{\phi}(t_{K+1},\tau)^T\,d\tau. \quad (6.28)$$

In general, Equation (6.28) must be approximated by numerical methods. For example, if $\underline{N}(t)$ is considered as a constant matrix $\underline{N}_K$ over the whole Kalman interval $(t_{K+1} - t_K)$ and if, according to Equation (6.24), this interval is divided into an even number p of sub-intervals of equal length $\Delta t = t_{s+1} - t_s$, then Equation (6.28) can very well be approximated by Simpson's rule:

$$\underline{H}_K \simeq \frac{\Delta t}{3}\,(\underline{\Phi}_{t_{K+1}=t_p/t_{p-1}} \cdots \underline{\Phi}_{t_1/t_0=t_K}\underline{N}_K\underline{\Phi}^T_{t_1/t_0=t_K} \cdots \underline{\Phi}^T_{t_{K+1}=t_p/t_{p-1}} +$$

$$+ 4\underline{\Phi}_{t_{K+1}=t_p/t_{p-1}} \cdots \underline{\Phi}_{t_2/t_1}\underline{N}_K\underline{\Phi}^T_{t_2/t_1} \cdots \underline{\Phi}^T_{t_{K+1}=t_p/t_{p-1}} +$$

$$+ 2\underline{\Phi}_{t_{K+1}=t_p/t_{p-1}} \cdots \underline{\Phi}_{t_2/t_2}\underline{N}_K\underline{\Phi}^T_{t_2/t_2} \cdots \underline{\Phi}^T_{t_{K+1}=t_p/t_{p-1}} +$$

$$+ \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad +$$

$$+ 2\underline{\Phi}_{t_{K+1}=t_p/t_{p-1}}\underline{\Phi}_{t_{p-1}/t_{p-2}}\underline{N}_K\underline{\Phi}^T_{t_{p-1}/t_{p-2}}\underline{\Phi}^T_{t_{K+1}=t_p/t_{p-1}} +$$

$$+ 4\underline{\Phi}_{t_{K+1}=t_p/t_{p-1}}\underline{N}_K\underline{\Phi}^T_{t_{K+1}=t_p/t_{p-1}} + \underline{N}_K) \,. \qquad (6.29)$$

Of course, this solution can only be of theoretical interest if $p > 2$, because the computation of Equation (6.29) becomes too laborious. In most cases, it will be sufficient to approximate Equation (6.28) by the simple expression

$$\underline{H}_K \simeq \frac{t_{K+1} - t_K}{2}\,(\underline{\Phi}_{K+1/K}\underline{N}_K\underline{\Phi}^T_{K+1/K} + \underline{N}_K) \,. \qquad (6.30)$$

No general expression can be derived for the measurement matrix $\underline{M}_K$, because it depends upon the mechanized system and the mode of operation.

In Section 5 of this chapter the conventional statistical error analysis of aided inertial systems was described. With the aid of the Kalman filter equations a statistical performance analysis with minimum error variances can be accomplished very effectively. For such an analysis only the last three of the five Kalman equations, namely (6.7) to (6.9), are required. It is an advantage that, in an ideal controlled system, the matrix $\underline{P}^*_K$ gives the variances of the states themselves, either directly or after simple transformations, for in such a system the error state variances will be reduced to a level equal to the variance with which each state is estimated. Therefore the following definitions are valid for the matrix $\underline{P}^*_K$:

Uncontrolled system (open loop operation)	*Ideal controlled system* (closed loop operation)
$\underline{P}^*_K = E[(\vec{x}_K - \vec{\hat{x}}^*_K)(\vec{x}_K - \vec{\hat{x}}^*_K)^T]$	$\underline{P}^*_K = E[\vec{x}^*_K \vec{x}^{*T}_K]$. (6.31)

As pointed out earlier, no ideal controllable system exists in reality. But a relation can also be derived for the covariances of the controllable states of a real system, in which the state vector is separated into controllable and uncontrollable states, as shown in Equation (6.13). By substituting Equation (6.14) into Equation (6.13) with negative sign (corresponding to the way in which the control signals enter the system) the following relation is yielded:

$$\vec{x}^{\,I}_{K+1} = [\underline{\Phi}^{II}_{K+1/K} \mid \underline{\Phi}^{IS}_{K+1/K}] \cdot \begin{bmatrix} \vec{x}^{\,I}_K - \vec{\hat{x}}^{\,I*}_K \\ \hline \vec{x}^{\,S}_K - \vec{x}^{\,S*}_K \end{bmatrix} + \vec{h}^{\,I}_K \,. \qquad (6.32)$$

From Equation (6.32), finally,

$$E[\vec{x}^{\,I}_{K+1}\vec{x}^{\,I\,T}_{K+1}] = [\underline{\Phi}^{II}_{K+1/K} \mid \underline{\Phi}^{IS}_{K+1/K}]\,\underline{P}^*_K\,[\underline{\Phi}^{II}_{K+1/K} \mid \underline{\Phi}^{IS}_{K+1/K}]^T + E[\vec{h}^{\,I}_K \vec{h}^{\,I}_K] \,. \qquad (6.33)$$

If a short Kalman cycle time is considered, as compared to the time constants of the navigator and the associated error processes, the matrix $\underline{\Phi}^{II}_{K+1/K}$ does not differ much from the identity matrix and the matrix $\underline{\Phi}^{IS}_{K+1/K}$ is near zero. Besides, the vector $\vec{h}^{\,I}_K$ is zero in all practical cases, because the random forcing functions of the controllable states consist of correlated error sources. Therefore the appropriate part of $\underline{P}^*_K$ can be used as a good description of the actual covariances of the controlled states. Using these facts it is not necessary to carry out the solution of Equation (6.33), and the computational burden during performance analysis can be reduced substantially.

The error behaviour of an aided navigation system will now be investigated, using the method described for system performance analysis. It should be noted that, by this method, optimal values for the elements of the weighting matrix $\underline{B}_K$ are automatically computed. In contrast to the conventional, deterministic method described in Section 4, the effects of these gain factors will certainly be more complex, since they do not enter the system at a single definite place, but because they influence in an optimal way every system state. Accordingly they are, of course, not exactly equivalent to the formerly used deterministic gain factors K_1, K_2 or K_3, for it is easily possible to include these factors also in the matrix $\underline{A}(t)$ of the system Equation (6.1) or (6.10), respectively, and, in addition, to carry out a Kalman filtering. The meaning of the matrix $\underline{B}_K$ is more comprehensive and more profound. Mathematically $\underline{B}_K$ has the task of weighting the estimates with the redundant measurements and, in the optimal case, of establishing an orthogonal relation between the linear independent random vectors $\vec{x}_K - \vec{\hat{x}}_K$ (estimation error) and $\vec{y}_K - \vec{\hat{y}}^*_K$ (measurement error), and thus to accomplish the orthogonality theorem mentioned earlier.

Figures 10 and 11 show the propagation of the position and velocity errors of an inertial navigation system in different modes used to navigate over the surface of the earth. It can be seen that here also considerable improvement of the accuracy can be obtained by additional redundant informations, as compared to pure inertial navigation. The velocity error in the pure inertial mode shows the characteristic Schuler period and the 24 hour period. In contrast to Figure 8, not only the influence of a single error source but the entire influence of all error sources is considered. The influence of any single error source could also be investigated by a change of the method described for the system performance error analysis, but, since this is an objective of suboptimal studies, it will not be discussed further.

In this example the state vector $\overline{x_k}$ consists of 24 variables. The first 7 states are the well-known controllable inertial navigation errors whose propagation equations are derived in several textbooks (References 7, 11, etc.). The other variables are the uncontrollable error sources which cannot be represented by white noise.

REFERENCES

1. Brammer, K. — *Zur optimalen linearen Filterung und Vorhersage instationärer Zufallsprozesse in diskreter Zeit.* Regelungstechnik, Heft 3, 16 Jahrg., 1968, pp.105-110.

2. Brown, R.G., Friest, D.T. — *Optimization of a Hybrid Inertial Solar-Tracker Navigation System.* Institute of Electrical and Electronic Engineers, International Convention Record, New York, 1964, Part 7, March 23-26, 1964, pp.121-135.

3. Davenport, W.B., Jr, Root, W.L. — *An Introduction to the Theory of Random Signals and Noise.* International Student Edition, McGraw-Hill, New York, 1958.

4. Fagin, S.L. — *Recursive Linear Regression Theory, Optimal Filter Theory, and Error Analysis of Optimal Systems.* Institute of Electrical and Electronic Engineers, International Convention Record, Vol.12, Part 1, New York, March 1964, pp.216-245.

5. Gunckel, T.F. Franklin, G.F. — *A General Solution for Linear Sampled Data Control Systems.* Transactions, American Society of Mechanical Engineers, Journal of Basic Engineering, June 1963, pp.197-203.

6. Joseph, P., Tou, J.T. — *On Linear Control Theory.* American Institute of Electrical Engineers, Transactions (Applications and Industry), Vol.80, 1961, pp.193-196.

7. Kachikas, G.A. — *Error Analysis for Cruise Systems.* Chapter 7 of "Inertial Guidance", edited by G.R.Pitman Jr, University of California Engineering and Physical Sciences Extension Series, John Wiley, New York, 1962.

8. Kalman, R.E. — *A New Approach to Linear Filtering and Prediction Problems.* Transactions of the American Society of Mechanical Engineers, Series D, Volume 82D, 1960, pp.35-45.

9. Kalman, R.E. — *Contributions to the Theory of Optimal Control.* Bulletin of the Mexican Mathematical Society, Second Series, Volume 5, 1960, pp.102-119.

10. Pentecost, E.E., Stubberud, A.R. — *Synthesis of Computationally Efficient Sequential Linear Estimators.* Institute of Electrical and Electronic Engineers, Transactions, Vol.AES-3, No.2, March 1967.

11. Pinson, J.C. — *Inertial Guidance for Cruise Vehicles.* Part 2 of "Guidance and Control of Aerospace Vehicles", edited by C.T.Leondes, University of California Engineering and Sciences Extension Series, McGraw-Hill, New York, 1963.

12. Schweidler, W. — *Vorträge über Determinanten und Matrizen mit Anwendungen in Physik und Technik.* Berlin, 1949, pp.137-140.

13. Sorenson, H.W. — *Kalman Filtering Techniques.* Contribution to the book "Advances in Control Systems, Theory and Applications", Volume 3, edited by C.T.Leondes, Academic Press, New York, 1966.

14. Tou, J.T. — *Optimum Design of Digital Control Systems.* Academic Press, New York, 1963.

15. Zurmühl, R. — *Matrizen.* Springer, Berlin-Göttingen-Heidelberg, 1964, pp.442-446.

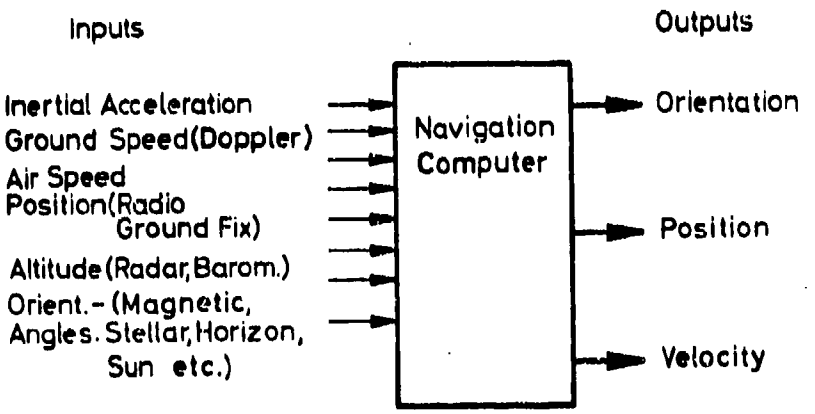


Fig. 1 Inputs and outputs of navigation systems

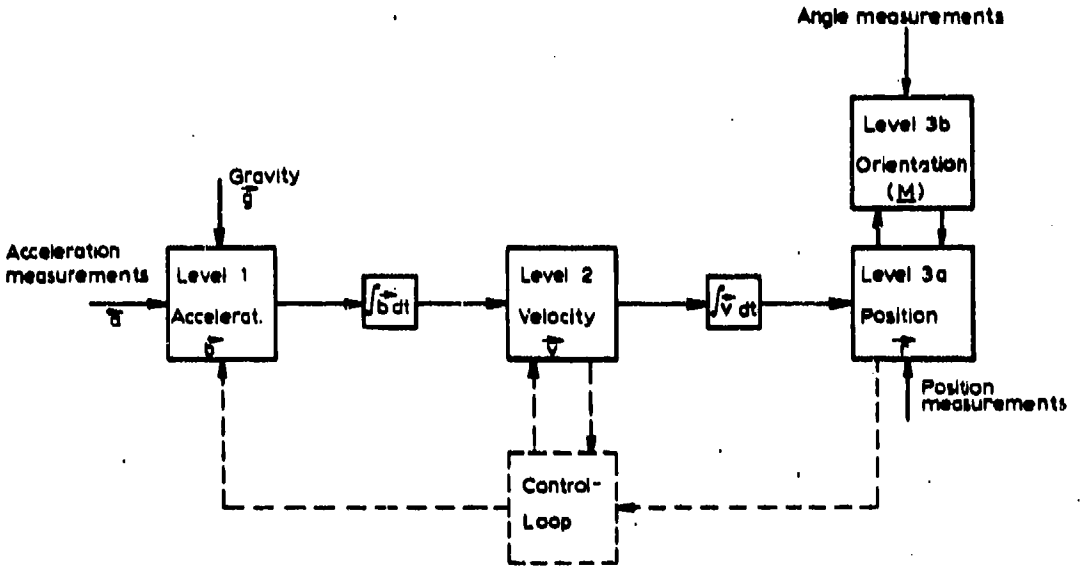


Fig. 2 Levels of navigation information

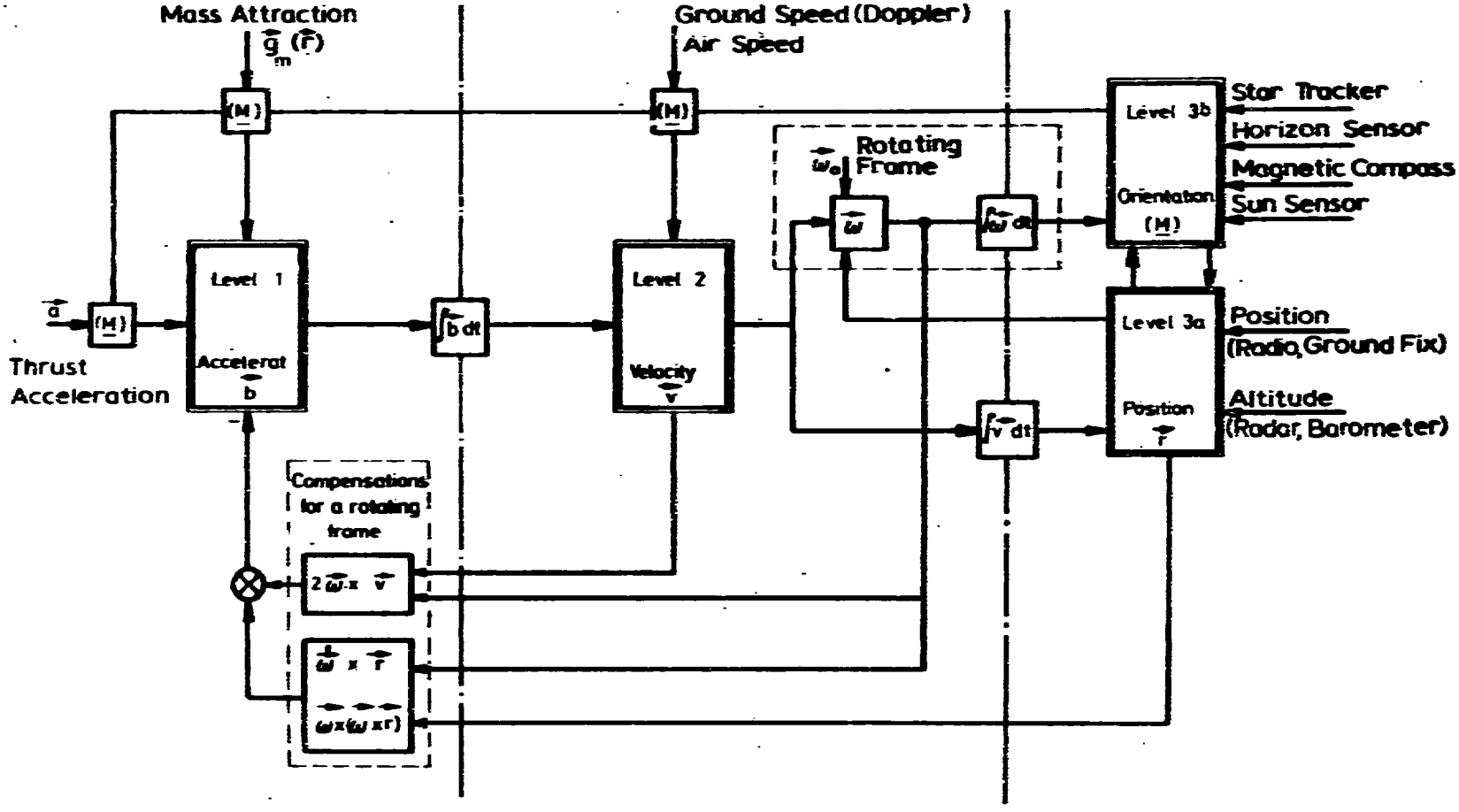


Fig.3 General block diagram of navigation systems

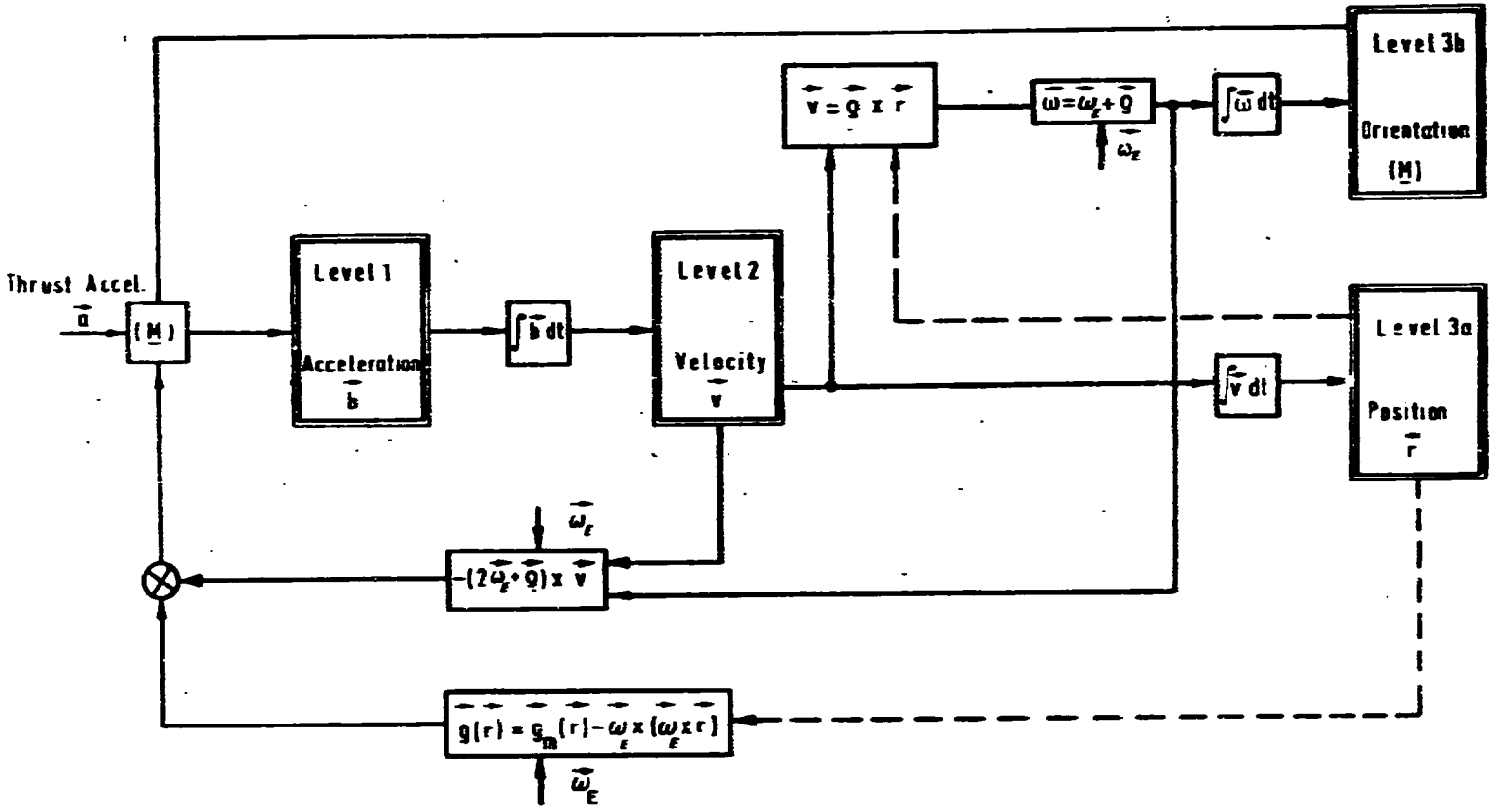


Fig.4 Vector block diagram of a pure inertial system used for earth navigation

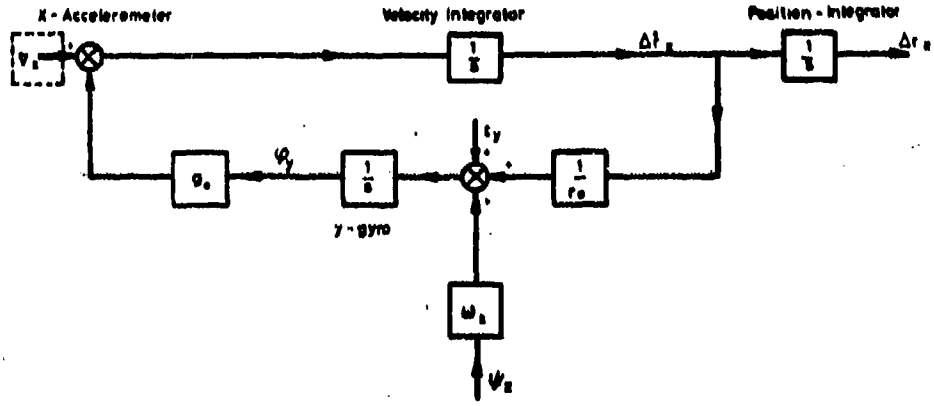


Fig. 5 Error block diagram for the x-channel of a Schuler-tuned system

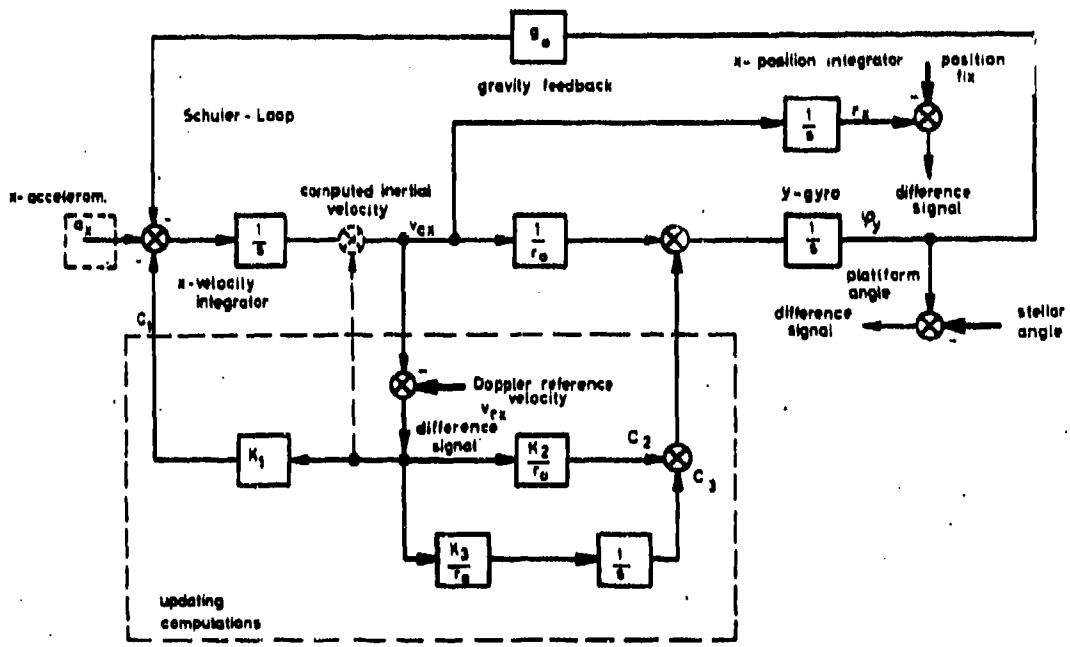


Fig. 6 Updating mechanisation

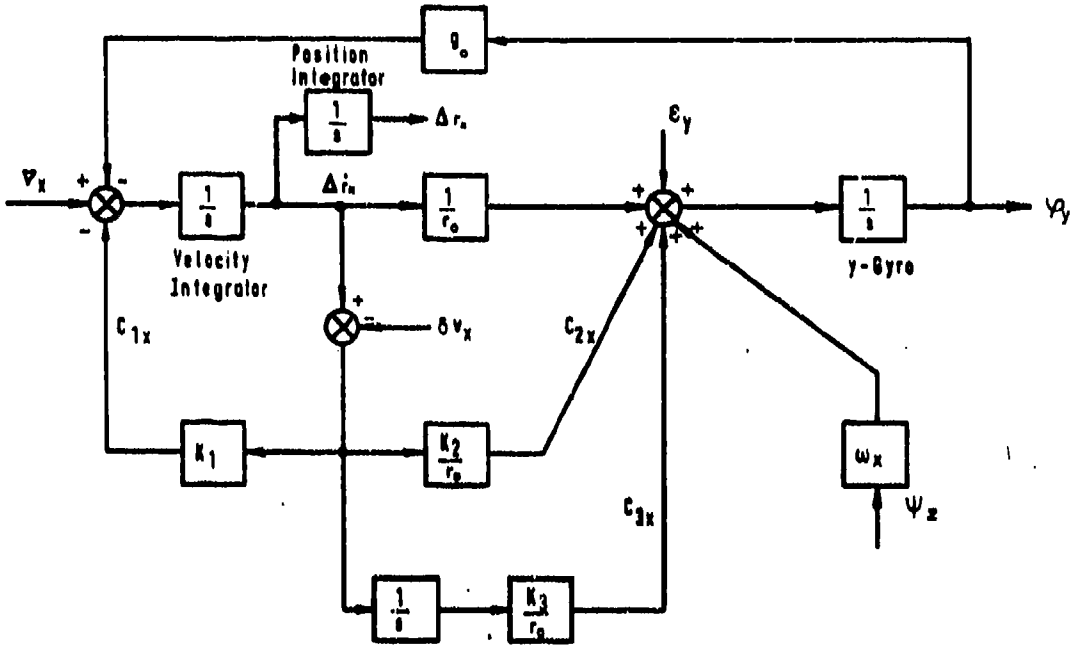


Fig.7 Error block diagram of a Doppler inertial system (x-channel)

