NONLINEAR STATE ESTIMATION
WITH APPLICATION TO COMMUNICATIONS
SATELLITES

By

© SANAA ABDEL-AZIM, B.Sc., M.Eng.

A Thesis
Submitted to the School of Graduate Studies
in Partial Fulfilment of the Requirements
for the Degree
Doctor of Philosophy

McMaster University
May 1979

NONLINEAR STATE ESTIMATION WITH APPLICATION TO COMMUNICATIONS SATELLITES

TO SAID AND DALIA

DOCTOR OF PHILOSOPHY (1979)
(Electrical Engineering)

McMASTER UNIVERSITY
Hamilton, Ontario

TITLE: Nonlinear State Estimation with Application to Communications Satellites

AUTHOR: Sanaa Abdel-Azim B.Sc. (Cairo University)
M.Eng. (McMaster University)

SUPERVISOR: N.K. Sinha
Professor of Electrical Engineering
B.Sc. Eng. (Banaras Hindu University)
Ph.D. (University of Manchester)

NUMBER OF PAGES: ix, 144

ABSTRACT

The problem of estimating the states of a dynamical system on the basis of output measurements is considered in detail. Some of the existing nonlinear estimation techniques are critically surveyed, these include the extended Kalman filter, the second-order filter, the innovations approach, and the invariant imbedding nonlinear filter. A new algorithm for nonlinear estimation is proposed which combines the invariant imbedding approach and the stochastic approximation algorithm for adaptively estimating the filter gain. The new algorithm is an iterative scheme which does not require knowledge of a priori input and measurement noise statistics. The proposed algorithm and the other techniques are used for the recursive state estimation of a satellite orbital trajectory. The results of simulation indicate the efficiency and reliability of the new algorithm. Convergence to the true state is achieved with much less computation when compared to the other methods.

ACKNOWLEDGEMENTS

With deep gratitude I wish to thank Professor N.K. Sinha for his guidance and encouragement during the preparation of this thesis.

I am also sincerely thankful to Mr. S. Altman and Dr. R. Mamen for providing the data and for their helpful suggestions and discussions.

Special thanks are due to my husband and my daughter for their patience and understanding.

I would like to acknowledge with thanks the useful discussions I had with H. El-Sherief, M. Rizk, Z. El-Razaz and M. El-Kady.

Financial support provided by the Communications Research Centre through Contract Number OSU4-0078 and by the Department of Electrical Engineering at McMaster University is gratefully acknowledged.

Special thanks are due to Miss P. Dillon of the Word Processing Centre for her invaluable help in typing this thesis.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The sequential estimation of the states of noisy nonlinear dynamical systems is of interest not only in automatic control but also in other areas of engineering where the system identification problem requires the processing of large quantities of data.

The problem of optimal state estimation in nonlinear dynamical systems has received considerable attention during the past few years. Since Wiener's pioneering work [1] on the theory of optimal filtering and prediction, many extensions and new developments have been made in the area of nonlinear filtering. Algorithms for filtering, smoothing and prediction of the states of a linear dynamical system have been derived by Kalman and Bucy [2,3]. These algorithms were extended to the estimation of states of nonlinear dynamical systems and the so-called first-order filter, or extended Kalman filter, was derived by several investigators including Cox [4], Mowery [5], and Ho and Lee [6]. The techniques utilized for generating these algorithms are based on orthogonal projection theory and maximum likelihood estimate. Most of these techniques employ a Taylor series expansion, neglect second and higher-order terms, and use linearized equations to compute the conditional error covariance matrix and filter gains.

From the point of view of the determination of the exact equations satisfied by the conditional probability density functions and conditional expectations, Kushner [7] indicated that the optimal filter can be realized by an infinite dimensional system using the stochastic Itô calculus. Bucy [8], Bass et al. [9], Denham and Pines [10] and many others have derived the filtering algorithms, by retaining second-order terms, using the probabilistic approach. Athans et al. [11] and Wishner et al. [12] have compared the performance of first-order and second-order filters, and showed that considerable improvement in the performance was achieved using the second-order filter. Detchmendy and Sridhar [13] and Kagiwada et al. [14] have derived filtering algorithms similar to the first-order filter for nonlinear esitmation problems using the least squares errors criterion and the invariant imbedding technique.