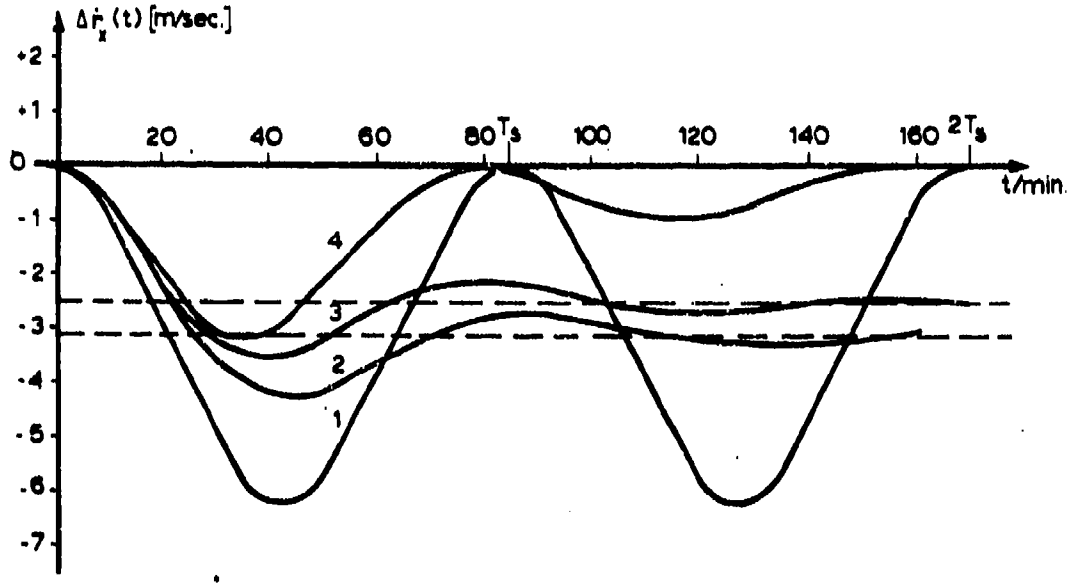


Fig.8 Velocity error due to fixed gyro drift ($\varepsilon_y = 0.1^0/h$) for different systems

1. Pure inertial system ($K_1 = 0$; $K_2 = 0$; $K_3 = 0$)
2. Doppler damped inertial system ($K_1 = 0.6\omega_s$; $K_2 = 0$; $K_3 = 0$)
3. Doppler inertial system ($K_1 = 0.6\omega_s$; $K_2 = 0.236$; $K_3 = 0$)
4. Doppler inertial system with an additional integrator ($K_1 = 0.6\omega_s$; $K_2 = 0.236$; $K_3 = 0.2312\omega_s$)

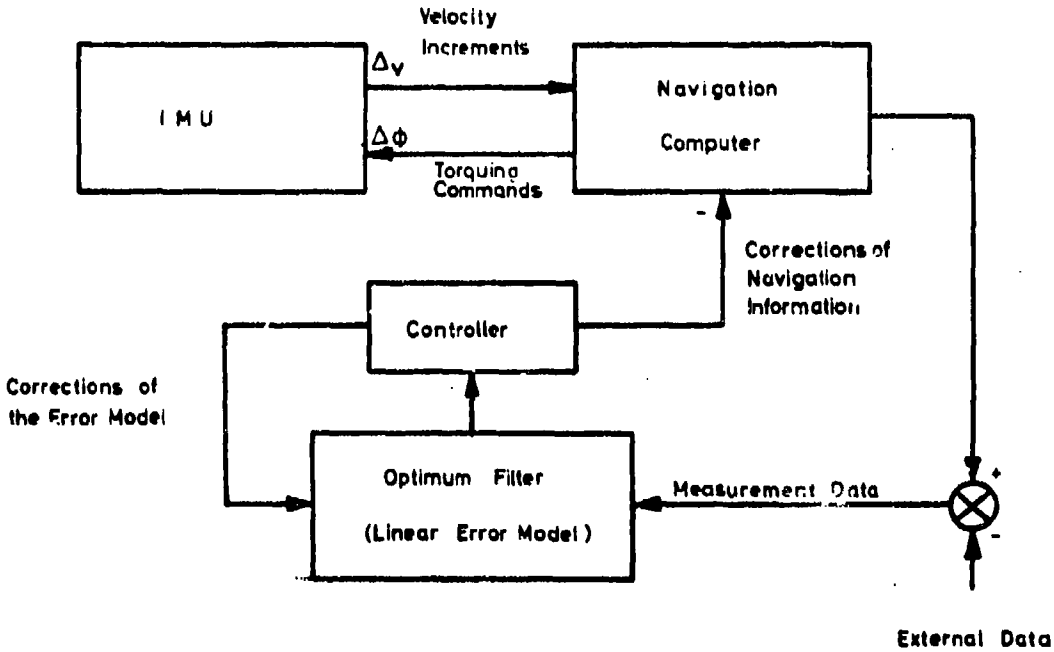


Fig.9 Block diagram of an inertial system with optimum filter

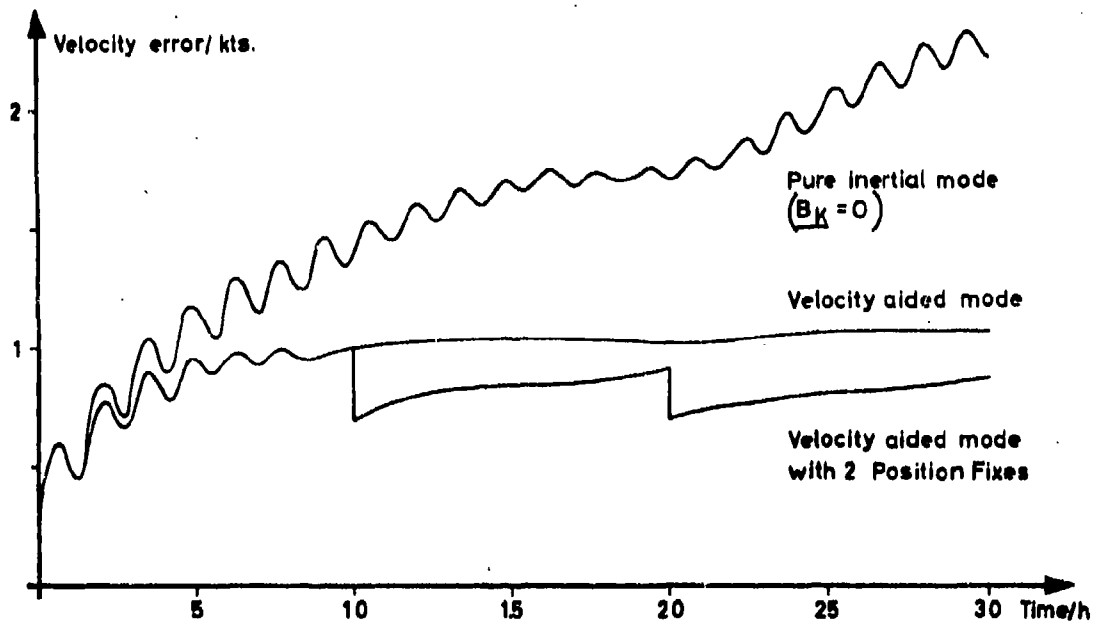


Fig. 11 Velocity error of an inertial system used for earth navigation

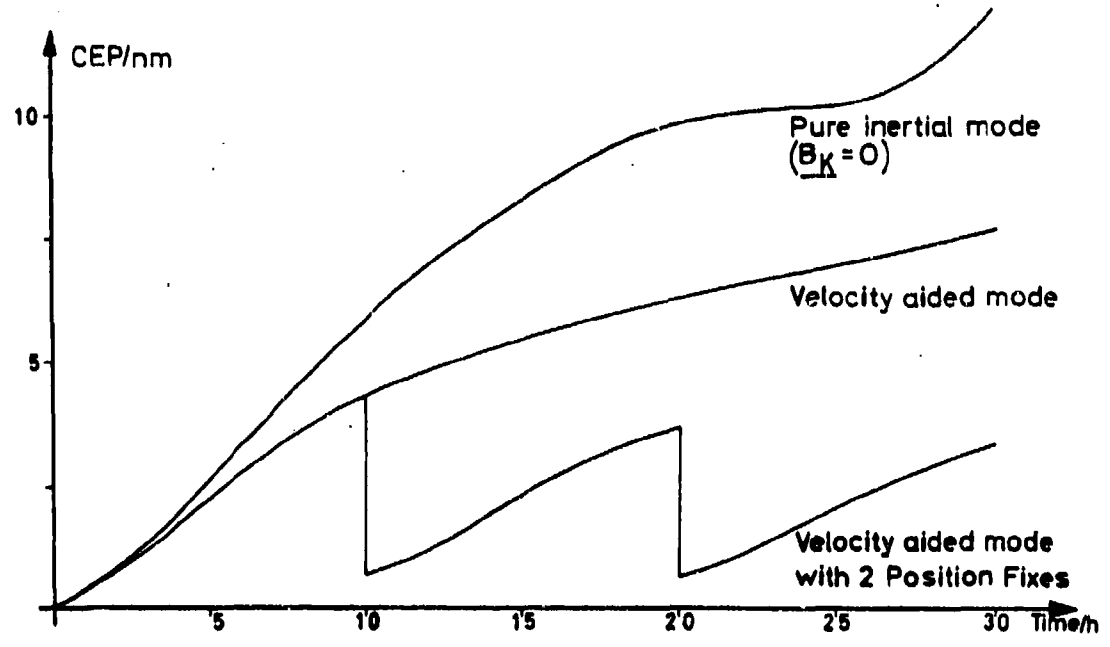


Fig. 10 CEP position error of an inertial system used for earth navigation

CHAPTER 19 - APPLICATION OF KALMAN FILTERING TECHNIQUES TO STRAPDOWN SYSTEM INITIAL ALIGNMENT

by

Earle B. Crocker and Leonard Rabins

Inertial Guidance Engineering Ordnance Systems
General Electric Company
100 Plastics Avenue
Pittsfield, Massachusetts, USA

CHAPTER 19 - APPLICATION OF KALMAN FILTERING TECHNIQUES TO STRAPDOWN SYSTEM INITIAL ALIGNMENT

Earle B. Crocker and Leonard Rabins

1. INTRODUCTION

Initial attitude alignment of systems involving inertial instruments is a requirement caused by the fact that sensors commonly used in inertial guidance, navigation, or attitude reference systems measure angular rates or integrals of angular rates (gyro outputs), or linear acceleration or integrals of linear accelerations (accelerometer outputs). When the need exists for continuous knowledge of absolute angular orientation of the frame or platform to which the sensors are mounted, then (just as initial values for integrals) the initial attitude must be established. This is the alignment problem which is considered in this chapter*. The basic objective is the investigation of digital methods for the determination of the initial alignment of a strapdown inertial navigation system from vibration and sway corrupted data on the launch pad.

The inertial components may be mounted on a stabilized platform or directly strapped down to an uncontrolled swaying body. In either case the essential mathematics of establishing alignment from outputs of the inertial components themselves is identical. In a strapdown inertial system, the measurements are made in the vehicle or body coordinate system. Angular information obtained from three single-axis platforms or three pulse-torqued gyros is used to update a coordinate transformation matrix relating a set of analytical or computing axes to the body axes. This coordinate transformation matrix is then used to transform the accelerometer outputs from the body axes to the computing axes. The initial alignment problem is concerned with determining the proper initial values for the elements of this coordinate transformation matrix.

If the initial computing axes are earth-fixed at the launch site, the problem is reduced to one of determining the orientation of these axes with respect to local vertical and an azimuth reference.

The azimuth angle can be determined from an optical measurement, as shown in Reference 1. The determination of vertical involves using the accelerometer outputs which have been resolved through the coordinate transformation matrix into a reference coordinate system. This system is maintained nominally earth-fixed at the launch site by using the gyro outputs to compensate for rotations caused by vehicle sway. This reference system will deviate from an earth-fixed system due to gyro drift because proper compensation for the earth's rate cannot be provided until the orientation of the reference system with respect to the earth is known.

If vehicle motions are zero, the determination of the orientation of the reference system with respect to vertical is trivial since the accelerometer outputs in the reference system are a simple function of the orientation and the acceleration of gravity. When there is vehicle motion, the output of the accelerometers due to gravity must be separated from the output due to vehicle motion by filtering.

Filtering methods investigated under this study were all based on the criterion of a minimum mean square error and differ mainly in the mechanization technique and the amount of *a priori* knowledge of the statistics which is assumed. A prime effort of the study is to investigate the accuracy of a Kalman filter in which both the state equations and noise are modeled differently from the "true" characteristics. A useful filter must exhibit little sensitivity to these modeling errors since the "true" characteristics, particularly the noise sources, are never known.

Methods considered included an optimal filter of the Kalman type, a simplified Kalman filter in which the effects of gyro drift and improper earth's rate compensation were neglected, a maximum likelihood approach, and a least-squares curve fit technique with and without preconditioning of the data. This chapter will emphasize the results and analysis of the Kalman filter, since the recommended approach was use of the simplified Kalman filter. The details of the analysis of the other techniques can be found in Reference 2.

2. ANALYSIS

2.1 Coordinate System Definition

Alignment of a coordinate system defining a strapdown or platform system implies knowledge of a basic reference system relative to which the alignment is performed. The reference system used here is earth fixed, with one axis along the local vertical corresponding to the average position of the swaying strapdown system. One horizontal

* This need obviously occurs for inertial guidance and other systems. See the book *Inertial Guidance* by D.H. Pitman.

axis is assigned an arbitrary azimuth. This coordinate system is designated CS(0). The strapdown reference analytic system is CS(1) and is the coordinate system in which all measurements by the strapdown instrument system CS(2) are referred. Alignment establishes the relationship between CS(0) and CS(1). The strapdown system computer further contains information expressing the relation between CS(1) and CS(2). In the following analysis CS(1) can be considered to coincide with CS(2) initially.

At the initial time CS(1) is misaligned from CS(0) by the small angle $(\alpha,\beta,\gamma) = \theta$, as shown in Figure 1.

While CS(0) rotates exactly at the earth's rotational rates, CS(1) cannot be correctly "torqued" since the misalignment angles α, β, γ are not known. This causes these angles to vary with time due to "earth's rate coupling". The angles, additionally, change due to error drift rate of the gyros.

Since x_0 is defined as locally vertical, a measurement of acceleration, A,* along axes z_1 and y_1 is proportional to the misalignment angles; thus,

$$A_{z_1} = g\beta \tag{1}$$

$$A_{y_1} = g\gamma . \tag{2}$$

In the strapdown system, the acceleration sensing instruments are actually located along axes of CS(2) (or some other axes whose relationship to CS(2) is known by physical alignment), but the measurements are transformed into CS(1) in the strapdown computer. The acceleration measured includes the effect of random horizontal accelerations caused by the missile sway. This is the principal noise input to which the filtering methods are applied.

2.2 Direction Cosine Relations Between Rotating Coordinate Systems

Certain basic direction cosine matrix equations are of importance in understanding relationships between rotating coordinate systems. Let ${}^1C_2$ denote the direction cosine matrix which transforms components (V_{2x}, V_{2y}, V_{2z}) of a vector V in coordinate system 2 to its components in coordinate system 1. Then,

$$V_1 = {}^1C_2 V_2 \tag{3a}$$

or, in detail,

$$\begin{pmatrix} V_{1x} \\ V_{1y} \\ V_{1z} \end{pmatrix} = \begin{pmatrix} C_{x_1,x_2} & C_{x_1,y_2} & C_{x_1,z_2} \\ C_{y_1,x_2} & C_{y_1,y_2} & C_{y_1,z_2} \\ C_{z_1,x_2} & C_{z_1,y_2} & C_{z_1,z_2} \end{pmatrix} \begin{pmatrix} V_{2x} \\ V_{2y} \\ V_{2z} \end{pmatrix} . \tag{3b}$$

The individual components of the ${}^1C_2$ matrix denote the cosine of the angle between the two subscripted axes.

Let coordinate system 1 be fixed in inertial space while system 2 rotates at a rate ω_2, where the angular rotation rate is measured or given as components along the axes of coordinate system 2. Then*

$$\frac{d\,{}^1C_2}{dt} = {}^1C_2\,\omega_2 , \tag{4}$$

where ω_2 is the skew-symmetric matrix

$$\omega_2 = \begin{bmatrix} 0 & -\omega_{z_2} & \omega_{y_2} \\ \omega_{z_2} & 0 & -\omega_{x_2} \\ -\omega_{y_2} & \omega_{x_2} & 0 \end{bmatrix} . \tag{5}$$

Finally, let coordinate system 0 be considered inertial, with systems 1 and 2 both rotating relative to the inertial system. Then, to establish the relation between systems 1 and 2, we note from Equation (4)

$$\frac{d\,{}^0C_1}{dt} = {}^0C_1\,\omega_1$$

and

$$\frac{d\,{}^0C_2}{dt} = {}^0C_2\,\omega_2 .$$

Since

$${}^0C_2 = {}^0C_1\,{}^1C_2 ,$$

it follows that

$$\frac{d\,{}^1C_2}{dt} = {}^1C_2\,\omega_2 - \omega_1\,{}^1C_2 . \tag{6}$$

---

* *Matrix Methods for Engineering* by L. A. Pipes, Prentice-Hall, 1963, p. 167.

2.3 System Dynamics

The equations describing the strapdown erection process are derived in this section. These include the differential equations governing the angles describing the misalignment between the earth-fixed launch-site coordinate system (CS) and the strapdown inertial reference CS, the transformed accelerometer outputs, and the missile sway velocity. The differential equations are then converted to a set of first-order difference equations, since this is the form appropriate for the application of the discrete-time Kalman filter equations. Certain assumptions made in this analysis are as follows:

(i) The initial misalignment angles are small enough so that the small angle approximations ($\sin\theta \approx \theta$, $\cos\theta \approx 1$) are valid.

(ii) Inertial component sensing axes are orthogonal to each other.

(iii) The gyro drift rates are constant.

(iv) Gyro torquing errors and accelerometer instrument errors are small and not considered.

(v) The sway velocity is adequately represented by a narrow-band noise process having zero mean value.

2.3.1 *Transformed Accelerometer Outputs*

The inertial instruments which measure linear acceleration are integrating devices, so that their outputs are proportional to the integral of acceleration over fixed time intervals. In particular, the outputs along y_1 and z_1 axes are given by

$$P_1(t_i) = \int_{t_{i-1}}^{t_i} g\beta(\tau)\, d\tau + V_1(t_i) - V_1(t_{i-1}) \tag{7}$$

$$P_2(t_i) = \int_{t_{i-1}}^{t_i} -g\gamma(\tau)\, d\tau + V_2(t_i) - V_2(t_{i-1}) , \tag{8}$$

where $P_1(t)$ is the output along z_1 at time t

$P_2(t)$ is the output along y_1 at time t

$V_1(t)$ is the sway velocity along z_1

$V_2(t)$ is the sway velocity along y_1

and g is the value of gravity (m/sec^2).

The horizontal accelerations caused by missile sway are seen to be included in the system measurements. While the missile sway causes spurious accelerations (noise) to be measured, the sway is limited to finite values whose average acceleration is zero. This allows the noise to be modeled as a zero mean process. In certain practical situations (ships, airplanes) in which the net average motion is not zero, additional inputs are required to describe the motion.

2.3.2 *Equations Governing the Angles α, β, γ Describing the Alignment Between the Earth-Fixed Launch-Site Coordinate System and the Strapdown Reference Coordinate System*

The three coordinate systems of concern in this analysis are

CS(0) - Earth-fixed launch-site CS

CS(1) - Strapdown reference CS

CS(2) - Strapdown instrument CS.

The transformation equation solved by the strapdown computing system relates CS(2) to CS(1). This equation follows from Equation (6),

$$\frac{d\,{}^1C_2}{dt} = {}^1C_2\,\omega_2 - \omega_0\,{}^1C_2 , \tag{9}$$

where ω_2 is the matrix of angular rates measured by the gyros.

ω_0 is the matrix of earth's rate in CS(0).

In Equation (9), the value of earth's rate in CS(0) is used rather than the proper values corresponding to the rate in CS(1), since there is as yet no information relating the alignment of CS(1) to CS(0).

The gyro measured rates are in error by the gyro drifts. Thus

$$\omega_2 = \dot{\theta} + d_2 , \tag{10}$$

where $\dot{\theta}$ represents the true angular rates of the swaying body and d_2 represents the gyro drift rates in the instrument axes.

The inverse equation to Equation (9) is

$$\frac{d\ {}^2C_1}{dt} = {}^2C_1\,\omega_0 - \omega_2\ {}^2C_1 \,. \tag{11}$$

Subsituting for the measured rate ω_2 from Equation (10),

$$\frac{d\ {}^2C_1}{dt} = {}^2C_1\,\omega_0 - \dot{\theta}\ {}^2C_1 - d_2\ {}^2C_1 \,. \tag{12}$$

In the same manner as for Equation (9),

$$\frac{d\ {}^0C_2}{dt} = {}^0C_2\,\dot{\theta} - \omega_0\ {}^0C_2 \tag{13}$$

and, with

$${}^0C_1 = {}^0C_2\ {}^2C_1 \,,$$

we find, using Equations (12) and (13) that

$$\frac{d\ {}^0C_1}{dt} = {}^0C_1\,\omega_0 - \omega_0\ {}^0C_1 - {}^0C_2\ d_2\ {}^2C_1 \,. \tag{14}$$

This is the basic equation which relates the strapdown reference CS(1) to the earth-fixed system CS(0).

The equation is simplified by first noting that the drift term can be written as

$${}^0C_2\ d_2\ {}^2C_1 = {}^0C_1\ {}^1C_2\ d_2\ {}^2C_1 = {}^0C_1\ d_1 \,. \tag{15}$$

Components of the drift will be assumed constant in CS(1) rather than CS(2).

Since CS(1) differs from CS(0) by small angles, as shown in Figure 1, ${}^0C_1$ can be written as

$${}^0C_1 = \begin{bmatrix} 1 & -\gamma & \beta \\ \gamma & 1 & -\alpha \\ -\beta & \alpha & 1 \end{bmatrix} . \tag{16}$$

Substituting Equation (16) in Equation (14) then causes Equation (14) to simplify to

$$\begin{bmatrix} \frac{d\,\alpha}{dt} \\ \frac{d\,\beta}{dt} \\ \frac{d\,\gamma}{dt} \end{bmatrix} = - \begin{bmatrix} 0 & -\omega_{z0} & \omega_{y0} \\ \omega_{z0} & 0 & -\omega_{x0} \\ -\omega_{y0} & \omega_{x0} & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} - \begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix} . \tag{17}$$

The equation for the full matrix $d\ {}^0C_1/dt$ reduces to only three equations, since the angles α, β, γ are small. The ω_0 matrix contains the values of earth's rate in CS(0), and the unknown drifts appear in the d matrix. Filtering will produce estimates for both the misalignment angles and the drifts.