One of the most recent techniques for solving nonlinear filtering problems, based on the innovations approach, was presented by Kailath [15] and Kailath and Frost [16]. The use of the innovations allows us to obtain formulas and simple derivations that are remarkably similar to those used for the linear case. This technique, therefore, distinguishes clearly the essential points in which the nonlinear problem differs from the linear one.

The main objective of this research is to develop an efficient algorithm for optimal nonlinear estimation. The principal obstacle to a complete derivation of the filter algorithm for nonlinear systems lies in the computation of the gain matrix. An adaptive scheme could be used which combines both the innovations property of an optimal filter and stochastic approximation. It is proposed to develop an algorithm combining the invariant imbedding approach with the stochastic approximation adaptive scheme for gain computation and state estimation.

One of the important applications of the estimation theory is the determination of satellite orbital trajectories. The proposed algorithm is used to estimate the orbital states of a communications satellite using actual tracking data supplied by the Spacecraft Mechanics Division of the Communications Research Centre in Ottawa.

The thesis is divided into two main parts. Chapters 2 and 3 are mainly tutorial in nature, whereas the main contributions and results of the work appear in Chapters 4, 5 and 6.

In Chapter 2, a brief critical survey of different linear estimation techniques is outlined. The work of

Kolmogorov [17], Wiener [1], Kalman [2] and Kalman and Bucy [3] are summarized. Adaptive state estimation and the use of the innovations approach in linear filtering are also briefly reviewed.

Chapter 3 gives a critical survey of three of the existing nonlinear estimation techniques. The methods reviewed are the extended Kalman filter, the second-order nonlinear filter and the innovations approach to nonlinear estimation.

Chapter 4 includes the principal theoretical contribution of this work. The invariant imbedding concept is described, a nonlinear estimation algorithm based on invariant imbedding is presented. A stochastic approximation algorithm for adaptive gain computation is combined with the previous algorithm to provide a new nonlinear estimation algorithm.

Chapter 5 is devoted to the study of the orbital trajectory state model proposed by Altman [18]. An observation model is derived to be used with the above state model for state estimation.

Chapter 6 presents the results of the model simulation on a CDC 6400 computer. The problem of satellite orbit determination is solved for two different cases of initial orbits. Three of the existing nonlinear estimation algorithms are used and compared with the new algorithm proposed by the author.

Chapter 7 includes some concluding remarks regarding the validity and efficiency of the proposed algorithm, based on the results presented in the previous chapter. Limitations of the method and suggestions for further research are also given in this chapter.

CHAPTER 2

OPTIMAL ESTIMATION FOR LINEAR SYSTEMS

2.1 Introduction

Physical systems are designed and built to perform certain defined functions. In order to determine whether a system is performing properly, and ultimately to control the system performance, an engineer must know the state of the system at every instant of time. Physical systems are often subject to random disturbances, so that the system state may itself be random.

In order to determine the state of the system, the engineer takes measurements or observations on this system. These measurements are generally contaminated with noise caused by the electronic and mechanical components of the measuring devices.

We shall be concerned with the problem of estimating the state of a system from noisy measurements of the output of the system. This chapter presents a critical survey of existing estimation techniques for linear systems.

Information theorists and communication engineers have been more familiar with problems in which covariance information is given about signal and noise, usually called Wiener filtering problems [1]. On the other hand, control engineers deal more often with problems where the signal and the noise are described by state-space models [2,19].

In 1961 Kalman and Bucy presented a new approach to the linear filtering problem [3]. The novelty of their formulation was the representation of all random processes by correlation functions. By restricting their attention to Gauss-Markov processes in particular, they derived a set of differential equations for the estimates. The main advantage of their approach is that it is much easier to solve a set of differential equations by analog or digital techniques than to solve an integral equation and perform a convolution.

A well-known limitation of the application of the Kalman-Bucy filter to real world problems is the assumption of known a priori statistics for the stochastic errors in both the state and observation processes. This approach leads to a nonadaptive filter and although the performance may be satisfactory over some global operating region, it will be inferior to that obtained when a priori statistics are known locally as a function of time. Therefore, in the presence of unknown system disturbances it may be desirable to adaptively estimate the a priori statistics simultaneously with the system state. There exist different schemes for adaptive Kalman filtering, an adaptive sequential estimator derived by Sage and Husa [20] is presented in this chapter.