2.3.3 *Equations Describing the Noise Inputs to the System Velocity Measurements*

The velocity equations (7) and (8) indicate the presence of noise terms. This noise is predominantly caused by missile sway. The noise cannot be represented as white, but rather a correlated velocity noise model is a better representation, due to the dynamics of the missile sway. The correlation function to be used is

$$R_{v_1}(\Delta) = \sigma_v^2\, e^{-|\Delta|/\tau} \cos \omega_n \Delta \,. \tag{18}$$

To generate this random variable from white noise inputs, let*

$$v_1(t) = n_1(t) \cos \omega_n t + n_2(t) \sin \omega_n t \,. \tag{19}$$

---

* *Probability, Random Variables, and Stochastic Processes* by A. Papoulis, McGraw-Hill, 1965, p. 373

where $n_1(t)$ and $n_2(t)$ are independent random variables with the autocorrelation function

$$R_{n_1}(\Delta) = R_{n_2}(\Delta) = \sigma_v^2 e^{-|\Delta|/\tau} , \tag{20}$$

Noise with this autocorrelation function can be considered as being the solution to the stochastic differential equations

$$\dot{n}_1(t) = (-1/\tau)n_1(t) + w_1(t) \tag{21}$$

$$\dot{n}_2(t) = (-1/\tau)n_2(t) + w_2(t) , \tag{22}$$

where $w_1(t)$ and $w_2(t)$ are white noise random variables with autocorrelation function

$$E[w_1(t_1)w_1(t_2)] = E[w_2(t_1)w_2(t_2)] = \frac{2\sigma_v^2}{\tau}\delta(t_1 - t_2) ; \tag{23}$$

$\delta(t_1 - t_2)$ is the Dirac delta function.

$v_1(t)$ is the sway velocity along the x axis. For the sway velocity along the y axis, $v_2(t)$, we have similar equations:

$$v_2(t) = n_2(t)\cos\omega_h t + n_4(t)\sin\omega_h t \tag{24}$$

$$\dot{n}_3(t) = (-1/\tau)n_3(t) + w_3(t) \tag{25}$$

$$\dot{n}_4(t) = (-1/\tau)n_4(t) + w_4(t) , \tag{26}$$

where

$$E[w_3(t_1)w_3(t_2)] = E[w_4(t_1)w_4(t_2)] = \frac{2\sigma_v^2}{\tau}\delta(t_1 - t_2) . \tag{27}$$

2.3.4 *Conversion of the System Equations to the Kalman State Equation Form*

The system equations are described by Equations (7), (8), (17), (19), (21), (22), (24), (25) and (26). These will be changed to a difference or state equation form directly recognizable as the common form for application of the Kalman filter equations. The process is fairly direct though tedious and intermediate results will not be shown*. All the differential equations in the set are linear; they are of the form

$$\frac{dx}{dt} = Fx(t) + \omega(t) \tag{28}$$

and have the general solution

$$x(t_i) = \phi(t_i, t_{i-1})x(t_{i-1}) + \int_{t_{i-1}}^{t_i} \phi(t_i, \tau)\omega(\tau)\, d\tau \tag{30}$$

which is of the proper state equation form. In Equation (30) the state transition matrix is the matrix exponential

$$\phi(t_i, t_{i-1}) = e^{F(t_i - t_{i-1})} . \tag{31}$$

The full state equation is finally obtained in the following form:

$$\begin{bmatrix} \alpha(t_i) \\ \beta(t_i) \\ \gamma(t_i) \\ d_x(t_i) \\ d_y(t_i) \\ d_m(t_i) \\ p_1(t_i) \\ p_2(t_i) \\ n_1(t_i) \\ n_2(t_i) \\ n_3(t_i) \\ n_4(t_i) \end{bmatrix} = \begin{bmatrix} & & & & & & & & & & & \\ & & & & & & & & & & & \\ \phi_{11} & \phi_{12} & 0 & 0 & & & & & & & & \\ & & & & & & & & & & & \\ 0 & I & 0 & 0 & & & & & & & & \\ & & & & & & & & & & & \\ \phi_{31} & \phi_{32} & I & \phi_{34} & & & & & & & & \\ & & & & & & & & & & & \\ 0 & 0 & 0 & \phi_{44} & & & & & & & & \end{bmatrix} \begin{bmatrix} \alpha(t_{i-1}) \\ \beta(t_{i-1}) \\ \gamma(t_{i-1}) \\ d_x(t_{i-1}) \\ d_y(t_{i-1}) \\ d_m(t_{i-1}) \\ p_1(t_{i-1}) \\ p_2(t_{i-1}) \\ n_1(t_{i-1}) \\ n_2(t_{i-1}) \\ n_3(t_{i-1}) \\ n_4(t_{i-1}) \end{bmatrix} + \begin{bmatrix} 0 & & & \\ & & & \\ 0 & & & \\ & & & \\ C(t) & & & \\ & & & \\ I & & & \end{bmatrix} \begin{bmatrix} u_1(t_i) \\ u_2(t_i) \\ u_3(t_i) \\ u_4(t_i) \end{bmatrix} \tag{32a}$$

* For a detailed derivation see Reference 2.

or
$$x(t_i) = \phi(t_i, t_{i-1})x(t_{i-1}) + G(t_i)u(t_i) \ , \tag{32b}$$

where
$$\phi_{11} = I - \frac{\sin \omega(t_i - t_{i-1})}{\omega} \omega_0 + \frac{1 - \cos \omega(t_i - t_{i-1})}{\omega^2} \omega_0^2 \tag{33}$$

$$\omega = (\omega_{x_0}^2 + \omega_{y_0}^2 + \omega_{z_0}^2)^{1/2}$$

ω_0 is the ω matrix of Equation (17)

I is the 3×3 unit matrix,

$$\phi_{12} = I(t_i - t_{i-1}) + \frac{1 - \cos \omega(t_i - t_{i-1})}{\omega^2} \omega_0 - \left[\frac{t_i - t_{i-1}}{\omega^2} - \frac{\sin \omega(t_i - t_{i-1})}{\omega^3}\right] \omega_0^2 \tag{34}$$

$$\phi_{21} = -\begin{bmatrix} 0 & g & 0 \\ 0 & 0 & -g \end{bmatrix} \phi_{12} \tag{35}$$

$$\phi_{22} = \begin{bmatrix} 0 & g & 0 \\ 0 & 0 & -g \end{bmatrix} \left\{ -\frac{(t_i - t_{i-1})^2}{2} I + \left(\frac{t_i - t_{i-1}}{\omega^2} - \frac{\sin(t_i - t_{i-1})}{\omega^3}\right)\omega_0 - \right. \tag{36}$$
$$\left. - \left(\frac{(t_i - t_{i-1})^2}{2\omega^2} - \frac{1 - \cos \omega(t_i - t_{i-1})}{\omega^4}\right)\omega_0^2 \right\}$$

$$\phi_{24} = G(t_i)\, \phi_{44} - G(t_{i-1}) \tag{37}$$

$$\phi_{44} = e^{-(t_i - t_{i-1})/\tau} I \tag{38}$$

I is the 4×4 unit matrix

$$G(t) = \begin{bmatrix} \cos \omega_h t & \sin \omega_h t & 0 & 0 \\ 0 & 0 & \cos \omega_h t & \sin \omega_h t \end{bmatrix} . \tag{39}$$

The noise terms which appear in the state equations are related to the white noise of Equations (21) to (28) by

$$\begin{bmatrix} u_1(t_i) \\ u_2(t_i) \\ u_3(t_i) \\ u_4(t_i) \end{bmatrix} = \int_{t_{i-1}}^{t_i} \phi_{44}(t_i, \tau) \begin{bmatrix} w_1(\tau) \\ w_2(\tau) \\ w_3(\tau) \\ w_4(\tau) \end{bmatrix} d\tau$$

and have the variance

$$Q(t_i) = E(uu') = \sigma_q^2 (1 - e^{-2(t_i - t_{i-1})/\tau}) I \ . \tag{40}$$

The data measurement consists of the two transformed accelerometer outputs. There is additional noise of the measurement due to quantization of the accelerometer outputs. The measurement equation is given by

$$y(t_i) = H(t_i)X(t_i) + r(t_i) \ , \tag{41}$$

where
$$H(t_i) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} . \tag{42}$$

$r(t_i)$ is the noise due to quantization. Its covariance matrix is given by

$$R(t_i) = E[r(t_i)r'(t_i)] = \frac{\Delta^2}{12} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} . \tag{43}$$

where Δ is the quantization level.

When a theodolite measurement is taken and processed, the resulting quantity is the azimuth misalignment plus an error which is equal to the leveling error about the projection of the LOS in the horizontal plane multiplied by the tangent of the elevation angle. Thus, when a theodolite measurement is taken, the $H(t_i)$ matrix is

$$H(t_i) = [1, \ \tan \gamma_0 \sin \zeta, \ \tan \gamma_0 \cos \zeta, 0, 0, 0, 0, 0, 0, 0, 0, 0] \ , \tag{44}$$

where ζ is the angle between the projection of the LOS in the horizontal plane and the z_1 axis.

3. OPTIMAL FILTER EQUATIONS

In Section 2, the system equations were derived. They are repeated here for convenience.

$$x(t_i) = \phi(t_i,t_{i-1})x(t_{i-1}) + G(t_i)\,u(t_i) \tag{45}$$

$$y(t_i) = H(t_i)\,x(t_i) + r(t_i)\ . \tag{46}$$

The well-known Kalman filter Equations can be applied to systems with this description. These equations give an estimate of the state vector $x(t_i)$, based on the measurements $\{y(t_1) \ldots y(t_i)\}$, which is optimal in the mean-square sense. The filter is described by the following difference equations:

$$\hat{x}(t_i,t_{i-1}) = \phi(t_i,t_{i-1})\hat{x}(t_{i-1}) \tag{47}$$

$$\hat{x}(t_i) = \hat{x}(t_i,t_{i-1}) + K(t_i)[y(t_i) - H(t_i)\hat{x}(t_i,t_{i-1})]\ . \tag{48}$$

The optimal filter gains $K(t_i)$ are calculated from the following set of difference equations:

$$P(t_i,t_{i-1}) = \phi(t_i,t_{i-1})P(t_{i-1})\phi(t_i,t_{i-1}) + G(t_i)Q(t_i)G'(t_i) \tag{49}$$

$$K(t_i) = P(t_i,t_{i-1})H'(t_i)[H(t_i)P(t_i,t_{i-1})H'(t_i) + R(t_i)]^{-1} \tag{50}$$

$$P(t_i) = [I - K(t_i)H(t_i)]\,P(t_i,t_{i-1})\ . \tag{51}$$

where $\hat{x}(t_i)$ is the optimal estimate of $x(t_i)$, given the measurements $y(t_1)$, $y(t_2)$, ... $y(t_i)$

$\hat{x}(t_i,t_{i-1})$ is the optimal estimate of $x(t_i)$, given the measurements $\{y(t_1),y(t_2),\ldots y(t_{i-1})\}$

$P(t_i)$ is the covariance matrix of the error in the estimate $\hat{x}(t_i)$.

$P(t_i,t_{i-1})$ is the covariance matrix of the error in the estimate $\hat{x}(t_i,t_{i-1})$.

These equations describe the optimal filter. It is desirable, however, to simplify and decouple the equations to reduce the computational requirements. The performance of the optimal filter is described in Section 5. It is also demonstrated that the degradation caused by certain simplifications is negligible.

These simplifications will be investigated in the following section.

4. SUBOPTIMAL FILTER EQUATIONS

4.1 Introduction

The filter equations presented in the previous section are rather complex. They involve 12 simultaneous difference equations and hence multiplication of 12×12 matrices. It is desirable, then, to investigate the possibility of simplifying these equations. In order to determine if the simplifications are acceptable, one must calculate the degradation in filter performance caused by them. In addition, the sensitivity of the filter to errors in the noise should be investigated. The equations required to evaluate the suboptimal filter performance are derived in this section. The approximations to the system dynamics which are used in the suboptimal filter are also presented.

The suboptimal filter is obtained by simplifying the state equation such that the state transition matrix $\phi(t_i,t_{i-1})$ of Equation (45) is replaced by an approximate and simpler form $\phi_s(t_i,t_{i-1})$.

A fairly general model for the suboptimal filter equations is then given by

$$\hat{x}_s(t_i,t_{i-1}) = \phi_s(t_i,t_{i-1})\hat{x}_s(t_{i-1}) \tag{52}$$

$$\hat{x}_s(t_i) = \hat{x}_s(t_i,t_{i-1}) + K_s(t_i)[y(t_i) - H(t_i)\hat{x}_s(t_i,t_{i-1})]\ . \tag{53}$$

The subscript s denotes suboptimal quantities. The gain $K_s(t_i)$ is obtained from

$$P_s(t_i,t_{i-1}) = \phi_s(t_i,t_{i-1})P_s(t_{i-1})\phi_s'(t_i,t_{i-1}) + G_s(t_i)Q_s(t_i)G'(t_i) \tag{54}$$

$$K_s(t_i) = P_s(t_i,t_{i-1})H'(t_i)[H(t_i)P_s(t_i,t_{i-1})H'(t_i) + R(t_i)]^{-1} \tag{55}$$

$$P_s(t_i) = [I - K_s(t_i)H(t_i)]\,P_s(t_i,t_{i-1})\ . \tag{56}$$

These equations are identical with the usual Kalman set except that they involve suboptimal quantities. To determine the true covariance matrix of the estimate errors, we first define the quantities

$$P_1(t_i) = E[x(t_i)x'(t_i)]$$

$$P_2(t_i,t_{i-1}) = E[\hat{x}_s(t_i,t_{i-1}),x'(t_i)]$$

$$P_3(t_i) = E[\hat{x}_s(t_i)x'(t_i)]$$

$$P_4(t_i,t_{i-1}) = E[\hat{x}_s(t_i,t_{i-1})\hat{x}_s(t_i,t_{i-1})]$$

$$P_5(t_i) = E[\hat{x}_s(t_i)\hat{x}_s'(t_i)] .$$

The covariance matrix of estimate errors for the suboptimal filter is given by

$$\Gamma(t_i) = E\{[x(t_i) - \hat{x}_s(t_i)][x(t_i) - \hat{x}_s(t_i)]'\} . \tag{57}$$

Equation (57) can be written, in terms of the covariance matrices defined above,

$$\Gamma(t_i) = P_1(t_i) - P_2'(t_i) - P_3(t_i) + P_5(t_i) . \tag{58}$$

By multiplying Equation (45) by its transpose and taking expected values of both sides, one immediately obtains

$$P_1(t_i) = \phi(t_i,t_{i-1})P_1(t_{i-1})\phi'(t_i,t_{i-1}) + G(t_i)Q(t_i)G'(t_i) . \tag{59}$$

Multiplying Equation (52) by the transpose of Equation (45) and taking expected values yields

$$P_2(t_i,t_{i-1}) = \phi_s(t_i,t_{i-1})P_1(t_{i-1})\phi'(t_i,t_{i-1}) . \tag{60}$$

The equation for $P_3(t_i)$ is obtained by multiplying Equation (53) by $x'(t_i)$. It is given by

$$P_3(t_i) = [I - K_s H(t_i)]P_2(t_i,t_{i-1}) + K_s H(t_i)P_1(t_i) . \tag{61}$$

Multiplying Equation (52) by its transpose and taking expected values gives

$$P_4(t_i,t_{i-1}) = \phi_s(t_i,t_{i-1})P_4(t_{i-1})\phi_s'(t_i,t_{i-1}) . \tag{62}$$

The equation for $P_5(t_i)$ is obtained by multiplying Equation (53) by its transpose and taking expected values. After some algebraic manipulation, one obtains

$$\begin{aligned} P_5(t_i) = {} & P_5(t_i,t_{i-1}) + K_s(t_i)H(t_i)[P_5'(t_i) - P_5(t_i,t_{i-1})] + \\ & + [P_3(t_i) - P_3(t_i,t_{i-1})]H'(t_i)K_s'(t_i) + \\ & + K_s(t_i)H(t_i)[P_3(t_i,t_{i-1}) - P_1(t_i)]H'(t_i)K_s'(t_i) + \\ & + K_s(t_i)R(t_i)K_s'(t_i) . \end{aligned} \tag{63}$$

This completes the derivation of the equations necessary to compute the covariance matrix of estimation errors.

4.2 Approximations to System Dynamics

If earth's rotation rate is neglected ($\omega = 0$), the system dynamics are considerably simplified.

The state transition matrix becomes

$$\phi(t_i,t_{i-1}) = \begin{bmatrix} I & \Delta(t_i - t_{i-1})I & 0 & 0 \\ 0 & I & 0 & 0 \\ B(t_i - t_{i-1}) & \dfrac{-B(t_i - t_{i-1})^2}{2} & I & \phi_{34}(t_i,t_{i-1}) \\ 0 & 0 & 0 & \phi_{44}(t_i,t_{i-1}) \end{bmatrix} \tag{64}$$

where $B = \begin{bmatrix} 0 & g & 0 \\ 0 & 0 & -g \end{bmatrix}$.

The result of this simplification is that the equations for α, β, and γ become uncoupled. If the drifts are neglected, a further simplification results. The equations describing β are given by

$$\begin{bmatrix}\beta(t_i)\\ p_1(t_i)\\ n_1(t_i)\\ n_2(t_i)\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0\\ g(t_i-t_{i-1}) & 1 & \phi_{33} & \phi_{34}\\ 0 & 0 & \phi_{33} & 0\\ 0 & 0 & 0 & \phi_{44}\end{bmatrix}\begin{bmatrix}\beta(t_{i-1})\\ p(t_{i-1})\\ n_1(t_{i-1})\\ n_2(t_{i-1})\end{bmatrix} + \begin{bmatrix}0 & 0\\ G_{31} & G_{32}\\ 1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}u_1(t_i)\\ u_2(t_i)\end{bmatrix}, \tag{65}$$

where $\phi_{33} = \phi_{44} = \exp[-(t_i - t_{i-1})/\tau]$

$\phi_{31} = \phi_{33}\cos\omega_n t_i - \cos\omega_n t_{i-1}$

$\phi_{34} = \phi_{33}\sin\omega_n t_i - \sin\omega_n t_{i-1}$

$G_{31} = \cos\omega_n t_i$

$G_{32} = \sin\omega_n t_i$.

The equations describing $\gamma(t_i)$ are of almost the identical form with a sign difference.

The equation for α is simply

$$\alpha(t_i) = \alpha(t_{i-1}) . \tag{66}$$

The filter corresponding to these uncoupled dynamics is also uncoupled. A summary of the suboptimal four-state filter equations follows.

$$\hat{x}_1(t_i,t_{i-1}) = \phi^*(t_i,t_{i-1})\hat{x}_1(t_{i-1}) \tag{67}$$

$$\hat{x}_2(t_i,t_{i-1}) = \phi^*(t_i,t_{i-1})\hat{x}_2(t_{i-1}) \tag{68}$$

$$\hat{x}_1(t_i) = \hat{x}_1(t_i,t_{i-1}) + K^*(t_i)[p_1(t_i) - H^*(t_i)\hat{x}_1(t_i,t_{i-1})] \tag{69}$$

$$\hat{x}_2(t_i) = \hat{x}_2(t_i,t_{i-1}) + K^*(t_i)[-p_2(t_i) - H^*(t_i)\hat{x}_2(t_i,t_{i-1})] \tag{70}$$

$$P^*(t_i,t_{i-1}) = \phi^*(t_i,t_{i-1})P^*(t_{i-1})\phi^{*\prime}(t_i,t_{i-1}) + G^*(t_i)Q^*(t_i)G^{*\prime}(t_i) \tag{71}$$

$$K^*(t_i) = P^*(t_i,t_{i-1})H^{*\prime}(t_i)[H^*(t_i)P^*(t_i,t_{i-1})H^{*\prime}(t_i) + R^*(t_i)]^{-1} \tag{72}$$

$$P^*(t_i) = [I - K^*(t_i)H^*(t_i)]P^*(t_i,t_{i-1}) , \tag{73}$$

where

$$\phi^*(t_i,t_{i-1}) = \begin{bmatrix}0 & 0 & 0 & 0\\ g(t_i-t_{i-1}) & 1 & \phi_{33} & \phi_{34}\\ 0 & 0 & \phi_{33} & 0\\ 0 & 0 & 0 & \phi_{44}\end{bmatrix}$$

$$G^*(t_i) = \begin{bmatrix}0 & 0\\ G_{31} & G_{32}\\ 1 & 0\\ 0 & 1\end{bmatrix}, \qquad Q^*(t_i) = \sigma_q^2[1 - e^{-(t_i-t_{i-1})/\tau}]\,I$$

$$\hat{x}^1(t_i) = \begin{bmatrix}\hat{\beta}(t_i)\\ \hat{\beta}_1(t_i)\\ \hat{n}_1(t_i)\\ \hat{n}_2(t_i)\end{bmatrix}, \qquad x_2(t_i) = \begin{bmatrix}\hat{\gamma}(t_i)\\ -\hat{\beta}_2(t_i)\\ \hat{n}_3(t_i)\\ \hat{n}_4(t_i)\end{bmatrix}$$

$$H^*(t_i) = [0 \quad 1 \quad 0 \quad 0], \qquad R^*(t_i) = \Delta^2/12 .$$

4.3 Bias Errors in the Suboptimal Filter

The approximations to the system dynamics can cause a bias to appear in the suboptimal estimate, and no direct method has yet appeared to determine this bias as readily as the determination of the covariance estimate of the suboptimal errors given by Equation (58). In general, a bias will occur if the non-noise associated part of the dynamic equations is approximated, or if the measurement equations are changed.

In the simplifications discussed in Section 4.2, the variation in time of the angles α, β, γ caused by the earth's cross-coupling is neglected. This causes a bias to occur in the estimates made by the simplified filter.

An approximate expression for the bias can be calculated by neglecting sway noise. In this case, the angle is varying essentially as a linear function of time due to earth's rate cross-couplings. However, the filter neglects time variation of the angle and assumes the angle to be a constant. It can be shown that the error in the estimate which minimizes the mean-square error is equal to one-half the change in angle over the measurement time. This fact can be used to write expressions for the measured angles in terms of earth's rate, the geometry of the optical measurement of azimuth and the actual angles. Then equations can be solved to yield a better estimate of the actual angles. Details are shown in Reference 2.

5. RESULTS SUMMARY

Figures 2 through 9 show the results of the analysis of the Kalman filter. These results are obtained from computer simulation of the appropriate equations. Figure 2 is the 1σ erection error for the optimal 12-state filter. This filter is described by Equations (47)-(51). The conditions used for this baseline case are given by

Initial integrated velocity uncertainty	$\sigma_p = 0$
Initial erection uncertainty	$\sigma_\theta = 1/2$ degree
Gyro drift uncertainty	$\sigma_d = 0.1$ meru (1 meru = 0.015^0/hr)
R.M.S. sway velocity	$\sigma_v = 0.5$ m/s
Accelerometer quantization level	$\Delta = 0.02$ m/s
Center frequency of sway velocity power spectrum	$f_n = 0.25$ Hz
Correlation time (reciprocal of bandwidth) of sway velocity power spectrum	$\tau = 200$ s
Sampling rate	1 sample/s .

Figure 3 shows the performance of the suboptimal 4-state filter for various initial erection angles. This filter is described by Equations (67)-(73). The main simplification employed in reducing the 12-state optimal filter to two identical 4-state filters was the neglect of earth's rate cross-coupling terms in the system dynamics used in the filter equations. In addition, gyro drifts were neglected in the simplified dynamics. The difference in errors for various initial angles is attributable to the earth's rate coupling. For initial angles less than 0.5 degrees the increase in error over the optimal filter is less than 1 second of arc at 60 seconds time. If the initial angle uncertainty is large, the 12-state optimal filter performance can be approached by updating the computer coordinate system after about 20 seconds to reduce the cross-coupling effects. Another approach to removing the cross-coupling terms is to use the correction matrix derived in Reference 2.

Figure 4 shows the degradation in performance due to the neglect of gyro drift. In Figure 4, the gyro drift numbers shown as parameters are the 1σ uncertainties in the true gyro drift. The filter equations assume zero gyro drift. It is seen that the drift uncertainty must be quite large (20 meru = 0.3 deg/hr) before the filter performance is degraded appreciably.

Figure 5 shows the erection error versus time for various sampling times. A sampling time of 1 second seems to provide a good trade-off between computational requirements and erection accuracy. The worst case occurs when the sampling time is 4 seconds. This is because the center frequency of the sway velocity power spectrum is 0.25 hertz. Since the power spectrum is very narrow, the sway velocity is almost sinusoidal, with a period of 4 seconds. The sampling is thus at the same part of the sine wave, and it becomes quite difficult to filter out the noise.

The correlation time (reciprocal of the bandwidth) of the narrow band noise has a significant effect on the estimation accuracy. This is demonstrated in Figure 6, where the correlation time assumed in the filter equations matches the correlation time of the actual noise. It was found that the filter was insensitive to errors in the assumed correlation time. When the assumed correlation time was 20 seconds rather than the true correlation time of 200 seconds, the degradation in filter performance was negligible. The important parameter is the actual noise correlation time and not the correlation time assumed for the filter calculations.

Since the frequency of the missile sway may not be known exactly, it is of interest to determine the sensitivity of the filter to that parameter. This is shown in Figure 7. In this case, the actual frequency is 0.25 hertz, while the number used in the filter calculations is 0.2 hertz. It is seen that the accuracy of the estimate is somewhat poorer, but still acceptable. The performance for an assumed frequency of 0.3 hertz is almost identical to the curve shown in Figure 7 and, therefore, has not been presented.

An alternative method of erecting the analytic coordinate system is to use a least-squares filter. The computational requirements of this method are much less than those of the optimal filtering technique. The estimate, of course, takes longer to converge to within acceptable limits. Figure 8 compares the performance of the least squares and the optimal 4-state filters of the case when $\tau = 200$ s. The optimal filter is significantly better than the least-squares filter for this case. Figures 9 and 10 show the same comparison for $\tau = 30$ s and 5 s respectively. It is seen that, as τ decreases, the least-squares filter performance approaches that of the optimal filter. This is to be expected since, for white noise ($\tau = 0$), the least-squares filter is optimal.

The preceding curves were calculated by determining the covariance matrices for the estimation errors. In addition, a simulation of the filter equations with input noise derived from a random number generator was performed. Figure 11 shows the results of one run of this simulation with the sway velocity noise spectrum lying in a narrow band around a center frequency. The $L\sigma$ value obtained from the covariance matrix calculation is shown for comparison. Figure 12 shows the results of a run in which the input noise is assumed to consist of three narrow band noises at different center frequencies. The filter still assumes that the noise is centered in a narrow band around a single frequency. The three-frequency noise is considered to be a more realistic representation of the velocity of a swaying missile with several bending modes on the launch pad. The center frequencies used were 1.82 rad/s, 2.14 rad/s, and 2.32 rad/s, with a correlation time of 200 seconds for each band. The total r.m.s. sway velocity was 0.5 m/s.

Certain "rules of thumb" which have been developed to allow extrapolation of the results given in this chapter to different noise characteristics and filtering times are as follows:

(a) The r.m.s. filtering error is proportional to the r.m.s. value of the sway velocity noise.

(b) If earth's rate cross-couplings are removed, by updating or by using the correction matrix, the filtering error is not a strong function of the initial angle error.

(c) The filtering error is approximately inversely proportional to the noise center frequency.

(d) For short correlation times, the filtering error for both the optimal 4-state and least-squares filter is approximately inversely proportional to the three-halves power of the filtering time (Error = $K/T^{3/2}$). This is shown by plotting the results of Figure 9 on log-log paper in Figure 13. For longer correlation times, the least-squares filter has a filtering error which is more nearly inversely proportional to the square of the filtering time, while the 4-state optimal remains inversely proportional to the three-halves power of the filtering time. This is shown by plotting the results of Figure 8 on log-log paper in Figure 14. It must be remembered, however, that the curves for the two filters can never intersect, so that the least-squares filter error must approach being inversely proportional to the three-halves power of filtering time as the time increases.

(e) The filtering error for the optimal 4-state filter is approximately inversely proportional to the square root of the sway velocity correlation time (Error = $K/\tau^{1/2}$). The least-squares filter error only approaches being inversely proportional to the square root of the correlation time at low correlation times, however. For higher correlation time the improvement in performance with increasing correlation time is much less than that of the optimal 4-state filter. This is shown by plotting the results of Figures 8, 9 and 10 at 60 seconds time on log-log paper in Figure 15. This rule obviously cannot be used as τ approaches zero, but it does hold at least down to $\tau = 1$ s.

ACKNOWLEDGEMENT

This chapter is based on work performed under Contract No. NAS-8-21273 to the NASA George C. Marshall Space Flight Center, July 1968. Principal investigators were Earle B. Crocker, Leonard Rabins, and Kevin A. Clements.

REFERENCES

1. Kennel, Hans F. — *Initial Alignment of a Strapdown Inertial Reference and Navigation System.* AIAA Paper 67-556. NASA Marshall Space Flight Center, August 1967.

2. – — *Final Report – Optimum Digital Filter Study for Strapdown Inertial Navigation System Initial Alignment.* Ordnance Systems, General Electric Company, Pittsfield, Massachusetts, July 8, 1968.

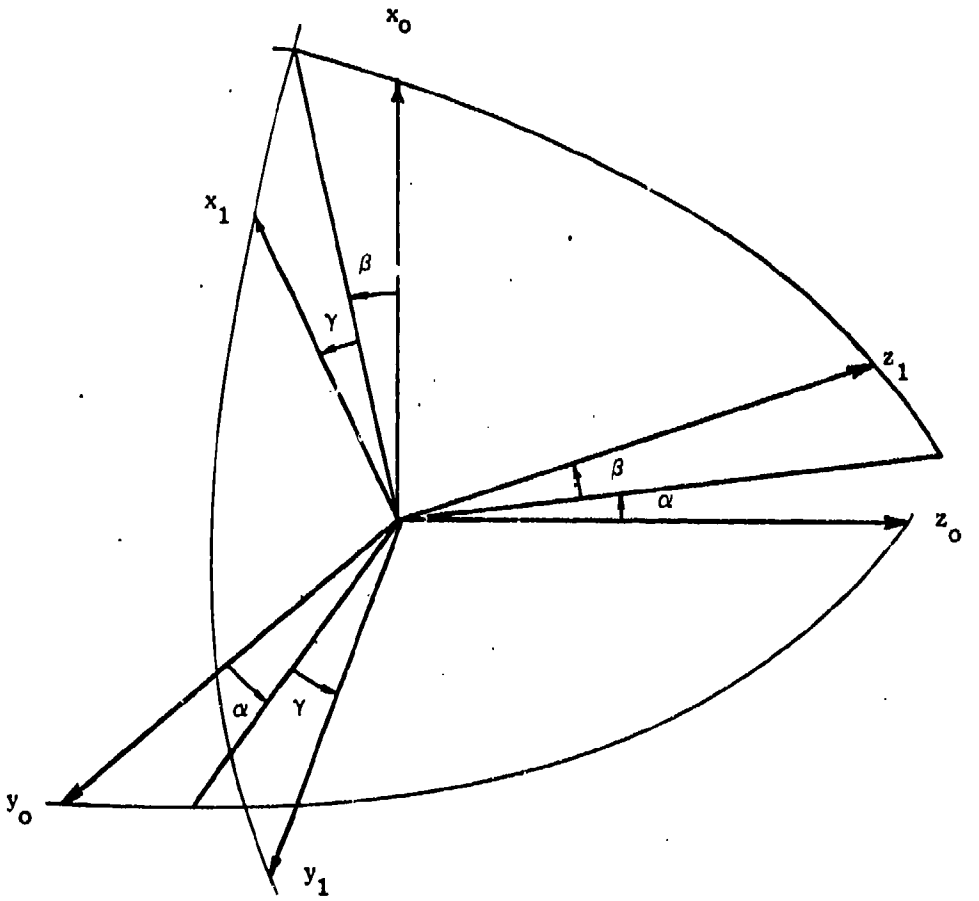


Figure 1

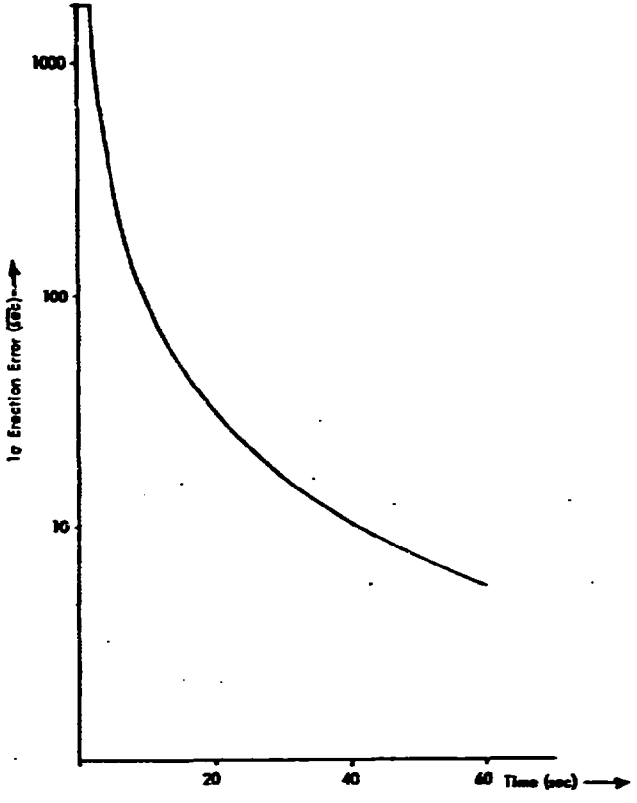


Fig. 2 The 1σ erection error for the optimal filter

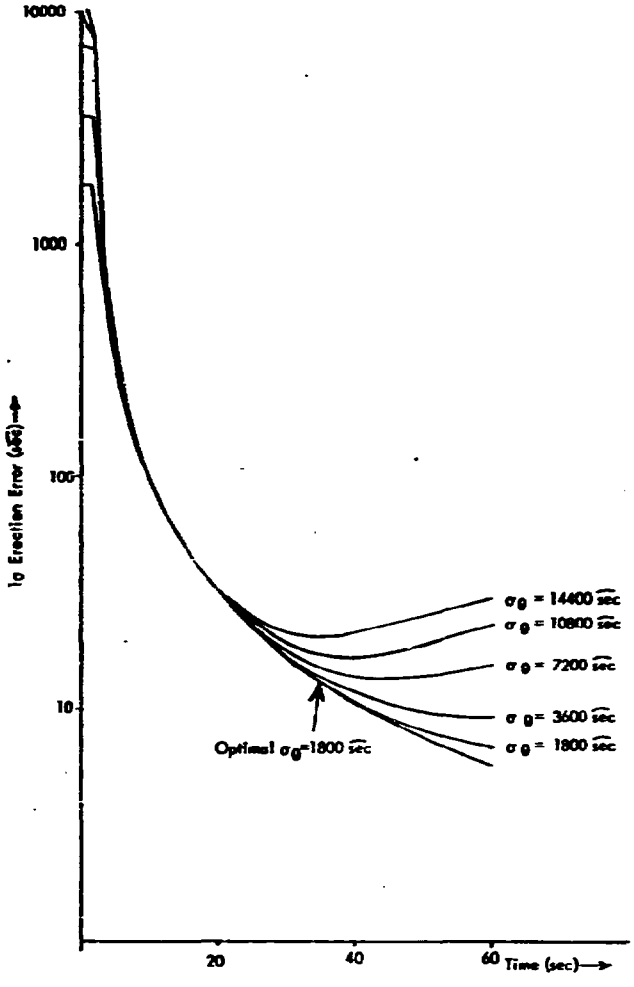


Fig. 3 Four-state filter erection error for various initial erection and alignment errors

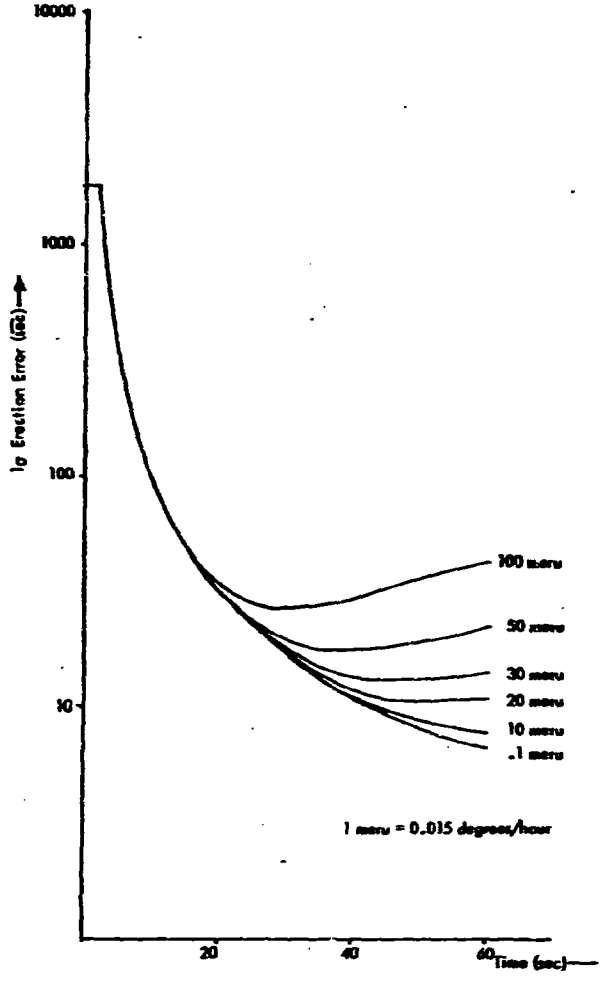


Fig.4 Four-state filter erection error for various gyro drifts

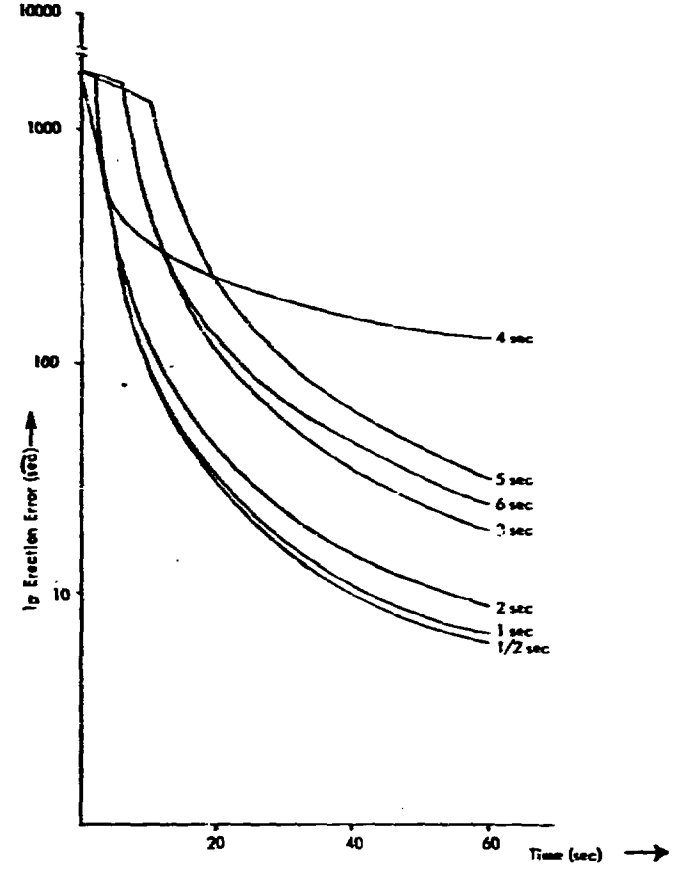


Fig.5 Erection error for various sampling times

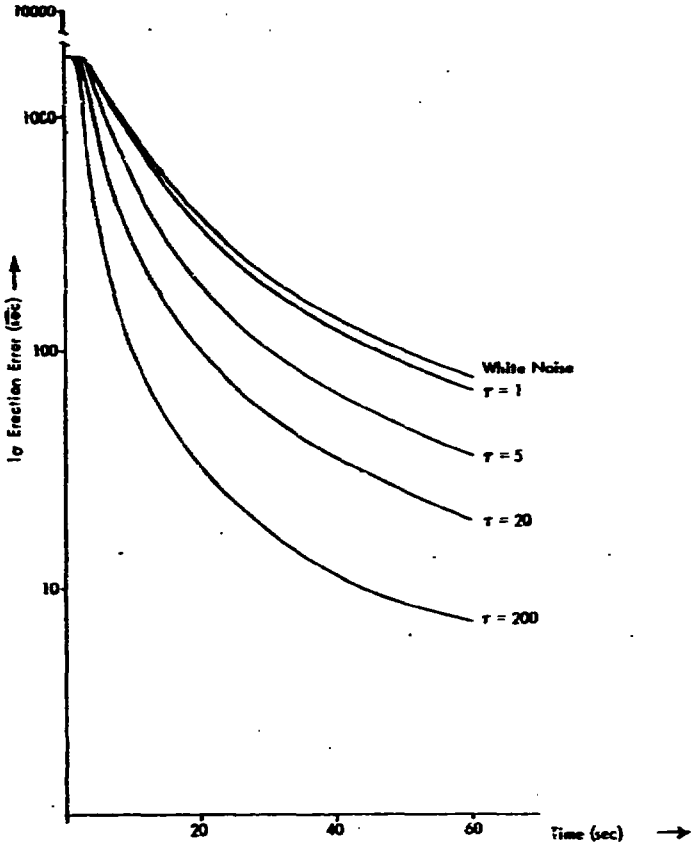


Fig. 6 Erection error for various sway velocity correlation times

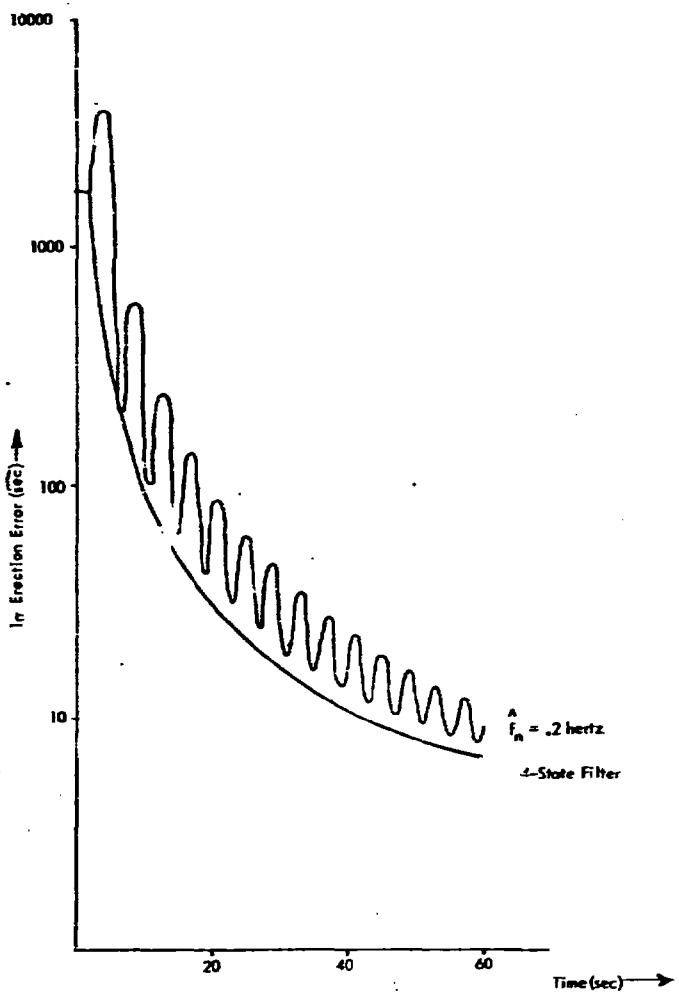


Fig. 7 Sensitivity of 4-state filter to errors in assumed center frequency of sway velocity

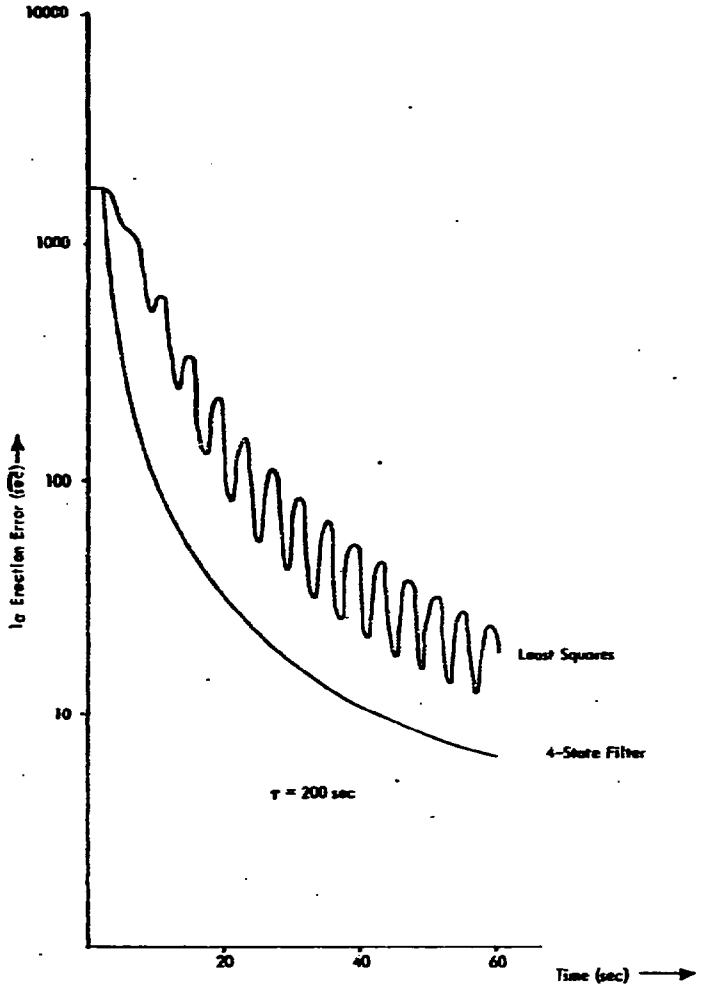


Fig. 8 A comparison of the 4-state and least-squares filters for a sway velocity correlation time of 200 seconds

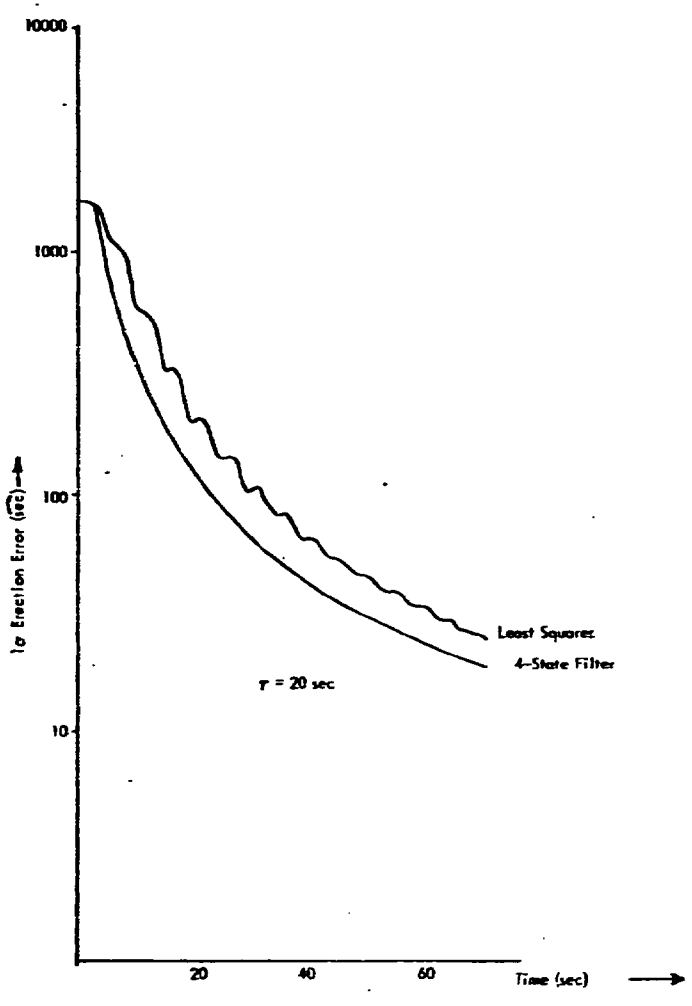


Fig. 9 A comparison of the 4-state and least-squares filters for a sway velocity correlation time of 20 seconds