The innovations approach has provided an elegant state estimation technique by replacing the observations by the innovations process and thus sidestepping the difficulties with the solution of the Wiener-Hopf equation. This approach also provides a convenient framework for adaptive improvement of Kalman gains. Initial guesses of the noise parameters are improved iteratively until the innovations process is sufficiently white. Problems of this nature have been treated by Mehra [21].

2.2 Kolmogorov-Wiener Filter

The optimum linear filtering and prediction problem first solved by Kolmogorov [17] and Wiener [1] marked the beginning of engineering awareness of the problem of state estimation. Unfortunately, because the results of this work were expressed in the frequency domain, they could not be directly extended to nonstationary problems. Although the general formulation of the nonstationary problem can also be developed, through the Wiener-Hopf equations, very few practical results were achieved. It was not until the development of the Kalman filter algorithm that the computational difficulties were overcome for the general nonstationary case.

The problem may be stated as follows: given the random process

$$y(t) = x(t) + n(t) \tag{2.1}$$

where $x(t)$ is the useful signal imbedded in the noise $n(t)$ and both are assumed to be random processes, determine a filter such that its output $\hat{x}(t)$ will be the best approximation to $x(t)$ in the mean square sense. That is, minimize $E[e^2(t)]$, where

$$e(t) \triangleq x(t) - \hat{x}(t) \tag{2.2}$$

and $E[\cdot]$ is the expectation operator.

Using variational arguments, Wiener showed that the impulse response $h(t)$ of the optimum linear filter satisfies the following Wiener-Hopf integral equation

$$R_{xy}(\tau) - \int_{-\infty}^{\infty} R_y(\tau-\sigma)\ h(\sigma)\ d\sigma = 0 \tag{2.3}$$

where R_{xy} is the cross-correlation function between $x(t)$ and $y(t)$, and R_y the autocorrelation of $y(t)$, defined as

$$\begin{aligned} R_{xy}(\tau) &\triangleq E\ [x(t)\ y(t-\tau)] \\ R_y(\tau) &\triangleq E\ [y(t)\ y(t-\tau)] \end{aligned} \tag{2.4}$$

The determination of the optimum filter transfer function requires knowledge of the correlation functions or the corresponding spectral densities as well as performing spectral factorization [22]. The formulation is cumbersome to implement and requires more computation as compared to the state-space formulation.

2.3 Kalman Filtering

In his work Kalman formulated and solved the Wiener problem using the state-space formulation. The classical filtering and prediction problem was reexamined using ~~the~~ Bode-Shannon representation of random processes [23] and the "state-transition" method of analysis of dynamic systems. The new results of this approach were:

(1) The formulation and methods of solution of the problem apply without modification to stationary and nonstationary statistics and to infinite memory filters.

(2) A nonlinear difference equation is derived for the covariance matrix of the optimal estimation error. From the solution of this equation the coefficients of the difference equation of the optimal linear filter are obtained without further calculations. This equation is the discrete-time counterpart of the Ricatti differential equation encountered in optimal

control [24].

(3) The filtering problem can be shown to be the dual of the noise-free regulator problem [2].

2.3.1 Discrete-Time Kalman Filter

The discrete-time Kalman filter results were actually the first to be obtained [2], partly because the major system-theory activity in the mid-fifties was in the field of sampled-data systems, which arose when modern digital computers were put into control and communication links. In his work Kalman combined state-space description and the concept of orthogonal projections to give a complete and elegant solution.

Consider the message model described by the linear vector difference equation

$$x(k) = \Phi(k,k-1)\ x(k-1) + \Gamma(k)\ w(k) \tag{2.5}$$

where the input noise w is a zero-mean white-noise process, with covariance

$$\mathrm{cov}\ \{w(k),\ w(j)\} = Q(k)\ \delta_K(k-j) \tag{2.6}$$

and δ_K is the Kronecker delta function.