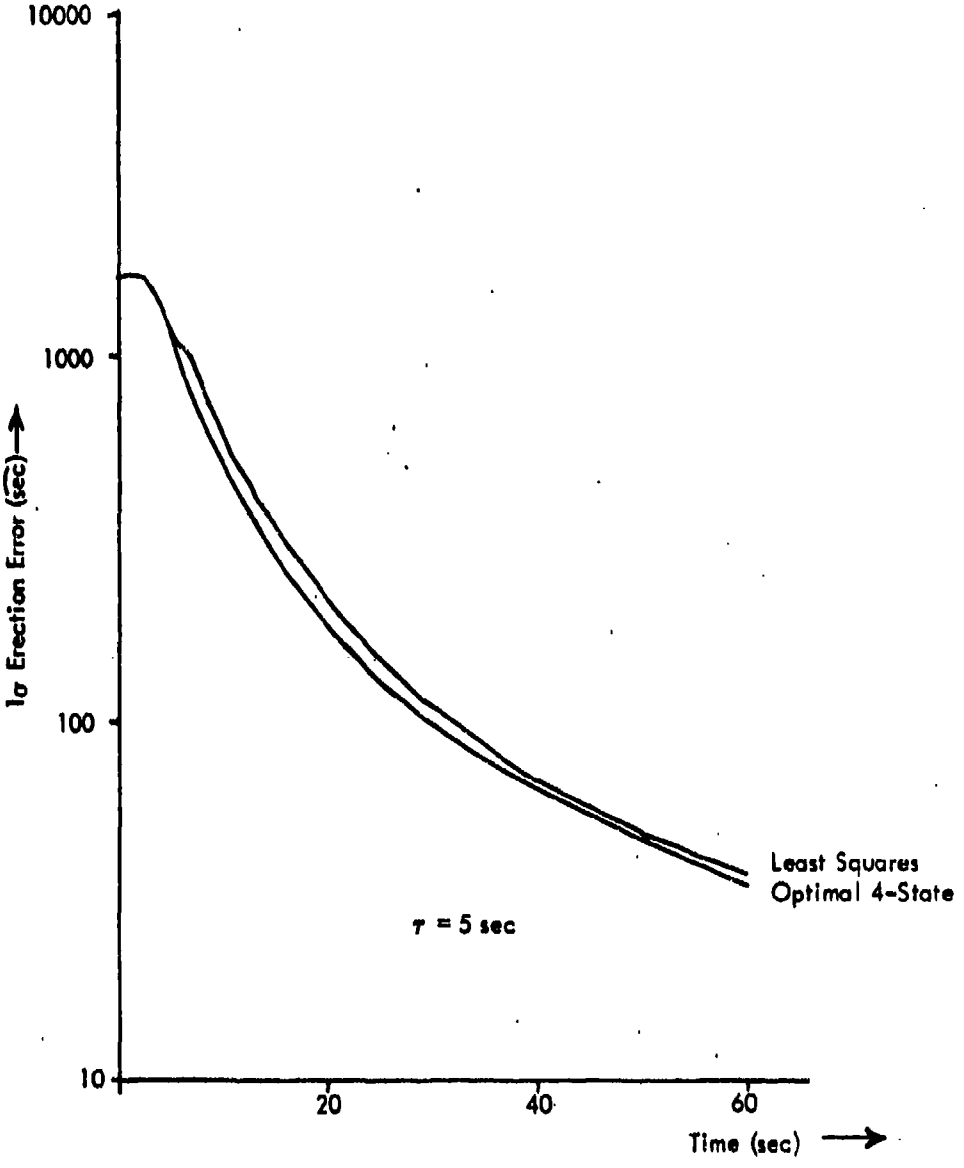


Fig. 10 A comparison of the 4-state and least-squares filters for a sway velocity correlation time of 5 seconds

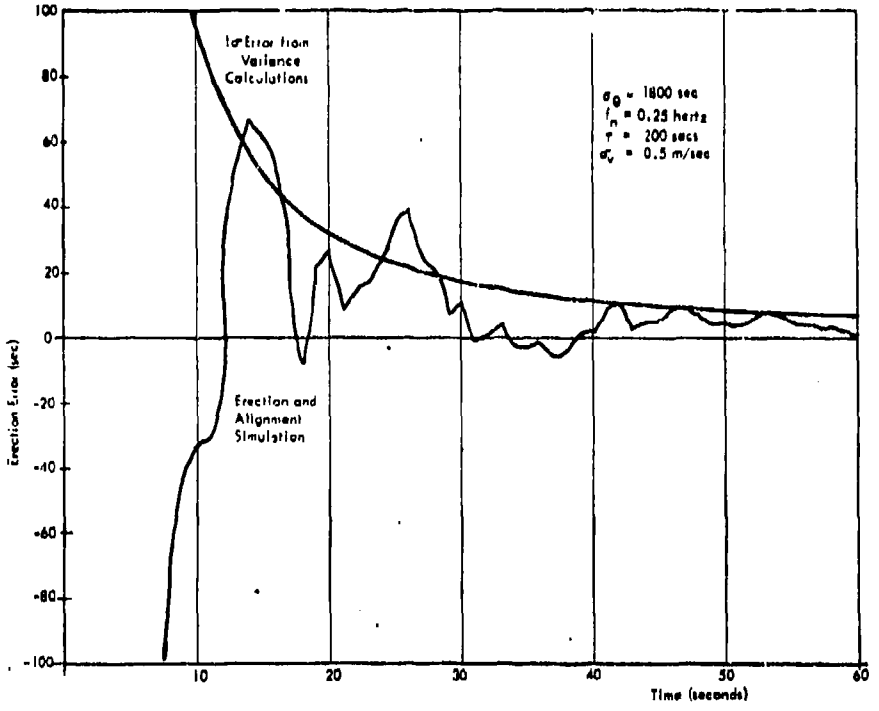


Fig. 11 Comparison of the results obtained from an estimation error variance calculation and those obtained from the erection and alignment simulation program (optimal 4-state filter)

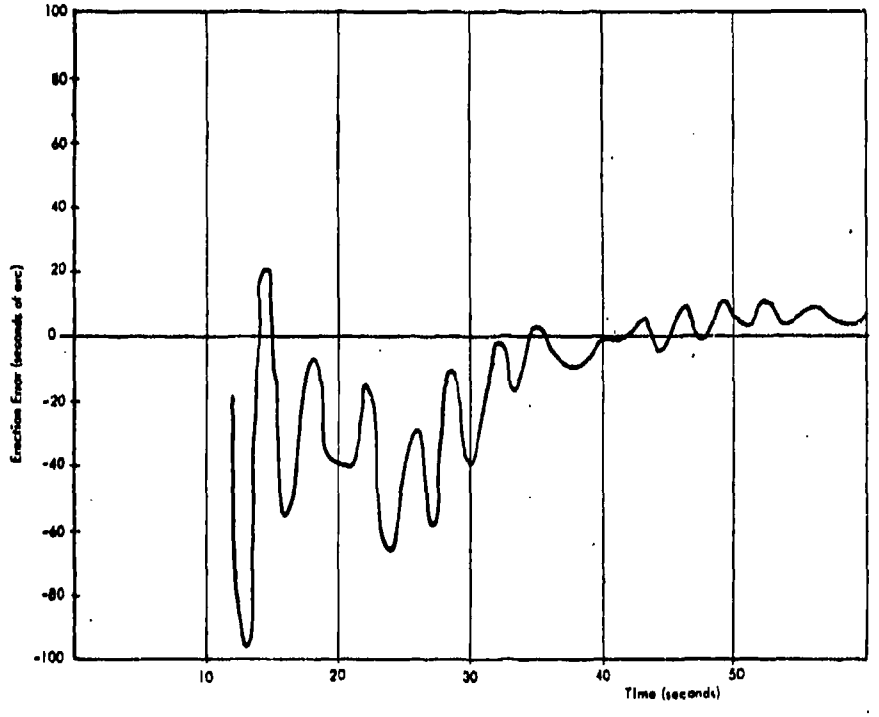


Fig. 12 Error when sway velocity noise has 3 narrow band components (4-state optimal filter)

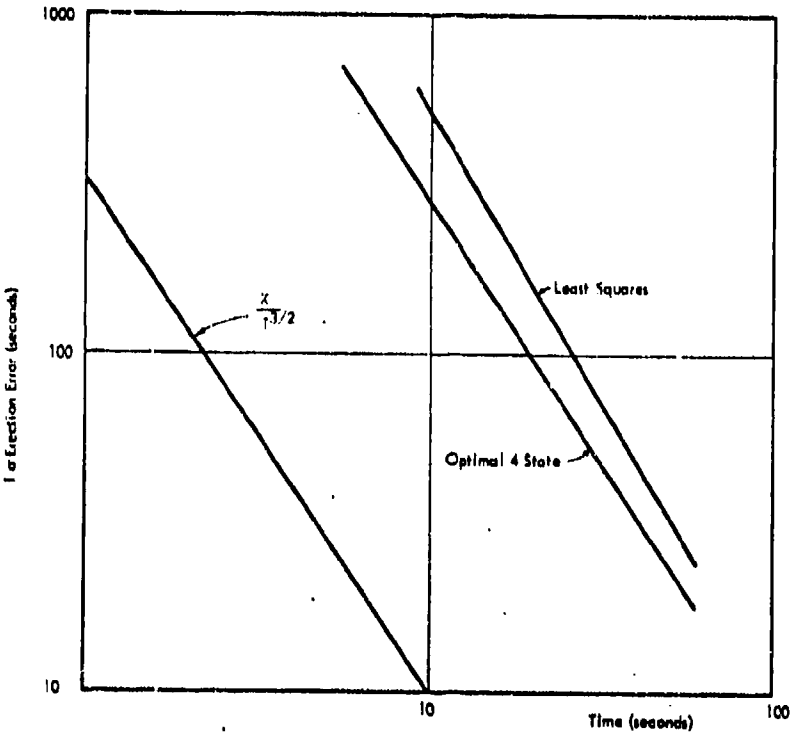


Fig. 13 Dependence of filtering error on filtering time (τ = 20 seconds)

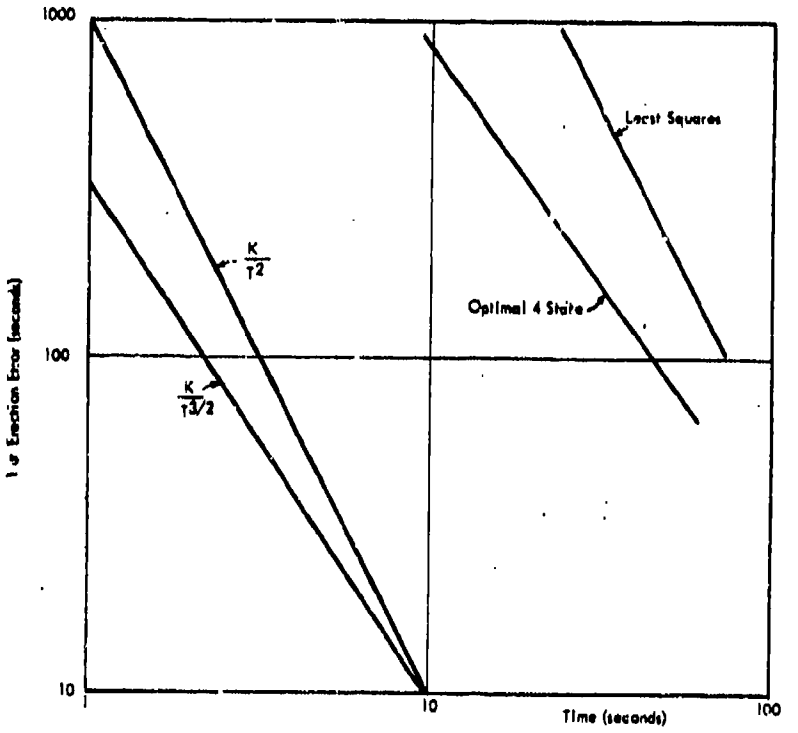


Fig. 14 Dependence of filtering error on filtering time (τ = 200 seconds)

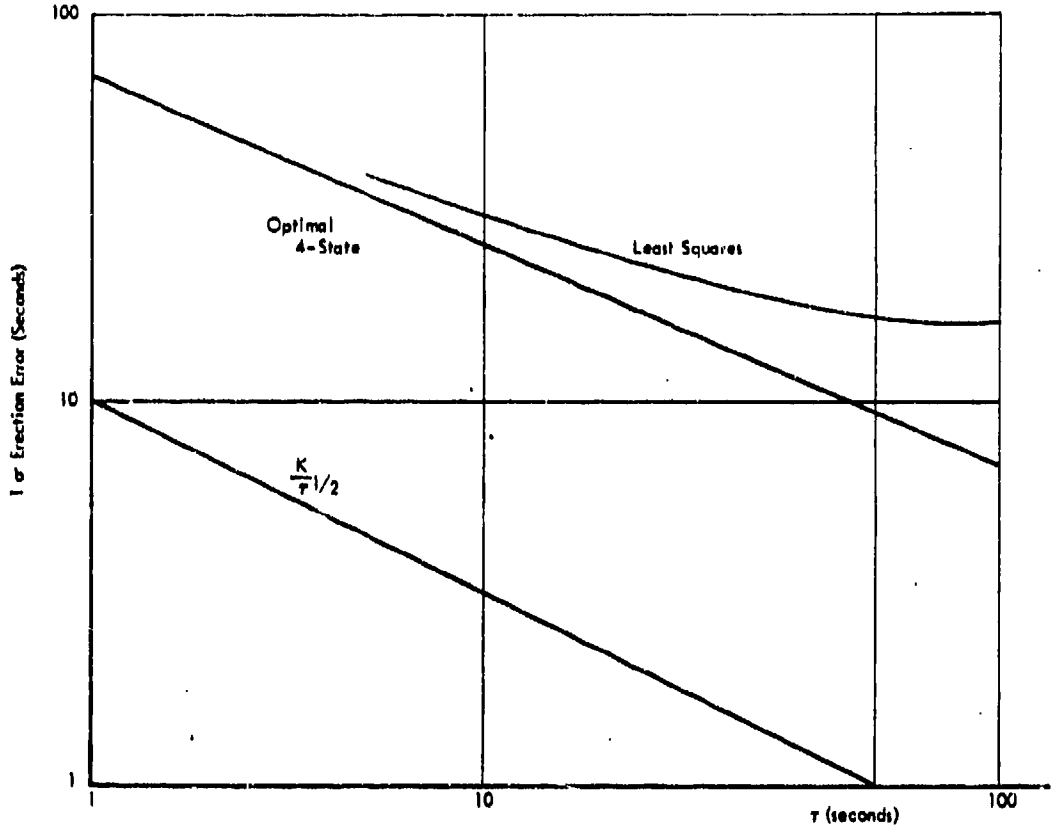


Fig. 18 Dependence of filtering error on noise correlation time

A KALMAN FILTER AUGMENTED MARINE NAVIGATION SYSTEM

by

H. Halamandaris and D. Gades

Satellite Positioning Corporation
16033 Ventura Blvd
Encino, California 91316, USA

NOTATION

P_i, T_i	satellite positions and time at the i^{th} interval
f_g	ground reference oscillator frequency
f_t	stable frequency transmitted from satellite
f_r	user received frequency
Δf	frequency difference (satellite)
$N_{i,j}$	Doppler count for satellite interval i , j
t	time
t_i	time at satellite i^{th} interval
Δt	incremental time
S_i	slant range at i^{th} interval
$\Delta t_{i,j}$	incremental slant range in i , j interval
C	speed of light
C_v	speed of sound in water
$J_{i,j}$	gravity harmonic (i) and order (j)
k	frequency scale constant
δ_L, δ_H	150 MHz and 400 MHz frequency distortions
γ	sonar Doppler beam transmitting angle
$\Delta\theta$	steady state attitude error
$\Delta\phi$	dynamic attitude error
δV	velocity error
T	temperature (water)
S	salinity
D	depth of transducer
$\bar{X}$	state variable vector
A	system dynamics matrix
U	state noise vector
Φ	state transition matrix
W	system noise
V_i	velocity components along i^{th} axis
T_i	coordinate transformation matrix
R_M, R_E	earth radii
$\delta\lambda$	perturbation quantities
A_i	i^{th} axis acceleration
Ω	earth rate
λ	latitude

Λ	longitude
g	gravity
ϵ	gyro drift
ψ_0, η, α	internal gyro angles
τ_0	gyro time constant
K_D	gyro damping gain
Δf_i	sonar Doppler frequency difference along i^{th} axis
f_i	i^{th} axis frequency for sonar Doppler
θ	ship's pitch angle
φ	ship's roll angle
ψ	ship's heading
δ_{ij}	Kronecker delta
Q(i)	variance of white noise
σ_i^2	variance of stochastic process
β_i	reciprocal of stochastic process correlation time
P	covariance matrix
K_k	Kalman gain
R	noise covariance matrix
$\hat{X}$	estimate of state vector
Z, $\hat{Z}$	observation and estimate of observation matrix
H	measurement matrix
ρ, $\dot{\rho}$	range and range-rate vectors
$\hat{i}$, $\hat{j}$, $\hat{k}$	unit vectors along x, y and z axes respectively
T_i	coordinate transformation matrix
$S_{1,i}$ $S_{2,i}$	satellite clock errors

Subscripts

N	north
E	east
D	down
u	up
ST	sea track
sc	sea current
s	starboard axis
f	fore/aft axis
a	aft axis
p	port axis

Superscripts

T	transpose
' "	reference coordinate frames
E	east
n	north
·	time differentiation

A KALMAN FILTER AUGMENTED MARINE NAVIGATION SYSTEM

H. Halamandaris and D. Ozdes

1. INTRODUCTION

With the technical horizons expanding very rapidly in recent years, new demands have been imposed on the performance indices of marine navigation systems. Typical indices are performance accuracy, reliability, maintainability, cost, etc. Reliability and maintainability are enhanced by utilizing proven technology and hardware while retaining a simple system physically. Performance accuracy of hardware commensurate with the reliability and maintainability constraints can be improved by completely integrating the navigation sensors via a digital interface with primary control in a general purpose computer with the controller being a Kalman filter.

The system described and analyzed herein employs a classical dead-reckoning navigation system to provide instantaneous position. Characteristic of all dead-reckoning systems is the monotonically growing error function; hence a navigation aid is required to constrain the long-term positional stability of the system. The Navy Navigation Satellite System (NNSS) provides an independent measure of positional information, and this information is used to optimally reset and calibrate the dead-reckoning system which contains

(i) A Doppler sonar to measure ship's velocity.

(ii) A gyrocompass to provide a measure of ship's heading.

(iii) A velocimeter to provide a measure of the speed of sound in water.

(iv) An inclinometer to provide a measure of ship's pitch and roll.

A general purpose digital computer in which all the mechanization equations are implemented is the central controller.

Section 2 contains a description of the operation for each of the four basic sensors and the NNSS. A generalization from the basic implementation to the format compatible with a Kalman filter implementation is presented in Section 3. Also contained is a development of the pertinent filter mechanization equations and real-time mechanization constraints. A performance summary for various environmental and dynamic conditions is presented in Section 4, which illustrates the feasibility of utilizing a Kalman filter to improve the integrated system performance accuracy.

2. SYSTEM DESCRIPTION

2.1 Integrated System Concepts

The primary objective for the integrated system as configured is to automatically provide continuous, accurate, real-time geodetic position while in no way compromising system operation. As a result, an improved self-contained, shipborne navigation system is synthesized, employing several discrete navigation subsystems. Figure 1 is a functional block diagram of the system which contains the basic constituents in solid blocks and optional navigational systems are shown in the dashed blocks. Instantaneous position velocity and heading are derived from the outputs of the Doppler sonar, gyrocompass and velocimeter with long-term position accuracy stability preserved via the satellite fixes. The nucleus of the system is the interface unit/computer complex. Within the computer/interface all signal transmission, conditioning, mixing, and optimal filtering is implemented.

The computer receives ship's fore/aft and port/starboard velocity (frequency) components from the Doppler sonar and corrects them for speed of sound anomalies derived from the velocimeter data. The component velocities are corrected for the ship's pitch and roll by the vertical reference unit and are resolved about the gyrocompass heading input to produce ship's North and East velocity components. A numerical integration is continuously performed to compute ship's latitude and longitude. An independent absolute position is calculated at the end of every satellite pass, and a system position update is automatically obtained. This update is derived using the optimal (Kalman) filter and the satellite position observation data.

The optional features which can be incorporated in the system to allow complete system automation are as follows:

Ship's steering commands can be generated in the computer and sent to the autopilot.

Point-to-point navigation options can be provided.

Automatic timing sequence controls as a function of distance traveled along course or elapsed time are generated for data acquisition systems.

Real-time data logging for post-mission analysis can be implemented through the use of a magnetic tape recorder.

Ship's track may be plotted through the use of an incremental plotter.

Each primary sensor's operation and theory is delineated below.

2.2 System Description and Theory of Operation

2.2.1 Navy Navigation Satellite System

The Navy Navigation Satellite (NNSS) is a world-wide, all-weather system from which accurate navigational position fixes can be obtained from the data transmitted from the orbiting satellite. The NNSS consists of four near-earth orbiting satellites, four tracking stations, two injection stations, the US Naval Observatory, and a computing center. Any number of shipboard navigational installations, such as that summarized above, can exist.

The navigation satellites are placed in a circular polar orbit at an altitude of approximately 600 nautical miles. The orbital planes of the satellites have a common point along the earth's rotational axis. The geometrical placement of the orbiting satellites allows an earth bound observer to cross directly under the satellite twice daily. The observer receives data from the satellite twice each time he is near the orbit because the satellites appear to traverse longitudinally as the earth rotates. A maximum of 16 fixes is possible at the equator. Realistically, about 12 fixes daily can be realized. The number of fixes available varies as a function of user latitude, as can be seen in Figure 2.

Each satellite orbits the earth in approximately 108 minutes, which is 90 minutes between passes plus roughly 18 minutes of tracking. Through its useful life, each satellite continuously transmits the following phase-modulated data as two-minute messages on two radio frequency carriers: two-minute mark synchronization signals, a 400 Hz reference signal, fixed and variable parameters describing the satellite's polar orbit.

The classical utilization of the satellite data to compute a position fix is similar in concept to any hyperbolic positioning system where the satellite simulates the multiple transmitting stations by its inherent motion relative to the user. Figure 3 illustrates the relative geometry of the user and the satellite. P_1, P_2 and P_3 are three distinct positions of the satellite during one orbital pass. If each of the P_i's differ by two minutes in time, the distance between points is about 960 km. This separation is comparable to baselines on the hyperbolic radio positioning systems. The earthbound navigator lies on the surface of a hyperboloid generated from the satellite information from any two successive points. The intersection of two or more hyperboloids uniquely locates the earthbound user's position.

Additional data points can be used to establish the user's position more accurately. A maximum of 9 data points (P_i) are available during a satellite pass; however, many factors such as geometry (low elevation), radio interference, etc., can cause data to be missed or be invalid.

In determining the P_i's and establishing the user's position, satellite orbital data is required, as well as some measure of relative distance between user and satellite. The satellite transmits both fixed and variable orbital parameters.

The fixed parameters describe the satellite's nominal orbits and are accurate only for a 12 to 16 hour interval. The variable parameters describe the fine structure in the satellite's nominal orbit as a function of time and are correct only for the time at which they are transmitted by the satellite. Thus the satellite's memory stores sufficient variable parameters to describe its orbit at two-minute intervals between subsequent injections of data into its memory. All data transmitted which do not change are permanently wired into the satellite's memory.

The satellite transmits a stable frequency f_t which is received by the user and designated as f_r. The user in turn compares the received frequency to a stable oscillator output frequency f_s to produce a frequency difference

$$\Delta f = f_s - f_r .$$

Figure 4 illustrates the relationship between the frequencies. Over each two-minute time interval, the number of beat cycles are counted or, essentially, an integration of the beat cycles is accomplished.

This information can be directly related to changes in slant range from the satellite to the user in that one frequency count is equivalent to one wavelength of distance traveled by the satellite. Reception of three two-minute satellite messages during an orbital pass, along with the related Doppler shift measured by the user's receiver, defines hyperbolic cones of differential slant range, intersecting at the user's ship's position. Taking data over a minimum of three intervals yields two equations in the two unknowns, and hence a position solution can be realized.

From Figure 4, and using the first two intervals,

$$N_{1,2} = \int_{t_1+\Delta t_1}^{t_2+\Delta t_2} (f_0 - f_r)\, dt = (f_0 - f_t)(t_2 - t_1) + f_0\,(\Delta t_2 - \Delta t_1)\,, \tag{2.1}$$

using the assumption that $f_t = f_r$. Also $\Delta t_2 - \Delta t_1$ is related to the slant range between satellite and receiver at P_1 and P_2,

$$\Delta t_2 - \Delta t_1 = \frac{1}{C}(S_2 - S_1)\,, \tag{2.2}$$

where C is the speed of light, and S_2 and S_1 are the slant ranges from the user to the satellite at P_2 and P_1, respectively. Substituting Equation (2.2) into Equation (2.1) and solving for the difference in distance from the earth-bound observer to two satellite positions, in effect, determines the baseline, e.g.,

$$\Delta S_{1,2} = S_2 - S_1 = \frac{C}{f_0} N_{1,2} - \left(1 - \frac{f_t}{f_0}\right) C\,(t_2 - t_1)\,. \tag{2.3}$$

Similarly,

$$\Delta S_{2,3} = S_3 - S_2 = \frac{C}{f_0} N_{2,3} - \left(1 - \frac{f_t}{f_0}\right) C\,(t_3 - t_2)\,. \tag{2.4}$$

Noting that $t_{i+1} - t_i$, where $i = 1, 1 \ldots n$, multiplied by C is the distance traveled by a particle of light in two minutes allows some simplifying assumptions to be made. However, $f_0 = f_t + \Delta f$ and Δf is very small compared to f_t; thus, to preserve computational accuracy, the above equations must be rearranged to avoid computations of the form $\Delta f / f_0$. This can be easily accomplished by using an additional point. Manipulating a third equation which is similar to the above yields

$$-S_1 + 2S_2 - S_3 = \frac{C}{f_0}(N_{1,2} - N_{2,3}) \tag{2.5}$$

$$-S_2 + 2S_3 - S_4 = \frac{C}{f_0}(N_{2,3} - N_{3,4})\,. \tag{2.6}$$

where the S_i are functions of the satellite position at each point i, where $i = 1, 2, \ldots n$, and the observer position. Hence, two equations in two unknowns allow the determination of the user's two unknown coordinates, latitude and longitude. An alternate implementation is to use a direct ranging approach which obviates the requirement to compute position from the slant range data. This approach tends to preserve accuracy, as the equivalent of GDOP (Geometric Dilution of Position) is not experienced.

The position determination accuracy was seen to be greatly affected by the earth's atmosphere and gravity. Gravity model refinements have continued since 1960. In 1964 the eighth harmonic and order term ($J_{8,8}$) in the gravity expansion model was incorporated, and on 8 January 1965, the terms $J_{12,12}$, $J_{14,14}$, $J_{16,16}$ and $J_{18,18}$ were added to the mechanization; a substantial reduction in position determination inaccuracy resulted. The information carried in the navigation messages, which are phase-modulated, is distorted by the troposphere and ionosphere. Thus, the earth's atmospheric effects were determined to be functions of frequency. The tropospheric refraction is proportional to frequency, as is the Doppler shift which contains the usable information, and can not be isolated. However, the ionospheric refraction was observed to be inversely proportional to frequency and could easily be determined. To enhance the system performance, the satellites transmit two coherent carrier frequencies (150 MHz and 400 MHz). These two frequencies are controlled by a reference oscillator and one frequency is an exact multiple (3/8) of the other. The actual transmission frequencies are 400 MHz offset by 32 KHz (399.968 MHz) and 150 MHz offset by 12 KHz (149.988 MHz). This allows a more precise Doppler count. Let f_H and f_L be the respective high and low frequencies transmitted by the satellite; then

$$f_H = k f_L\,, \tag{2.7}$$

where $k = 8/3$. The frequencies received on the high and low channels respectively are

$$f_{R,H} = f_H + \Delta f_H + \delta_H \tag{2.8}$$

$$f_{R,L} = f_L + \Delta f_L + \sigma_L\,. \tag{2.9}$$

Employing the observed results of frequency distortion allows

$$f_{R,H} = k f_L + k\Delta f_L + \frac{1}{k\delta_L}\,. \tag{2.10}$$

Solving for δ_H from Equations (2.8) and (2.9) yields

$$\delta_H = \frac{1}{1 - k^2} (f_{R,H} - k f_{R,L}) , \qquad (2.11)$$

which is the Doppler refraction correction.

Since the satellite transmits circularly polarized signals, either circularly or linearly, polarized receiving antennas can be used. However, if highly directional, narrow-band antennas are used, two antennas are required, one for each operating frequency, or a single broadband antenna designed for dual frequency operation may be used. The antennas used are larger and more difficult to handle than the simplest antenna arrangement which is composed of two whip antennas, one for each frequency; however, this should be avoided, if possible, since the whip has an overhead null, and satellite data would, therefore, be lost for satellite passes of more than 70^0 elevation angle. The elevation angle limits on satellite coverage do exceed this value.

The satellite signals are received by the user from the time the approaching satellite rises above the radio horizon until the time it sets. The receiving antenna patterns should, therefore, provide complete overhead coverage.

2.3 Velocity Measuring Subsystems

2.3.1 *Sonar Doppler*

The sonar Doppler provides a precise measurement of ship's fore/aft, port/starboard, and up/down velocity relative to the earth for water depths up to 600 feet. The sonar Doppler is a two-axis, bottom referenced speedometer which derives its information from high frequency sonar signals emitted from a transmitter on the boat. The basic operation is delineated below.

The Doppler system employs beams of ultrasonic energy, directed obliquely at the ocean floor at an angle of γ degrees relative to the vertical, to obtain true velocity measurements in the fore-aft and port-starboard directions and the up-down axis. Because the pitch, roll and heave of the ship add other apparent motions relative to the ocean bottom, the Doppler system should employ pairs of beams (one pair angled fore and aft, the other angled to port and starboard, see Figure 5) to compensate for these motions. By averaging the frequency difference between fore and aft, and port and starboard, the true velocities are determined. The concept of four beams is an application of the Janus configuration which is common in airborne Doppler radars. A primary reason for using this configuration is for the cancellation of errors which results from averaging the return signals in the signal processing.

Typical sonar Doppler sensors utilize a 300 KHz sonar signal in a pulsed transmission mode of operation. Signal feedthrough from transmitter to receiver is eliminated and excellent signal-to-noise ratio characteristics are obtained with receiving circuit gating and time varied gain. The pulsed operating mode provides a signal amplitude sufficient to supply accurate signal data at vessel over-the-bottom operating heights up to 600 feet. Beyond 600 to 650 feet the system automatically locks on to the water scatter return; the sonar energy is reflected from small scatterers present in all oceans. The electronic gating is set to observe this reverberation return only from the undisturbed water below the ship, thus excluding the effects of turbulence and boundary layers. The result is a highly accurate and reliable indication of the ship's true velocity relative to the ocean bottom up to 600 feet and relative to the water mass at all greater depths.

Thus, the water mass return represents vehicle motion relative to the sea and the expected accuracy is degraded by the unknown sea currents which are typically on the order of 0.5 to 1.0 knots for short-term correlated noise. The sonar accuracies are tabulated below.

Water Depth (ft)	*Mode*	*Typical (1σ) Accuracy (%)*
1 - 400	Bottom	0.2
400 - 600	Bottom	0.5
Greater than 600	Volume reverberation	0.5*

An error source which is additive to the one above is the receiver/transmitter electronics noise, which is approximately 0.02 knots. The primary contributor is the crystal controlled oscillator which is stable to within 0.01%. The biggest concern in sonar performance is that in the reverberation mode the sea current motion must be statistically added to the sonar error. Typical random sea current drifts are on the order of 0.5 knots to 1.0 knots, depending on the area, with typical correlation distances of 18-20 nautical miles, assuming a Gaussian distribution with exponential autocorrelation function.

---

* Measured relative to sea currents.

Major contributing error sources in a conventional unaugmented Doppler mechanization, which are eliminated in the integrated mechanization, are the errors induced by the vehicle pitch and roll dynamics. The pitch and roll angular resolution will be mechanized and will yield negligible errors. If the resolution is not implemented, the following velocity errors per axis result:

(i) Due to steady state pitch (trim) and roll (list),

$$\delta V \simeq \Delta\theta^2/6570 \ ,$$

where $\Delta\theta$ is the pitch or roll angle in degrees.

(ii) Due to dynamic pitch and roll (sinusoidal)

$$\delta V \simeq 0.101\,\Delta\phi^2/6570 \ ,$$

where $\Delta\phi$ is the amplitude of the pitching or rolling motion in degrees.

2.3.2 Vertical Reference Unit

The vertical reference unit measures ship's pitch and roll through which the sonar Doppler outputs are resolved to essentially create a stabilized, locally-level coordinate frame in which the system velocities are expressed. The basic sensors in the vertical reference unit are two single-axis fluid-damped pendulums (inclinometers) which output the deflection of the ship from the local vertical. Their outputs are whole angle indications of ship's pitch and roll. These angles are then used to generate the Euler transformation matrix consistent with the rotational dynamics of the ship.

2.3.3 Velocimeter

The velocimeter is a precision instrument for measuring the velocity of sound in water to an accuracy of one part in ten thousand. Its output is used to weight the sonar Doppler output pulses.

The operation of the velocimeter is based upon the sing-around principle. A pair of piezoelectric ceramic transducers and two reflectors are mounted to form a sound path of fixed length in the water. In operation, a pulse of acoustic energy is transmitted through the water, received, amplified and used to generate another pulse of acoustic energy. The repetition frequency of this regenerative action is dependent upon the transit time of the signal pulse and is therefore a measure of the propagation velocity. Errors resulting from the flow of water along the sound path length are minimized by folding the sound path.

The basic parameters which cause a variation in the speed of sound are water temperature, pressure and salinity. A typical expression for C_v is

$$C_v = 4422 + 11.25T - 0.045T^2 + 0.0182d + 4.3(S-34) \ , \qquad (2.12)$$

where C_v is in ft/sec, T is temperature in degrees Fahrenheit, d is depth in feet and S is salinity in parts per thousand. The primary items of concern in this expression are as follows:

(a) Temperature - a 20 - 25 degree F variation can be expected between low latitudes to polar areas.

(b) Salinity - large variations can occur if the vehicle is in combining salt water currents, or near river outlets.

2.3.4 Gyrocompass

The gyrocompass, with its integral gyroscope, is controlled such that it continually aligns itself with the meridian and tracks true north. Hence, its natural output is ship's heading. Sonar Doppler velocity components in the stabilized, locally-level plane are resolved about heading to provide ship's velocity components in north and east components. Figure 1 indicates a feedback torquing signal from the computer to compensate for the ship's instantaneous position and spatial rate. In the gyrocompass error model which is presented in the following section, all gyro drifts are assumed to be either deterministic or to possess exponentially autocorrelated noise functions.

3. SYSTEM ANALYTICAL MODEL

3.1 Introduction

The integrated multi-sensor marine navigation system defined in Section 2 employs optimal signal processing and multiply-redundant navigation data, in the form of range and/or range-rate or range-difference and/or range-rate-difference, in synthesizing minimum-variance state vector estimates.

The state variables of the state vector consist of position, velocity, heading and the stochastic models of the sensor error sources. The optimal estimate of the navigation variables are employed in updating (i.e., servoing or correcting) the navigation system; the prediction of the sensor error characteristics leads to cross-calibration of sensors; thus the navigation performance is further enhanced.

The physical dynamics involved in optimally implementing the multisensor navigation system require a precise definition of the propagation of linearized error differential equations characteristic to the Doppler-Sonar gyrocompass system augmented by the satellite data processor.

In the optimum linear continuous estimators the system error state vector is defined by the relationship

$$\dot{\bar{X}} = A\bar{X} + U , \tag{3.1}$$

where X is the n-dimensional error state vector, A is an $n \times n$ coefficient matrix and U is state noise vector. The corresponding discrete equation is

$$[\bar{X}]_{k+1} = [\Phi]_{k+1,k}[X]_k + [W]_k , \tag{3.2}$$

where Φ is the well known state transition matrix characterizing the error transition from k^{th} to $(k+1)^{th}$ iteration.

3.2 System Sensor Error Models

The primary system sensor (i.e., Doppler-Sonar (D/S), gyrocompass and satellite receiver) error models are usually obtained by a first-order perturbation of mechanization equations.

The (D/S) mechanization equations are

$$[\bar{V}]_s = [T_1][V'']_s , \tag{3.3}$$

where $(V_s)^T = [V_f V_s V_D]$ is the velocity vector defined in the ship's static coordinate system, $(V''_s)^T = [V''_f V''_s V_D]$ is the velocity vector defined in the ship's dynamic (i.e., pitched and rolled) coordinate system* and the matrix $[T_1]$ provides transformation from ship's dynamic to state coordinate systems.

The ship's velocities in the navigation coordinate system are obtained by

$$[\bar{V}]_{sT} = [T_2][V_s]_s , \tag{3.4}$$

where $[V]^T_{sT} = [V_E V_N V_D]$ is the ship's sea-track velocity vector defined in a (east, north, up) coordinate system sharing the same origin of the ship's coordinate systems. The matrix $[T_2]$ provides transformation from ship's static to navigation coordinate systems. The ship's velocities are resolved through the sea-track heading angle ψ_{sT} to northerly and easterly components.

Combining Equations (3.3) and (3.4) yields

$$[V_{sT}] = [T_3][V''_s] , \tag{3.5}$$

where

$$[T_3] = [T_2][T_1] . \tag{3.6}$$

However, if the D/S system is operated in the volume reverberation mode, a sea current vector, $[V]^T_{sc} = [V^E_{sc} V^N_{sc} 0]$, must† be subtracted from Equation (3.5). Hence

$$[\bar{V}]_{sT} = [T_3][V''_s] - [\bar{V}]_{sc} K \tag{3.7}$$

where $K = i$ $(i = 0, 1)$ in the bottom and sea volume reverberation modes, respectively.

The latitude and longitude rates of the system are obtained by dividing sea-track velocity components by the radii of curvature of the earth. First-order perturbation of these rates, neglecting second-order terms

$$\frac{\delta R_N}{R_N^2} \simeq \frac{\delta R_E}{R_E^2} \simeq 0$$

$$\delta R_N \simeq \delta R_E \simeq 0 ,$$

yields

$$\delta\dot{\lambda} = \frac{\delta V_N}{R_N} \tag{3.8}$$

$$\delta\dot{\Lambda} = \frac{\delta V_E}{R_E \cos\lambda} + \frac{\delta\lambda V_E \sin\lambda}{R_E \cos^2\lambda} \tag{3.9}$$

where R_i $(i = E, N)$ are the radii of curvature of the earth[1].

---

* All coordinate systems used are Cartesian, right hand oriented.

† Sea current velocity vector is defined in the navigation coordinate system and is resolved through the sea current heading angle to the northerly and easterly components.

The velocity error rate equations obtained from Equation (3.7) in scalar form are

$$\begin{aligned}\delta\dot{V}_N = {} & \{\delta\dot{V}_f \cos\psi_G - \delta V_f\dot{\psi}_G \sin\psi_G - \delta\dot{\psi}_G V_f \sin\psi_G - \delta\psi_G\dot{V}_f \sin\psi_G - \\ & - \delta\psi_G V_f\dot{\psi}_G \cos\psi_G - \delta\dot{V}_s \sin\psi_G - \delta V_s\dot{\psi}_G \cos\psi_G - \\ & - \delta\dot{\psi}_G V_s \cos\psi_G - \delta\psi_G\dot{V}_s \cos\psi_G + \delta\psi_G V_s\dot{\psi}_G \sin\psi_G - \\ & - \delta\dot{V}_{sc} \cos\psi_{sc} - \delta V_{sc}\dot{\psi}_{sc} \sin\psi_{sc} - \delta\dot{\psi}_{sc}V_{sc} \sin\psi_{sc} - \\ & - \delta\psi_{sc}\dot{V}_{sc} \sin\psi_{sc} - \delta\psi_{sc}V_{sc}\dot{\psi}_{sc} \cos\psi_{sc}\} \end{aligned} \tag{3.10}$$

$$\begin{aligned}\delta\dot{V}_E = {} & \{\delta\dot{V}_f \sin\psi_G + \delta V_f\dot{\psi}_G \cos\psi_G + \delta\dot{\psi}_G V_f \cos\psi_G + \delta\psi_G\dot{V}_f \cos\psi_G - \\ & - \delta\psi_G V_f\dot{\psi}_G \sin\psi_G + \delta\dot{V}_s \cos\psi_G - \delta V_s\dot{\psi}_G \sin\psi_G - \\ & - \delta\dot{\psi}_G V_s \sin\psi_G - \delta\psi_G\dot{V}_s \sin\psi_G - \delta\psi_G V_s\dot{\psi}_G \cos\psi_G - \\ & - \delta\dot{V}_{sc} \sin\psi_{sc} + \delta V_{sc}\dot{\psi}_{sc} \cos\psi_{sc} + \delta\dot{\psi}_{sc}V_{sc} \cos\psi_{sc} + \\ & + \delta\psi_{sc}\dot{V}_{sc} \cos\psi_{sc} - \delta\psi_{sc}V_{sc}\dot{\psi}_{sc} \sin\psi_{sc}\} \end{aligned} \tag{3.11}$$

$$\delta\dot{V}_U = -\delta\dot{V}_D \,. \tag{3.12}$$

$\delta\psi_{sc}$ and δV_{sc} are modeled as time-correlated noise sources. The $\delta\dot{\psi}_{sc}$ and $\delta\dot{V}_{sc}$ terms have second-order effects which are neglected. The $\delta\dot{V}_i$ $(i = s,f,D)$ are modeled as time-correlated noise sources and δV_i $(i = s,f,D)$ terms are presented in Section 3.2.2.

3.2.1 *Gyrocompass Rate Error Equations*

A typical two-degree-of-freedom gyrocompass is shown in Figure 6. The rates sensed by the gyrocompass are contained in the bracketed terms. In theory the compensation of the bracketed terms is implemented by applying signals of the same magnitude but of opposite sign to the corresponding summing functions. In practice, however, the inaccessibility of the summing functions requires compensating terms to be applied to the level and azimuth or to level (only) torquers of the gyrocompass. In the analysis below a theoretical compensation method is assumed; this deviation from the practice simplifies the computer simulation of error equations and, most importantly, yields an universal error model compatible with most gyrocompasses available.

The error performance in each case, with few exceptions, yields the same results.

Referring to Figure 6, the first-order perturbation of the mechanization equations, and assuming δR_i $(i = E,N) \simeq 0$, $\delta K_D \simeq 0$, $\delta\Omega \simeq 0$, $\delta K_M \simeq 0$, $\delta\tau \simeq 0$, $A_E\delta\alpha >> \alpha\delta A_E$, $g\delta\theta >> k_D\delta A_N$, yields

$$\delta\dot{\lambda} = -\left\{\left[\delta\lambda\left(\Omega\cos\lambda + \frac{V_E}{R_E}\sec^2\lambda\right) + \frac{\delta V_E}{R_E}\tan\lambda\right] + g\delta\alpha + \epsilon_\psi\right\} \tag{3.13}$$

$$\delta\dot{\eta} = \left\{-\delta\psi\left(\Omega\cos\lambda + \frac{V_E}{R_E}\right) - K_D g\delta\eta + A_E\delta\alpha - \frac{\delta V_N}{R_N} - \delta\dot{\psi}_T + \epsilon_\eta\right\} \tag{3.14}$$

$$\delta\dot{\alpha} = \left\{\frac{gK_M}{\tau_g}\delta\eta - \frac{1}{\tau_g}\delta\alpha\right\}. \tag{3.15}$$

The $\delta\psi_T$ term in Equation (3.14) is the residual torquing error. The ϵ_i $(i = \psi,\eta)$ are the total instrument drift errors due to bias, noise (short and/or long correlated) warm-up noise, scale factor error, non-orthogonality, mass unbalance (g-sensitive), anisoelasticity (g^2-sensitive). The orthogonality error angles and the scale factor errors are multiplied by the corresponding total rates (i.e., vehicle and earth rate). *The Equations (3.13) and (3.14) are valid whether the gyrocompass consists of two single-degree-of-freedom gyros or a two-degree-of-freedom gyro.* The number of elements to be implemented in the on-board computer is dictated by the accuracy requirements, memory allotments and the type of compensation used.

3.2.2 *Doppler/Sonar Error Equation*

The Doppler/Sonar output is the calibrated velocity vector defined in the ship's dynamic coordinate frame, i.e.,

$$(\bar{V}'')^T = (V_f'' \; V_{sc}'' \; V_D'') \,.$$

Referring* to Section 2,

* In Equations (3.16) through (3.18) the sonar beam angles are assumed to be 30°, measured from the vertical (down) direction.

$$V_f'' = \frac{\Delta f_f c_v}{2f_t} \quad (\Delta f_f = f_f - f_a) \tag{3.16}$$

$$V_a'' = \frac{\Delta f_a c_v}{2f_t} \quad (\Delta f_a = f_a - f_p) \tag{3.17}$$

$$V_0'' = \frac{c_v}{\sqrt{3}} \left(\frac{\Sigma f_a}{2f_t} - 1 \right) = \frac{c_v}{\sqrt{3}} \left(\frac{\Sigma f_f}{2f_t} - 1 \right) \tag{3.18a}$$

$$\therefore \frac{c_v}{4f_t\sqrt{3}} (\Sigma f_f + \Sigma f_a - 4f_t) \ , \ (\Sigma f_f = f_f + f_a) \ , \ (\Sigma f_a = f_a + f_p) \ , \tag{3.18b}$$

where f_f , f_a , f_s and f_p are the fore, aft, starboard and port transducer frequency outputs of the transducer; c_v is the velocity of propagation of sound in the sea medium. The f_t is the transmitting frequency.

Combining Equations (3.16) through (3.18b) with (3.3) and first-order perturbation of the result, with the assumptions $\delta(\Delta f_f) \simeq 0$, $\delta(\Delta f_a) \simeq 0$, yields

$$\begin{aligned} \delta V_f = {} & \left\{ \left(\frac{c_v}{2f_t} \right) \left(\frac{\delta c_v}{c_v} - \frac{\delta f_t}{f_t} \right) \left[\Delta f_f \cos\theta + \frac{\sin\theta}{2\sqrt{3}} \left(\Sigma f_f + \Sigma f_a - 4f_t \right) \right. \right. - \\ & - (\delta\theta) \left[\Delta f_f \sin\theta - \frac{\cos\theta}{2\sqrt{3}} \left(\Sigma f_f + \Sigma f_a - 4f_t \right) \right] + \\ & \left. \left. + \left[\delta\Sigma f_f + \delta\Sigma f_a - 4\delta f_t \right] \frac{\sin\theta}{2\sqrt{3}} \right\} \end{aligned} \tag{3.19}$$

$$\begin{aligned} \delta V_s = {} & \left(\frac{c_v}{2f_t} \right) \left\{ \left(\frac{\delta c_v}{c_v} - \frac{\delta f_t}{f_t} \right) \left[\Delta f_f \sin\theta \sin\varphi + \Delta f_a \cos\varphi - \frac{1}{2\sqrt{3}} (\Sigma f_f + \right. \right. \\ & + \Sigma f_a - 4f_t) \sin\varphi \cos\theta \Big] + \delta\theta \left[\Delta f_f \cos\theta \sin\varphi + \sin\varphi \sin\theta \times \right. \\ & \times \left. \left(\frac{\Sigma f_f + \Sigma f_a - 4f_t}{2\sqrt{3}} \right) \right] + \delta\varphi \left[\Delta f_f \cos\varphi \sin\theta - \Delta f_a \sin\varphi - \right. \\ & \left. - \cos\varphi \cos\theta \left(\frac{\Sigma f_f + \Sigma f_a - 4f_t}{2\sqrt{3}} \right) \right] - \left(\frac{\sin\varphi \cos\theta}{2\sqrt{3}} \right) \left[\delta\Sigma f_f + \right. \\ & \left. \left. + \delta\Sigma f_a - 4\delta f_t \right] \right\} . \end{aligned} \tag{3.20}$$

Equations (3.19) and (3.20) contain the $\delta V_0''$ term. φ and θ , of the trigonometric terms shown in the equations, are the roll and pitch angles, respectively. The corresponding pitch and roll errors (i.e, $\delta\theta$ and $\delta\varphi$) are due to attitude readout or inclinometer readout errors and pitch and roll amplitude* and frequency errors, i.e.,

$$\delta\theta = \delta\theta_0 + \delta\theta_s \sin\omega_\theta t + \delta\omega_\theta \theta_s \cos\omega_\theta t + n(\theta) \tag{3.21}$$

$$\delta\varphi = \delta\varphi_0 + \delta\varphi_s \sin(\omega_\varphi t + \gamma) + \delta\omega_\varphi \varphi_s \cos(\omega_\varphi + \gamma) + n(\varphi) \ . \tag{3.22}$$

The $[\delta\omega_i (i = \theta, \varphi) = 2\pi\delta f_i]$ is a function of the ship's dynamics; γ_i is an arbitrary phase angle. The noise terms $n(i)$ $(i = \theta, \varphi)$ are defined by

$$E\{n_k(i)[n_j(i)]^T\} = \delta_{kj} Q(i) \ , \tag{3.23}$$

where δ_{kj} is the Kronecker delta and $Q(i)$ is the so-called white-noise variance† .

3.2.3 *Velocimeter Error Equations*

Referring to Section 2.3.3, the velocimeter output is defined by

$$C_v = \{4422 + 11.25T - 0.045T^2 + 0.0182D + 4.3(S-34)\} \ , \tag{3.24}$$

---

* A constant plus sinusoid is assumed.

† Typical values for $f_i = [(6 \text{ to } 10) \text{ sec}]^{-1}$.

where

C_v = speed of sound in water (ft/sec)

T = water temperature (°F)

D = depth of transducer (ft)

S = salinity of water (parts per thousand).

The first-order perturbation of Equation (3.24) yields

$$\delta C_v = 11.25\delta T - 0.09T\delta T + 0.016\delta D + 4.3\delta S \ . \tag{3.25}$$

The errors due to measurement of the depth of the transducer and salinity of water are small and hence may be neglected. Hence

$$\delta C_v \simeq (11.25 - 0.09T)\,\delta T \ . \tag{3.26}$$

The rate of change of the temperature error is modeled as an exponentially correlated noise process.

3.2.4 Models for Exponentially Correlated Noise Sources

The exponentially correlated noise sources k_i, stated in previous sections, are modeled by

$$E[k_i k_i^*] = \sigma_i^2 e^{-\beta_i|\tau|} \ , \tag{3.27}$$

where σ_i^2 is variance of the process and β_i is the inverse correlation time. Note that Equation (3.27) is obtained by passing a white-noise process through a first-order filter.