The observation (measurement) model is given by the linear algebraic relation

$$y(k) = H(k)\ x(k) + v(k) \tag{2.7}$$

where the measurement noise v is a zero-mean white-noise process, with covariance

$$\text{cov}\ \{v(k),\ v(j)\} = R(k)\ \delta_K(k-j) \tag{2.8}$$

The input and observation noise are assumed to be uncorrelated, so that

$$\text{cov}\ \{w(k),\ v(j)\} = 0 \qquad \text{for all } j,k \tag{2.9}$$

The initial value of x is a random variable with mean x_o and variance P_o.

Based on a set of sequential observations defined by $Y(k) = \{y(1),\ y(2),\ \ldots,\ y(k)\}$, we wish to determine an estimate $\hat{x}(k|k)$ of $x(k)$. The estimation error will be denoted by

$$e(k|k) \triangleq x(k) - \hat{x}(k|k) \tag{2.10}$$

Of the many possible unbiased linear estimators, we wish to select the one which gives the minimum error variance. Kalman's original derivation of the optimum linear minimum-error-variance estimator used the orthogonal projection approach [25].

As shown by Kalman, the optimal estimate $\hat{x}(k|k)$ of the state vector $x(k)$ may be interpreted geometrically as the projection of $x(k)$ onto $Y(k-1)$. Algebraically, the estimates are given by

$$\hat{x}(k|k) \triangleq \text{Proj}\,\{x(k) \mid Y(k)\} \tag{2.11}$$

After several mathematical manipulations [2] equation (2.11) becomes

$$\hat{x}(k|k) = \Phi(k,k-1)\,\hat{x}(k-1|k-1) + K(k)\,\tilde{y}(k) \tag{2.12}$$

where $K(k)$ is a weighting of observation for optimal estimation, called the Kalman gain matrix. The vector $\tilde{y}(k)$ is defined as

$$\tilde{y}(k) = y(k) - H(k)\,\hat{x}(k|k-1) \tag{2.13}$$

where $\quad \hat{x}(k|k-1) = \phi(k,\, k-1)\,\hat{x}(k-1|k-1)$

The optimal linear Kalman estimator is generated by the following recursive algorithms.

<u>Filter algorithm</u>

$$\hat{x}(k|k) = \Phi(k,k-1)\,\hat{x}(k|k-1) + K(k)\,\tilde{y}(k) \tag{2.14}$$

<u>Gain algorithm</u>

$$K(k) = P(k|k-1)\,H^T(k)\,[H(k)\,P(k|k-1)\,H^T(k) + R(k)]^{-1} \tag{2.15}$$

Prediction error variance algorithm

$$P(k|k-1) = \Phi(k,k-1)\, P(k-1)\, \Phi^T(k,k-1) + \Gamma(k-1)\, Q(k-1)\, \Gamma^T(k-1) \tag{2.16}$$

Error variance algorithm

$$P(k) = [I - K(k)\, H(k)]\, P(k|k-1) \tag{2.17}$$

Initial conditions

$$\hat{x}(0) = x_o, \qquad P(0|0) = P_o \tag{2.18}$$

A block diagram of the discrete Kalman filter is given in Fig. 2.1.

The great merit in the Kalman algorithms for the filtering problem lies in the fact that the solution directly specifies a practical implementation of the results. In many practical problems, real-time computation is realizable. An additional feature of the approach is that the error variance $P(k)$ is computed as a direct part of the estimator, and may be used to judge the accuracy of the estimation procedure. The Kalman filter is an unbiased estimator, assuming that the correct prior statistics of the plant and measurement noises and the initial state are used to implement the algorithm. In his original work Kalman proves the stability of the filter under certain conditions [26].

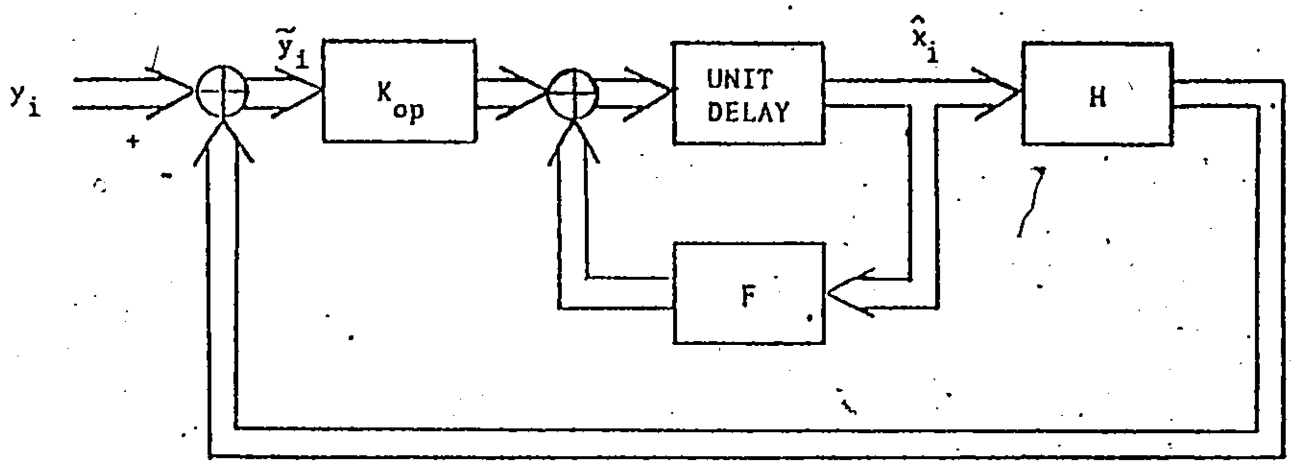


Figure 2.1

Block diagram of discrete-time Kalman filter

Different approaches are used to derive the discrete Kalman filter given above, e.g., least squares method [27], maximum likelihood approach [28] and the probabilistic approach [29].

2.3.2 Continuous-Time Kalman-Bucy Filter

Kalman considered the continuous-time problem as the limiting case of the discrete-time problem as the sampling interval is reduced to zero [19]. Bucy's important contribution to the joint paper [3] was a derivation using the finite-time Wiener-Hopf equation (2.3).

The message model for the continuous problem takes the form of a first-order differential equation,

$$\dot{x}(t) = F(t)\ x(t) + G(t)\ w(t) \tag{2.19}$$

where w(t) is a zero-mean white-noise process, with

$$\mathrm{cov}\ \{w(t),\ w(\tau)\} \triangleq Q(t)\ \delta_D(t-\tau) \tag{2.20}$$

where δ_D is the Dirac delta function.

The initial mean and variance of the state x(t) are known as

$$E[x(0)] = x_o, \qquad \mathrm{var}\ [x(0)] = P_o \tag{2.21}$$

The observation model is given as

$$y(t) = H(t)\ x(t) + v(t) \tag{2.22}$$

The measurement noise is white and of zero mean, with

$$\mathrm{cov}\ \{v(t),\ v(\tau)\} \triangleq R(t)\ \delta_D(t-\tau) \tag{2.23}$$

It is assumed that the input and measurement noises are uncorrelated.

We are interested in determining the linear minimum-error-variance sequential estimator of the state $x(t)$, given the measurements $y(\tau)$ for $0 \leq \tau \leq t$, that is, $Y(t)$.

Using the limiting approach or the Wiener-Hopf equation, the following continuous-Kalman filter is derived.

Filter algorithm

$$\dot{\hat{x}}(t) = F(t)\ \hat{x}(t) + K(t)\ [y(t) - H(t)\ \hat{x}(t)] \tag{2.24}$$

Gain algorithm

$$K(t) = P(t)\ H^T(t)\ R^{-1}(t) \tag{2.25}$$

Error-variance algorithm

$$\begin{aligned}\dot{P}(t) = {} & F(t)\ P(t) + P(t)\ F^T(t) - P(t)\ H^T(t)\ R^{-1}(t)\ H(t)\ P(t) \\ & + G(t)^T Q(t)\ G\ (t)\end{aligned} \tag{2.26}$$

A special case of the continuous Kalman filter, the stationary Kalman filter, should satisfy the following assumptions.

(1) The message and observation models are time-invariant

$$\dot{x}(t) = F\,x(t) + G\,w(t)$$
$$y(t) = H\,x(t) + v(t) \tag{2.27}$$

where F, G and H are constant matrices.

(2) The input and measurement noises are at least wide-sense stationary.