3.2.5 The System State Transition Matrix

The equations presented in Sections 3.1 and 3.1.1 through 3.1.3 are in the form of

$$\dot{\underline{X}} = A\underline{X} + \underline{W} \ . \tag{3.28}$$

The state transition matrix Φ is defined by[2]

$$\exp\left(\int_0^{\Delta t} A(\tau)\,d\tau\right) = \left\{I + \int A(\tau)\,d\tau + \frac{1}{2!}\left(\int A(\tau)\,d\tau\right)^2 + \frac{1}{3!}\left(\int A(\tau)\,d\tau\right)^3 \cdots\right. \tag{3.29}$$

Expansion of each term yields

$$A(\tau) = \left\{A_0 + \tau\dot{A}_0 + \frac{\tau^2}{2!}\ddot{A}_0 \cdots\right\} \tag{3.30}$$

$$\int_0^{\Delta t} A(\tau)\,d\tau = \left\{A_0\Delta t + \dot{A}_0\frac{\Delta t^2}{2!} + \ddot{A}_0\frac{\Delta t^3}{3!} \cdots\right\} \tag{3.31}$$

$$\left(\int_0^{\Delta t} A(\tau)\,d\tau\right)^2 = \left\{A_0^2(\Delta t)^2 + (A_0\dot{A}_0 + \dot{A}_0A_0)\frac{\Delta t^3}{2!}\right\} \tag{3.32}$$

$$\left(\int_0^{\Delta t} A(\tau)\,d\tau\right)^3 = \left\{A_0\Delta t^3 \cdots\right\} . \tag{3.33}$$

Hence

$$\exp\left(\int_0^{\Delta t} A(\tau)\,d\tau\right) = \left\{I + \Delta tA_0 + \frac{\Delta t^2}{2!}(\dot{A}_0 + A_0^2) + \frac{\Delta t^3}{3!}\left[\ddot{A}_0 + \frac{3}{2}(A\dot{A}_0 + \dot{A}_0A_0) + A_0^3\right] + \cdots\right\} \tag{3.34}$$

Such expansion, however, has some restrictions[2]. Assuming Δt to be small and $A(t)$ constant during the interval, Equation (3.34) reduces to a simple Taylor series expansion. The computation of Equation (3.34) contains only a sufficient number of terms so that additional terms are negligible by comparison with the partial sum to that point[3]. For this marine navigation system, the state transition matrix is computed by truncating the series after the third term*.

Using the above approximations allows writing the discrete form

$$\underline{X}_{k+1} = \Phi\underline{X}_k \ . \tag{3.35}$$

* It is noted, however, that truncation beyond two terms for up to 4 hours of navigation, with Δt = 60 sec, yields less than 1% error. Such a method is equivalent to simple rectangular integration.

The state transition matrix obtained is in the form of

$$\Phi = \left[\begin{array}{c|c} \Phi_{11} & \Phi_{12} \\ \hline 0 & \Phi_{22} \end{array}\right], \tag{3.36}$$

where Φ_{11} and Φ_{12} are derived by forming the A matrix from the equations given in previous sections. The Φ_{22} matrix contains exponential terms $e^{-\beta_i \Delta t}$. The simulation state vector is presented in Table I.

The ship-borne mechanization of a large state vector, as presented in Table I, is constrained by memory and the computational speed of the computer. Judicious selection of state variables through the use of a sophisticated simulation program will (and indeed does) reduce the size of the state vector to less than one half of the size shown.

3.3 Kalman Mechanization

The Doppler/Sonar-gyrocompass navigation system is augmented by the NNSS generated Doppler signals received and processed by the receiver. Optimal mixing of sensor signals are implemented by the Kalman filter (Ref.4, Vol.III, pp.223-280). The Kalman equations are

(a) Extrapolate

$$P_k = \Phi P_{k-1} \Phi^T + Q_{k-1} . \tag{3.37}$$

(b) Filter

$$K_k = P_k H_k^T (H_k P_k H_k^T + R_k)^{-1} \tag{3.38}$$

$$P_k^* = (I - K_k H_k) P_k \tag{3.39}$$

$$\hat{X}_k^* = \Phi \hat{X}_k + K_k (n_k - \hat{n}_k) . \tag{3.40}$$

Equations (3.37) and (3.39) are the covariance of errors in the estimate before and after the updating process, respectively[4]. The estimate vector is set to zero following a correction after each update.

3.3.1 The State Noise Covariance Matrix Q

The state noise covariance matrix Q is in the form of

$$Q = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & Q_{22} \end{array}\right]. \tag{3.41}$$

The Q_{22} matrix consists of diagonal elements

$$q_i = \sigma_i^2 (1 - e^{-2\beta_i \Delta t}) , \tag{3.42}$$

where σ_i^2 is the variance of the process and β_i is the inverse correlation time. If the correlation is defined spatially, it is normalised by the ship's sea-track velocity.

3.3.2 The Kalman Extraction Matrix H

The satellite dynamic signal processing will yield range and range-rate measurements. The processing of the signals is a characteristic of the receiver used. Correspondingly, the user has the options of available optimal data-averaging techniques and implementing range and/or range-rate or range-difference and/or range-rate-difference methods.

Let $Z = f(x)$ define the actual observation, where x is the dynamic state of the satellite and the navigator. If $Z = f(x)$ is expanded in a Taylor series and higher order terms are neglected,

$$\delta Z = z = H\delta X + v_i , \tag{3.43}$$

where z is the measurement error vector and v_i is the observation noise. The H matrix shows how the observed errors are distributed in the state space. The H matrix is defined by

$$H = \frac{\partial Z}{\partial x} = \left[\begin{array}{c} H_\rho \\ \hline H_{\dot\rho} \end{array}\right]. \tag{3.44}$$

Let the dynamic measurement consist of range and range-rate vector, i.e.,

$$Z = \left[\begin{array}{c} \rho \\ \hline \dot\rho \end{array}\right], \tag{3.45}$$

where

$$\bar{\rho} = (X_{si} - x)\hat{i} + (Y_{si} - y)\hat{j} + (Z_{si} - z)\hat{k} \quad (3.46)$$

$$\dot{\bar{\rho}} = (\dot{X}_{si} - \dot{x})\hat{i} + (\dot{Y}_{si} - \dot{y})\hat{j} + (\dot{Z}_{si} - \dot{z})\hat{k} . \quad (3.47)$$

The $\hat{i}$, $\hat{j}$, $\hat{k}$ are unit vectors in the geocentric, right-handed Cartesian coordinate frame*.

Correspondingly, the absolute range is

$$|\bar{\rho}| = \rho = \sqrt{(\Delta x^2 + \Delta y^2 + \Delta z^2)} , \quad (3.48)$$

where

$$\Delta i \ (i \neq x, y, z) = (i_s - i) , \quad (3.49)$$

and the absolute range rate is

$$|\dot{\bar{\rho}}| = \dot{\rho} = \frac{\bar{\rho} \cdot \dot{\bar{\rho}}}{|\bar{\rho}|} . \quad (3.50)$$

Let the state vector consist of the ship's and satellite's position and velocity errors, i.e., a total of 12 state variables. Hence

$$\frac{\partial \rho}{\partial (x, y, z)} = \left[\left(\frac{x - X_{si}}{\rho} \right) \left(\frac{y - Y_{si}}{\rho} \right) \left(\frac{z - Z_{si}}{\rho} \right) \right] = H_{i1} , \quad (3.51)$$

where i is the index of the observed satellite.

$$\frac{\partial \rho}{\partial (\dot{x}, \dot{y}, \dot{z})} = H_{i2} = 0 \quad (3.52)$$

$$\frac{\partial \rho}{\partial (X_{si}, Y_{si}, Z_{si})} = - H_{i1} = H_{i3} \quad (3.53)$$

$$\frac{\partial \rho}{\partial (\dot{X}_{si}, \dot{Y}_{si}, \dot{Z}_{si})} = H_{i4} = - H_{i2} = 0 . \quad (3.54)$$

Hence

$$H_\rho = \left[H_{i1} \middle| \ H_{i2} = 0 \middle| \ H_{i3} = - H_{i1} \middle| \ H_{i4} = 0 \right] . \quad (3.55)$$

Similarly,

$$\frac{\partial \dot{\rho}}{\partial (x, y, z)} = \left[\left(\frac{\dot{x} - \dot{X}_{si}}{\rho} - \frac{x - X_{si}}{\rho^2} \dot{\rho} \right) \left(\frac{\dot{y} - \dot{Y}_{si}}{\rho} - \frac{y - Y_{si}}{\rho^2} \dot{\rho} \right) \left(\frac{\dot{z} - \dot{Z}_{si}}{\rho} - \frac{z - Z_{si}}{\rho^2} \dot{\rho} \right) \right]$$

$$= H_{s1} \quad (3.56)$$

$$\frac{\partial \dot{\rho}}{\partial (\dot{x}, \dot{y}, \dot{z})} = \left[\left(\frac{x - X_{si}}{\rho} \right) \left(\frac{y - Y_{si}}{\rho} \right) \left(\frac{z - Z_{si}}{\rho} \right) \right] = H_{s2} = H_{i1} \quad (3.57)$$

$$\frac{\partial \dot{\rho}}{\partial (X_{si}, Y_{si}, Z_{si})} = H_{s3} = - H_{s1} \quad (3.58)$$

$$\frac{\partial \dot{\rho}}{\partial (\dot{X}_{si}, \dot{Y}_{si}, \dot{Z}_{si})} = H_{s4} = - H_{s2} . \quad (3.59)$$

Hence

$$H_{\dot{\rho}} = \left[H_{s1} \middle| \ H_{s2} \middle| \ H_{s3} = - H_{s1} \middle| \ H_{s4} \right] . \quad (3.60)$$

---

* The satellite and navigator position (velocity) are defined in terms of orbital and geodetic parameters, respectively. In order to compute range and range-rate vectors, all positions and velocities are transformed to the geocentric coordinate frame.

These universal models are applicable to any navigation satellite system. For example, when the NNSS is used, only one satellite per observation is available; thus the user may prefer to process range only, or range-rate only or range and range-rate models. If a constellation of satellites is available, as in the TRW-Navstar system, the H_ρ and $H_\dot{\rho}$ matrices are computed one per satellite; for example, if three satellites are simultaneously available,

$$H^T = [H_\rho^1 H_\rho^2 H_\rho^3 \mid H_{\dot{\rho}}^1 H_{\dot{\rho}}^2 H_{\dot{\rho}}^3] , \tag{3.61}$$

where each H_ρ^i $(i = 1, 2, 3)$ and $H_{\dot{\rho}}^i$ $(i = 1, 2, 3)$ are in the form of Equations (3.55) and (3.60), respectively. Note that the number of satellites determines the number of rows of the extraction matrix. If a differencing scheme is employed in a constellation type system,

$$H^T = [(H_\rho^1 - H_\rho^2) \mid (H_\rho^1 - H_\rho^3) \mid (H_{\dot{\rho}}^1 - H_{\dot{\rho}}^2) \mid (H_{\dot{\rho}}^1 - H_{\dot{\rho}}^3)] . \tag{3.62}$$

The differencing scheme, with slight modifications in the transition matrix, is also applicable to the NNSS (single satellite per observation), i.e.,

$$H = \begin{bmatrix} H_\rho(t = t_a) - H_\rho(t = t_b) \\ \hline H_{\dot{\rho}}(t = t_a) - H_{\dot{\rho}}(t = t_b) \end{bmatrix} . \tag{3.63}$$

Equation (3.56), i.e., H_{12}, must be computed to evaluate its numerical significance: for example, for high altitude satellites it may be neglected.

For the NNSS, the satellite errors are not carried in the extraction matrix H; hence H_{13}, H_{14}, H_{23}, and H_{24} are deleted, i.e.,

$$H = \begin{bmatrix} H_{11} & 0 \\ \hline H_{21} & H_{22} = H_{11} \end{bmatrix} . \tag{3.64}$$

The satellite errors which are not lumped to the observation noise, v, are added to the observation vector, i.e.,

$$z = H\delta\vec{X} + L(si) + v , \tag{3.65}$$

where $L(si)$ is defined by two time-dependent bias and noise processes:

$$L(si) = B_{1,1}t + B_{2,1}t^2 + n_{si} . \tag{3.66}$$

The n_{si} are modeled as time-correlated noise sequences.

$$E[n_{si}n_{si}^T] = \sigma_{si}^2 e^{-\beta_{si}\Delta t} , \tag{3.67}$$

where $B_{1,1}$, $B_{2,1}$ and n are approximated identically for all NNSS satellites* and are added as state variables 46, 47, 48 to the state vector presented in Table I.

Recall that the state transition matrix was derived in terms of the geodetic coordinate parameters; thus the covariance of errors and the corresponding error vector are defined in the geodetic coordinate system. On the other hand, the extraction matrix H is derived in terms of geocentric coordinates; consequently, the observation vector for a six-element system is modified as

$$z = H^* \delta\vec{X} + L(si) + v \tag{3.68}$$

$$H^* = H \begin{bmatrix} T_q & 0 \\ \hline 0 & T_q \end{bmatrix} , \tag{3.69}$$

where T_q is a geodetic to geocentric transformation matrix and is defined by

$$T_q = \begin{bmatrix} -\sin\Lambda & (-\sin\lambda\cos\Lambda) & (\cos\lambda\cos\Lambda) \\ \cos\Lambda & (-\sin\lambda\sin\Lambda) & (\cos\lambda\sin\Lambda) \\ 0 & \cos\lambda & \sin\lambda \end{bmatrix} = \begin{bmatrix} C_{xe} & C_{xn} & C_{xu} \\ C_{ye} & C_{yn} & C_{yu} \\ C_{ze} & C_{zn} & C_{zu} \end{bmatrix} . \tag{3.70}$$

To this point all H elements corresponding to the position and velocities are defined. The partial derivatives of the range and range-rate with respect to the remaining variables are zero. Referring to Table I, the x, y position errors are defined in terms of longitude and latitude, and in units of radians, respectively. After converting the pertinent variables to feet, the system H matrix is defined by

$$H_{1,4} = [h_a h_b h_c] \begin{bmatrix} C_{xe} \\ C_{ye} \\ C_{ze} \end{bmatrix} \cos\lambda \tag{3.71}$$

---

* In the sea-borne mechanization, $B_{1,1}$ and the noise process are modeled.

$$H_{1,6} = [h_a h_b h_c] \begin{bmatrix} C_{xn} \\ C_{yn} \\ C_{zn} \end{bmatrix} \tag{3.72}$$

$$H_{1,u6} = [h_a h_b h_c] \begin{bmatrix} C_{xu} \\ C_{yu} \\ C_{zu} \end{bmatrix} \tag{3.73}$$

$$H_{1,u4} = \frac{t}{R_k} \tag{3.74}$$

$$H_{1,u7} = \frac{t^2}{R_k} \tag{3.75}$$

$$H_{1,u8} = \frac{t}{R_k} \tag{3.76}$$

$$H_{2,1} = H_{1,6} \tag{3.77}$$

$$H_{2,2} = H_{1,u} \tag{3.78}$$

$$H_{2,4} = [h_d h_e h_f] \begin{bmatrix} C_{xe} \\ C_{ye} \\ C_{ze} \end{bmatrix} R_E \cos\lambda \tag{3.79}$$

$$H_{2,6} = [h_d h_e h_f] \begin{bmatrix} C_{xn} \\ C_{yn} \\ C_{zn} \end{bmatrix} R_N \tag{3.80}$$

$$H_{2,u7} = H_{1,u6} \tag{3.81}$$

$$H_{2,u8} = [h_d h_e h_f] \begin{bmatrix} C_{xu} \\ C_{yu} \\ C_{zu} \end{bmatrix} R_k \tag{3.82}$$

$$H_{2,u4} = 1 \tag{3.83}$$

$$H_{2,u7} = 2t \tag{3.84}$$

$$H_{2,u8} = 1 . \tag{3.85}$$

where R_E , R_N are shown in Reference 1 and

$$R_k = R_0 (1 - e \sin^2\lambda) \tag{3.86}$$

$$[h_a h_b h_c] = \left[\left(\frac{\dot{x} - \dot{X}_{si}}{\rho}\right)\left(\frac{y - Y_{si}}{\rho}\right)\left(\frac{z - Z_{si}}{\rho}\right)\right] \tag{3.87}$$

$$[h_d h_e h_f] = \left[\left(\frac{\dot{x} - \dot{X}_{si}}{\rho} - \frac{x - X_{si}}{\rho^2}\dot{\rho}\right)\left(\frac{\dot{y} - \dot{Y}_{si}}{\rho} - \frac{y - Y_{si}}{\rho^2}\dot{\rho}\right)\left(\frac{\dot{z} - \dot{Z}_{si}}{\rho} - \frac{z - Z_{si}}{\rho^2}\dot{\rho}\right)\right] \tag{3.88}$$

$$\dot{\rho} = \frac{1}{\rho}\left[(X_{si} - x)(\dot{X}_{si} - \dot{x}) + (Y_{si} - y)(\dot{Y}_{si} - \dot{y}) + (Z_{si} - z)(\dot{Z}_{si} - \dot{z})\right] \tag{3.89}$$

$$\rho = \sqrt{\left[(X_{si} - x)^2 + (Y_{si} - y)^2 + (Z_{si} - z)^2\right]} . \tag{3.90}$$

3.3.3 The Observation Error Covariance Matrix R

The V_i term presented in Equation (3.43), the observation noise vector, is an additive random sequence with known statistics, i.e.,

$$E[v_i] = 0 \tag{3.91}$$

$$E[V_{it_k} V^T_{it_{k+1}}] = 0 \tag{3.92}$$

$$E[V_i \bar{X}] = 0 \tag{3.93}$$

$$E[V_i V_j^T] = \delta_{ij} R_i , \tag{3.94}$$

where δ_{ij} is the Kronecker delta and $R_i = \sigma^2_{wi}$ is a non-negative definite matrix. The R matrix, for non-differencing cases, is a diagonal matrix. If, however, the satellite system provides more than one satellite and a differencing scheme is used, the R matrix contains off-diagonal elements as well.

In our model the R matrix is defined by

$$R = \begin{bmatrix} (\sigma_{w\rho}/R_k)^2 & 0 \\ 0 & \sigma^2_{w\dot{\rho}} \end{bmatrix} . \tag{3.95}$$

3.4 Real Time Processing Considerations

Since Kalman mechanizations consist of sets of matrix operations, it can readily be seen that if the number of state variables in X is large, the matrix manipulations can become unwieldy. Hence, by judicious choice of state variables, the system dimension can be minimized; however, the degree of suboptimality can be constrained within limits. Also certain pseudo-states can be implemented to approximate the effects of noise in the system rather than explicitly modeling all noise functions as states; for example, in the form of

$$P_{k\ominus} = \Phi_{k,k-1} P_{k-1} \Phi^T_{k,k-1} + B_{k,k-1} Q_{k-1} B^T_{k,k-1} , \tag{3.96}$$

extrapolation of the noise terms is incorporated in the Q matrix through a B coefficient matrix. Obviously, such modeling eliminates the capability of estimating the subject noise processes. The inversion required in the Kalman gain matrix can be simplified if the R matrix is a diagonal matrix. Under such circumstances the inversion is performed serially (i.e., row by row); this reduces each row to a scalar (1 × 1) value or simply to a constant or fraction. If, however, the Z matrix has more than 3 rows, the serial processing becomes a time consuming process*. The extrapolation routine requires n × n matrix multiply operations, which consumes large computer memory. The collapsible-matrix-multiply routine†, based on element indexing, however, reduces memory requirements more than one half.

Each of the above simplifications aids in reducing the digital computer memory requirements for real time operation.

Still another scheme to aid in preserving accuracy in a fixed-point mechanization is a real time rescaling module. This can be used to keep the covariance and related computations optimally scaled by periodically checking the elements of the covariance matrix and shifting if necessary. A typical situation where a rescaling module would be useful is postulated below. Assume that unfavorable geometric conditions prevail and three successive satellite passes are missed; near the equator this would mean about six to eight hours without a position fix. If positions (latitude error and longitude error) are state variables and assuming a linear degradation of position, a ratio of 64 to 1 on the respective covariance matrix elements would result if the fix occurred at 8 hours instead of one hour. If the covariance matrix were scaled to accommodate the worst case condition, then, when the system is operating at optimum accuracy, significance is lost since the data are shifted to the right in the "register". Essentially the scheme described above is a "pseudo-floating-point" mechanization. A true floating point mechanization could be implemented if system processing time and memory constraints were not severe. Computational precision can still be a problem in floating point if sufficient data bits are not carried.

Another concept to consider is the effect of the finite word length of the on-board computer on the Kalman filter-estimator mechanization.

All these problems may be evaluated through use of a sophisticated simulation program which includes the following:

Simulation of filter/sensor mismatches and their effects on system navigation performance.

Simulation of the suboptimal filter and of its effects on system navigation performance (i.e., the effects of neglecting some of the state variables).

---

* The auto-invert routine is the optimal solution for up to 3 × 3 matrices. The concept may be extended to multiples of 3 or modified for 2 × 2 and its multiples.

† The collapsible-matrix-multiply routine has been simulated and used in numerous occasions in air/seaborne Kalman filter mechanizations.

Simulation of the finite word length of the on-board computer and its effects on system navigation performance.

The problems presented above may be rigorously analyzed through the use of the REAL-WORLD-COVARIANCE concept* and addition of word length masking instructions to each arithmetical statement.

4. DYNAMIC SIMULATION RESULTS

In order to evaluate the system performance, several configurations were simulated. The mission profiles and performance results are presented below.

4.1 Mission Profile

Two missions were simulated. The first mission assumed is described by a course originating near the equator which is consistent with about ninety minutes between satellite passes. The ship's speed is 10 knots at a heading of 45°. Pitching and rolling motions of the ship were assumed to be 3° and 5°, respectively.

A variable water profile was assumed to illustrate the degradation in water depths of 0-400 ft and 400-600 ft. Depths greater than 600 ft were not considered in the error analyses described in the following section because the sea current drifts swamp the remaining system error sources. However, under such conditions, a local sea current biasing can be realized by implementing corresponding states in the state vector.

The second mission originated at 30 degrees North latitude and 118 degrees West longitude. Ship's speed was initially zero and it accelerated to 5 knots while simultaneously turning from a heading of 309 degrees to 315 degrees. For the remainder of the five and one half hour mission, the ship remained on a straight course. The ship's dynamics are as described above.

4.2 Error Analysis Results

Basic assumptions relating to the system mechanization in the first mission simulated are as follows:

(i) Ship's attitude compensation is implemented. Inclinometer errors are about 0.3 degrees per axis.

(ii) Gyro motion compensation is included. A standard heading error of 0.5 degrees for the non-Kalman filtered mechanizations is shown. The same gyro damped with Kalman filter corrections results in a heading error of less than 0.15 degrees.

(iii) A one sigma position error of 300 ft/axis is assumed for the static satellite receiver subsystem error.

(iv) The sonar Doppler errors are a function of water depth and are described in Section 2.

(v) Satellite elevation angles at closest approach are between 20 and 60 degrees.

Figure 7 is a plot illustrating the radial position time history for three configurations in the first mission. The initial error is assumed to be zero. This is justifiable since the satellite yields position fixes which are zero-mean processes. Hence, by taking many fixes at the same static location, the position error in the limit approaches zero. In the first ninety minutes the radial error grows monotonically, at which point the satellite is assumed overhead. At this point the system is reset. Note that the Kalman filtered mechanization is reset to a lower value and, in the subsequent ninety minutes, the error growth is at a reduced rate relative to the non-filtered systems. This results from the cross-calibration of the navigation sensors upon receipt of satellite fix. As shown in Figure 7, the residual position error at each fix gets progressively better, from the increased knowledge of the system errors.

The error profile for the second mission is shown in Figure 8. These curves were generated using a dynamic simulation with the 48 state variables described in Section 3. The observation vectors employed were range and range-rate, as indicated on the figure. The range error is 170 feet per observation and the range-rate error is 0.1 ft/sec per observation. A gradual update is indicated, since a range and/or range-rate observation is made every one minute. Passes of variable duration and spacing were simulated to approximate a real-world environment. As cited above, the general improvements in accuracy are also present in this plot with one explainable exception – the fix at five hours. The residual for this fix is higher than the previous fix residual, due to the satellite/user geometry.

The dead reckoning error is larger for the second mission, due to the additional errors relating to the equipment that were implemented. For instance, frequency offset errors were assumed for the sonar Doppler and velocimeter. Scale factor errors, non-orthogonalities, mass unbalance and aniscelasticity were initialized to describe the gyro compass as well as gyro drifts. The initial heading error was set at 0.75 degrees and was corrected down to 0.08 degrees at the fix.

---

* The concept is to simulate an optimal extrapolator in parallel with the suboptimal one. The Kalman gain computed in the suboptimal filter is applied to the real-world (full-size) covariance matrix. Hence the scientist conducting the study has at his disposal two (i.e., real-world and system) covariance matrices. Obviously, the real-world covariance matrix computation is the sole indication of the suboptimally computed Kalman gain effects in real world.

REFERENCES

1. Pitman, G.R. — *Inertial Guidance*. John Wiley, New York, 1962, pp.455-460.

2. DeRusso, P.M. — *State Variables for Engineers*. John Wiley, New York, 1966, pp.364-366.

3. Ogata, K. — *State Space Analysis of Control Systems*. Prentice-Hall, New Jersey, 1967, p.315.

4. Leondes, C.T. et al. — *Advances in Control Systems*, Academic Press, 1966.

5. Chernoff, J. — *Utilization of Satellite Navigation Techniques in Oceanographic Operations*. Marine Sciences Instrumentation, Vol.4, Plenum Press, 1966.

6. Kalman, R.E. — *A New Approach to Linear Filtering and Prediction Problems*. Journal of Basic Engineering, Vol.82, No.1, March 1960.

7. Kershner, R.B. — *Present State of Navigation by Doppler Measurement from Near Earth Satellites*. APL Technical Digest, November-December, 1965.

8. Newton, R.R. — *The Navy Navigation Satellite System*, Space Research VII, North-Holland Publishing Company, Amsterdam, 1965.

9. Stansell, T.A. — *The Navy Navigation Satellite System: Description and Status*. Journal of the Institute of Navigation, Washington, DC, Vol.15, No.3, Fall, 1968.

TABLE I

The State Vector

Number	*Symbol*	*Name of Error State Variable*
1	δV_N	North velocity
2	δV_E	East velocity
3	$\delta\psi_G$	Gyro heading
4	$\delta\lambda$	Longitude
5	$\delta\Lambda$	Latitude
6	$\delta\eta$	Gyro tilt
7	$\delta\alpha$	Gyro gravity reference
8*	δc_V	Velocimeter
9	$\delta\varphi_C$	Roll bias
10*	$\delta\varphi_S$	Roll sinusoidal
11	$\delta\theta_C$	Pitch bias
12*	$\delta\theta_S$	Pitch sinusoidal
13*	$\delta\omega_\theta$	Pitch frequency
14*	$\delta\omega_\varphi$	Roll frequency
15	δV_{SC}	Sea current velocity
16	$\delta\psi_{SC}$	Sea current heading
17	$\delta\psi_T$	Gyro azimuth loop residual torquing
18	$\epsilon_{\psi B}$	Gyro bias A
19*	$\epsilon_{\psi N1}$	Gyro noise No. 1A
20*	$\epsilon_{\psi N2}$	Gyro noise No. 2A
21*	k_N	Scale factor A
22*	η_{NT}	Orthogonality A
23*	η_{NS}	Orthogonality A
24*	γ_{N1}	Mass unbalance A
25*	γ_{N2}	Mass unbalance A
26*	ξ_N	Anisoelasticity A
27*	$\epsilon_{\theta B}$	Gyro bias, T
28*	$\epsilon_{\theta N1}$	Gyro noise No. 1, T
29*	$\epsilon_{\theta N2}$	Gyro noise No. 2, T
30*	η_{TS}	Orthogonality, T
31*	γ_{T1}	Mass unbalance, T
32*	γ_{T2}	Mass unbalance, T
33*	ξ_T	Anisoelasticity, T
34*	k_T	Scale factor
35*	$\delta\dot{V}_S$	Forward velocity rate
36*	$\delta\dot{V}_C$	Starboard velocity rate
37	δV_D	Vertical velocity
38*	$\delta\dot{V}_D$	Vertical velocity rate
39	δf_t	Sonar transmitter frequency
40	$\delta\Sigma f_c$	Cross axis frequency sum
41	$\delta\Sigma f_f$	Along axis frequency sum
42*	M_N	Warm-up bias A
43*	M_T	Warm-up bias T
44	$\dot{\theta}_T$	Gyro tilt loop residual torquing
45	δh	altitude (geoidal)

Notes: 1. Element 44 is added for gyro model modification.

2. Asterisked numbers are not used in the sea-borne computer.

3. State variables due to satellite errors are added and are shown in Section 3.3.3.

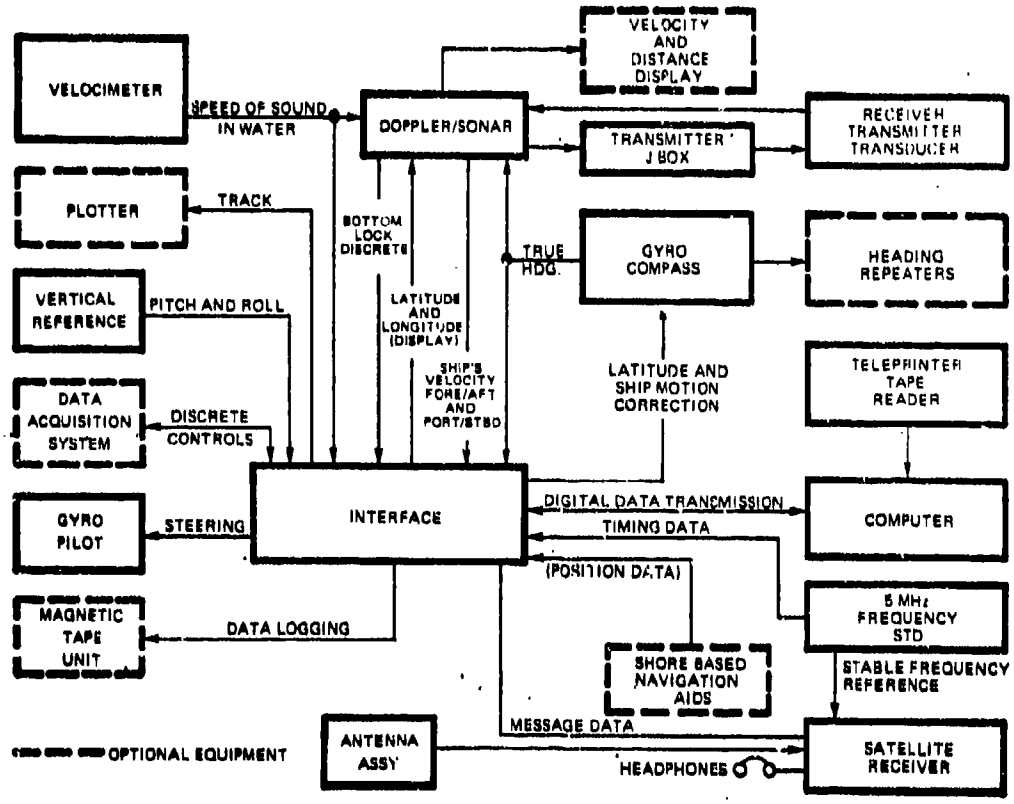


Fig. 1 Functional block diagram for a fully integrated system

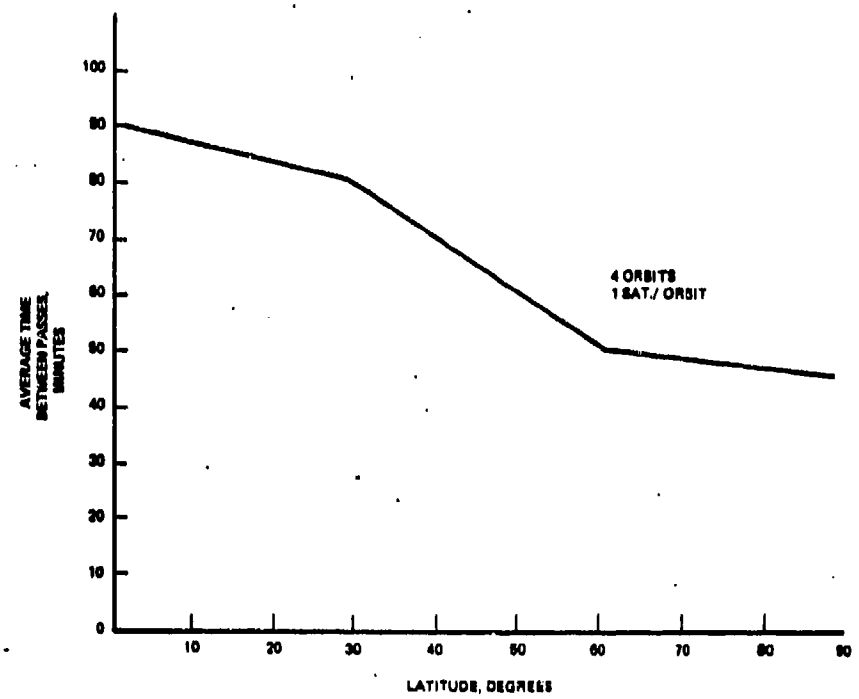


Fig. 2 Satellite availability chart

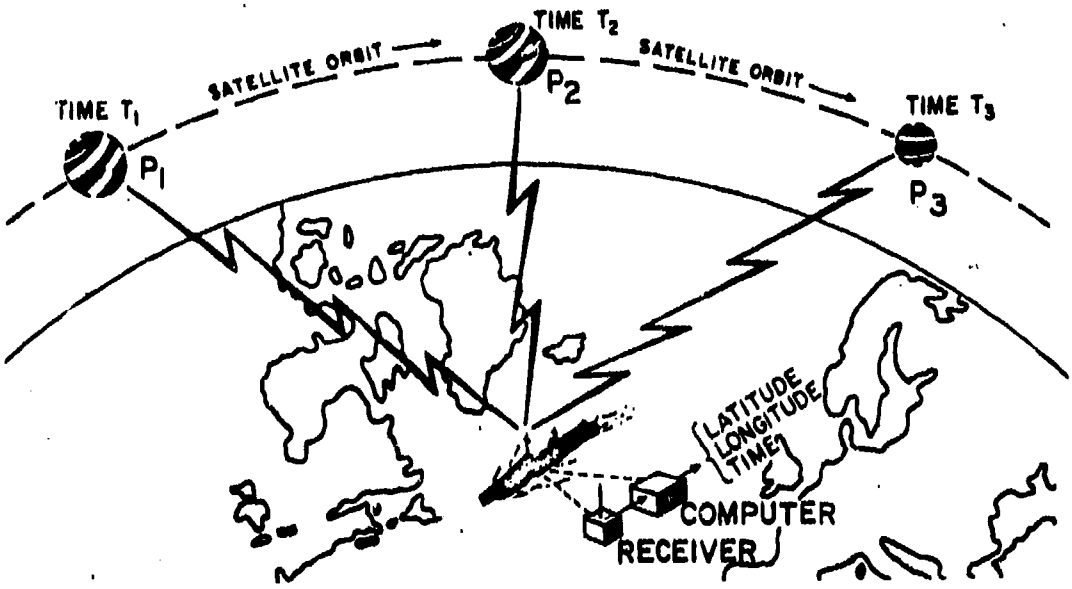


Fig. 3 Relative geometry of shipboard user and satellite

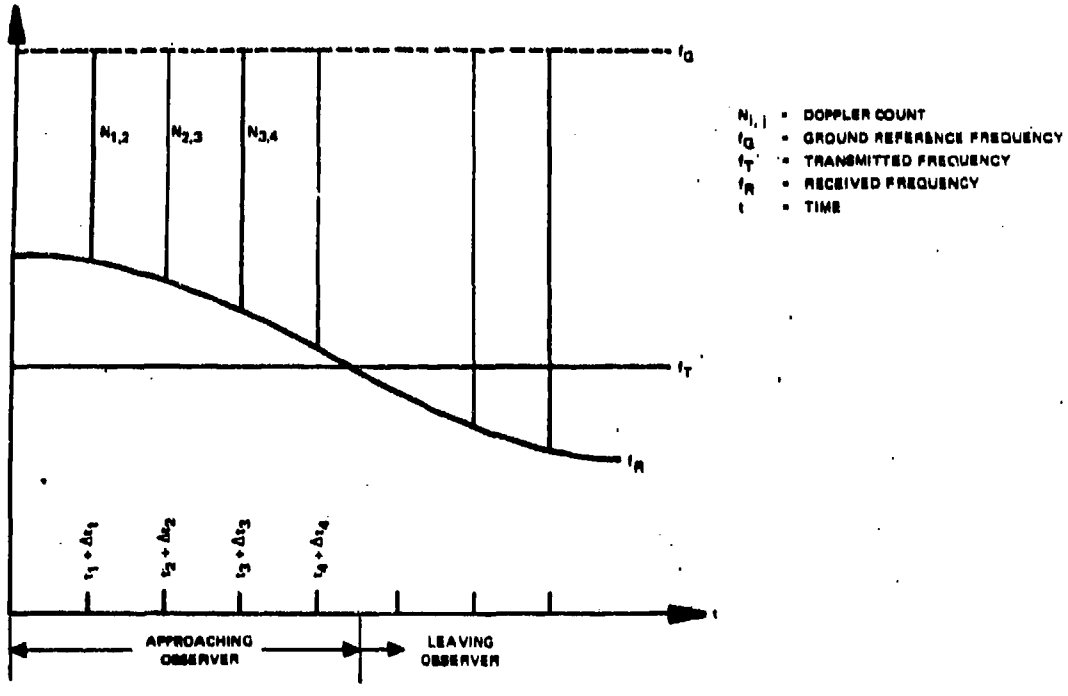


Fig. 4 Illustration of integrated Doppler count

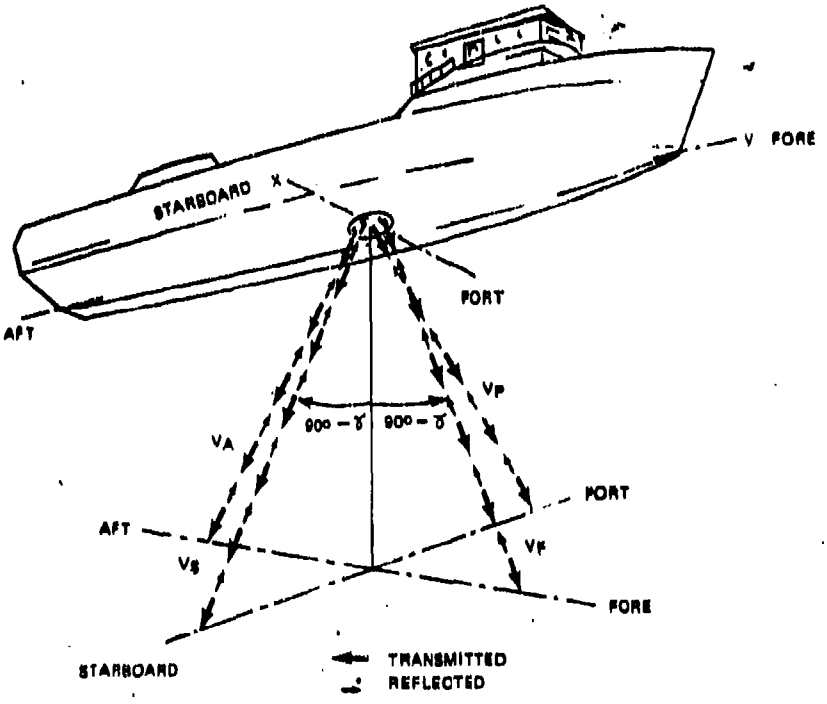
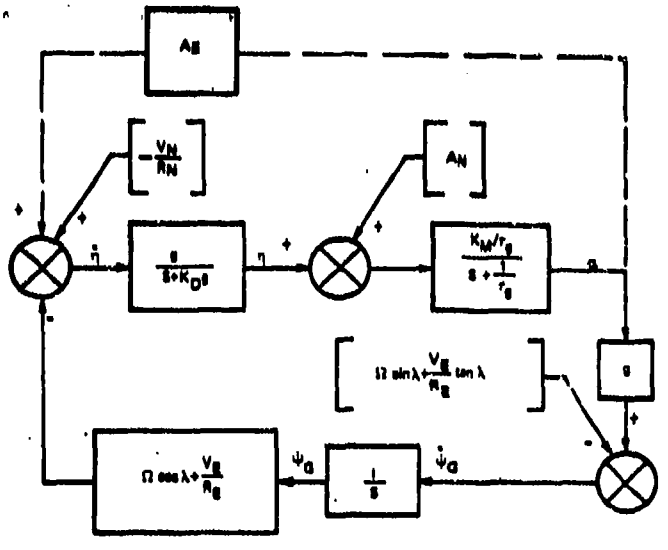


Fig.5 Basic Doppler navigation principle



$$\dot{\psi}_G = -(\Omega \sin\lambda + \frac{V_E}{R_E}\tan\lambda) + \omega_z$$

$$\eta = \left\{-\psi\left[\Omega \cos\lambda + \frac{V_E}{R_E}\right] + K_D\,(A_N + \omega_y) + \alpha A_E - \frac{V_N}{R_N}\right\}$$

$$\dot{\alpha} = \frac{K_M}{r_g}(A_N + \omega_x - \frac{\alpha}{r_g})$$

Fig.6 A typical gyrocompass mechanization block diagram. Note: Compensating terms are not included

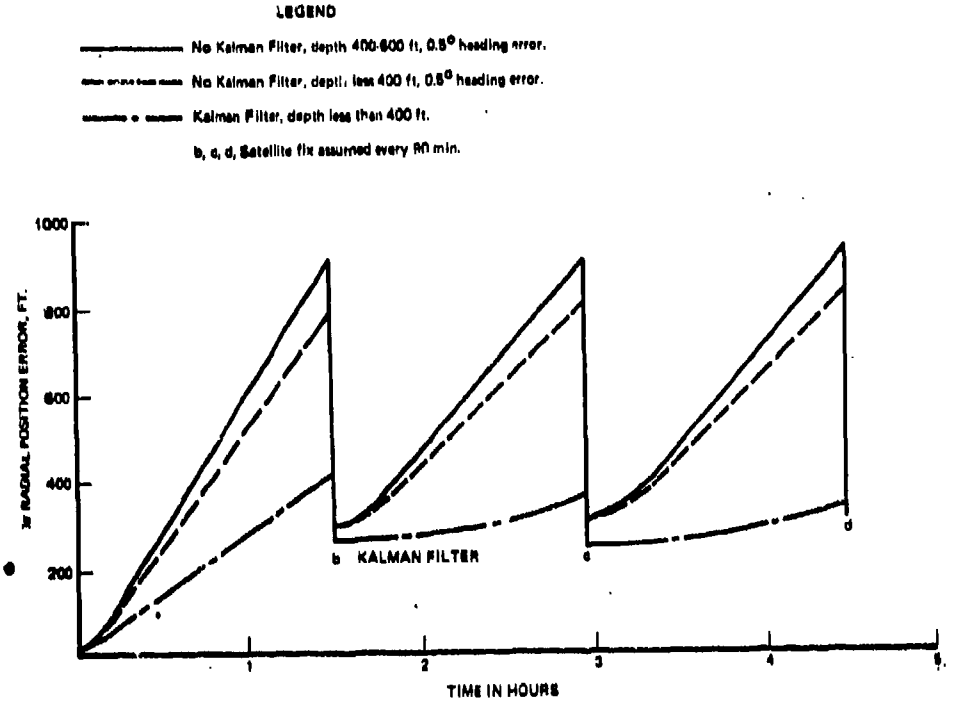


Fig.7 Satellite augmented marine navigation system performance for 10 knot ship's velocity and attitude compensation

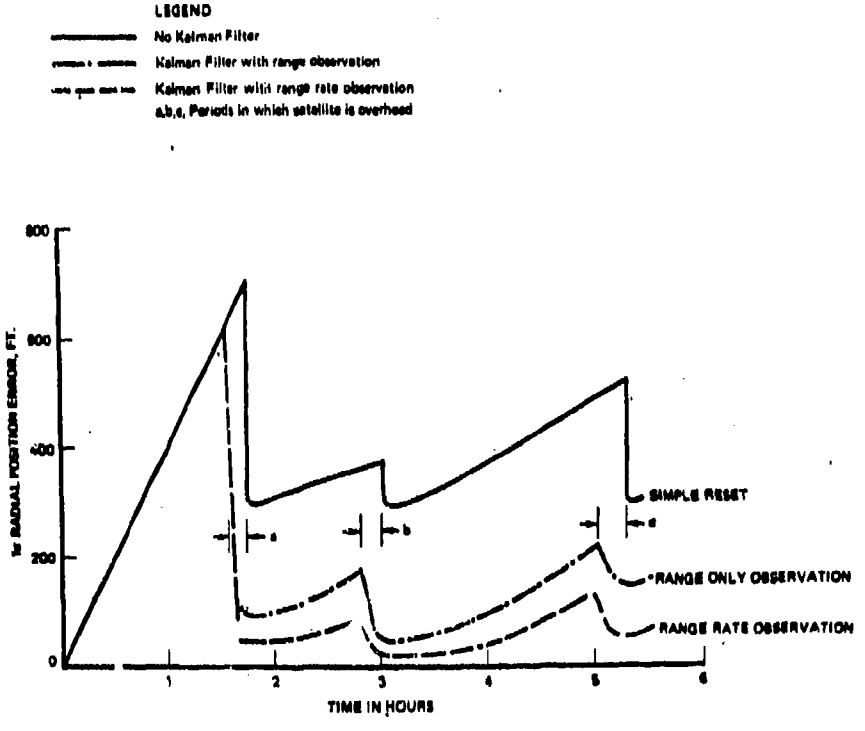


Fig.8 Dynamic simulation results for 5 knot ship's speed with attitude compensation

GPO 861-142

NATIONAL DISTRIBUTION CENTRES FOR AGARD PUBLICATIONS

AGARD publications are distributed to NATO Member Nations through the National Distribution Centres listed below.

BELGIUM	General J.TELMATE Coordinateur AGARD - V.S.L. Etat Major Forces Aériennes Caserne Prince Baudouin Place Dailly Bruxelles 3
CANADA	Director of Scientific Information Services Defence Research Board Department of National Defence - 'A' Building Ottawa, Ontario
DENMARK	Danish Defence Research Board Østerbrogades Kaserne Copenhagen Ø
FRANCE	O.N.E.R.A. (Direction) 29, Avenue de la Division Leclerc 92, Châtillon sous-Bagneux
GERMANY	Zentralstelle für Luftfahrtdokumentation und Information Maria-Theresia Str. 21 8 München 27 Attn: Dr Ing. M.J.RAUTENBERG
GREECE	Greek National Defence Staff D. JSS Athens
ICELAND	Director of Aviation Flugrad Reykjavik
ITALY	Ufficio del Delegato Nazional all' AGARD Ministero Difesa - Aeronautica Roma
LUXEMBOURG	Obtainable through BELGIUM
NETHERLANDS	Netherlands Delegation to AGARD National Aerospace Laboratory, NLR Attn: Mr A.H.GELLEKER P.O. Box 126 Delft
NORWAY	Norwegian Defense Research Establishment Main Library, c/o Mr P.L.EKERN P.O. Box 25 N-2007 Kjeller
PORTUGAL	Direccao do Servico de Material da Forca Aerea Rua da Escola Politecnica 42 Lisboa Attn. Brig. General Jose de Sousa OLIVEIRA
TURKEY	Turkish General Staff (ARGE) Ankara
UNITED KINGDOM	Ministry of Technology T.I.L. (F) Room 461 Block A, Station Square House St.Mary Cray Orpington, Kent
UNITED STATES	National Aeronautics and Space Administration (NASA) Langley Field, Virginia Attn: Report Distribution and Storage Unit

This AGARD publication has been issued with a limited number of copies, and distribution has been made through the AGARD distribution centres set up in the NATO countries, a list of which is given in the inside cover.

If spare copies are not available at these centres, microfiche or hard copy (printed facsimile, or reproduction from microcopy) may be purchased from:

Clearinghouse for Federal
Scientific and Technical
Information (CFSTI)
Springfield,
Virginia 22151, USA

ESRO/ELDO Space
Documentation Service
European Space Research
Organization
114, Avenue de Neuilly
92, Neuilly-sur-Seine, France

The request for microfiche or hard copy of an AGARD document should include the AGARD serial number when available, title, author (or editor) and publication date; if known, the NASA Accession Number should also be quoted.

The unit price for a microfiche provided by CFSTI is $0.65, and for a hard copy $3.00. Payment should be made by check and must accompany order. Remittances from foreign countries should be made by international money order or draft on an American bank, payable to CFSTI.

The unit cost for a microfiche provided by ESRO/ELDO is French francs 2.50 up to 60 pages, and for hard copy French francs 0.50 per page.

Full bibliographical references and abstracts of the newly issued AGARD publications are given in the following bi-monthly abstract journals with indexes:

Scientific and Technical Aerospace Reports (STAR) published by NASA, Scientific and Technical Information Facility, P O Box 33, College Park, Maryland 20740, USA.

and

National Aeronautics and Space Administration
WASHINGTON, D. C. 20546
OFFICIAL BUSINESS

POSTAGE AND FEES PAID
NATIONAL AERONAUTICS AND
SPACE ADMINISTRATION

100 001 [illegible] 00 3DS 70099 02677
DEFENSE DOCUMENTATION CENTER
SCIENTIFIC AND TECHNICAL INFORMATION
CAMERON STATION, BLDG. 5
ALEXANDRIA, VIRGINIA 22314
ATT: DDC RESEARCH LIBRARY

M B KHAN
STC' 21

[illegible]
Harford House, 7-9 Charlotte St. London. W1P 1HD