When the stationary problem is formulated in the frequency domain, the Wiener filter is obtained. So the stationary Kalman filter and the Wiener filter are two different means of solving the stationary estimation problem. The Kalman filter is expressed in the time-domain and state-variable notation, whereas the Wiener filter is expressed in frequency-domain transfer-function notation. Obviously there are significant computational advantages of the Kalman formulation for most control problems.

Kalman has often stressed that the major contribution of this work is the proof that under certain technical conditions called "controllability" and "observability" [30]:

(1) The optimum filter is "stable", in the sense that the effects of initial errors and round-off and other

computational errors will die out asymptotically.

(2) Every solution of the variance equation (2.17) or (2.26) starting at a symmetric, non-negative matrix P_o converges to a unique steady-state solution, independent of the initial condition and of errors introduced during the computation.

However, in some practical applications, one finds that the actual estimation errors greatly exceed the values which would be theoretically predicted by the error variance. This phenomenon, referred to as divergence, causes instability of the filter algorithm even though, theoretically, it is computationally stable.

One way to eliminate divergence is to use adaptive filtering algorithms [31,32,33].

2.4 Adaptive Estimation

Two problems which occur when methods of optimal estimation are applied to an actual problem are the choice of the prior statistics and the choice of a mathematical model that represents the system.

The model must be sufficiently complete for an adequate description of the system and also sufficiently simple such that the resulting algorithms are computationally feasible.

Typically, the a priori statistics are selected through analysis of empirical data or computer simulation and are assumed constant. This leads to a nonadaptive filter, and although estimation performance may be satisfactory over some global operating regime, it will be inferior to that obtained when the a priori statistics are known locally as a function of time. Therefore, in the presence of unknown system disturbances, it is desirable to adaptively estimate the a priori statistics simultaneously with the system state.

Mehra [21] has proposed an algorithm for adaptive Kalman filtering based on first obtaining estimates of the noise covariance matrices which are asymptotically unbiased and consistent. Another algorithm has been proposed by Carew and Belanger [33] which estimates the optimum gain matrix by utilizing the autocorrelation function of the sub-optimum innovations sequence generated by using a suboptimum gain matrix. However, both of these algorithms suffer from the drawback that considerable data must be accumulated before they can be used. Furthermore, their approach may be called "open-loop", and will not be effective if the noise covariance is varying slowly.

A "closed-loop" adaptive approach to estimating the optimal gain matrix, based on stochastic approximation has been proposed by Scharf and Alspach [34], Sinha [35], Sinha and Mukherjee [36] and others. This approach will be discussed in detail in Chapter 4, and used for nonlinear estimation.

Sage and Husa derived an adaptive filtering algorithm which provides sequential estimation of the first two moments of the plant and the measurement noise [20]. This algorithm is suitable for on-line application and also nonlinear estimation problem when combined with the extended Kalman filter.

Sage and Husa consider the discrete message and observations models given by equations (2.5) and (2.7). Let $r(k)$, $R(k)$ be the mean and variance of the measurement noise, $q(k)$ and $Q(k)$ be the mean and variance of the plant noise. In their work Sage and Husa developed a suboptimal adaptive Bayes estimation algorithms that estimate the states as well as the noise statistics.

The following is a summary of the filter for discrete systems, the continuous case could be obtained by a limiting argument.

Message model

$$x(k) = \Phi(k,k-1)\ x(k-1) + \Gamma\ w(k) \tag{2.28}$$

Observation model

$$y(k) = H(k)\ x(k) + v(k) \tag{2.29}$$

Filter algorithm

$$\hat{x}(k|k) = \hat{x}(k|k-1) + K(k)\ [y(k) - \hat{r}(k) - H(k)\ \hat{x}(k|k-1)] \tag{2.30}$$

Prediction algorithm

$$\hat{x}(k|k-1) = \Phi(k,k-1)\ \hat{x}(k-1|k-1) + \Gamma\ \hat{q}(k-1) \tag{2.31}$$

Filter gain algorithm

$$K(k) = P(k|k-1)\ H^T(k)\ [H(k)\ P(k|k-1)\ H^T(k) + \hat{R}(k)]^{-1} \tag{2.32}$$

Prediction error variance algorithm

$$P(k|k-1) = \Phi(k,k-1)\ P(k-1|k-1)\ \Phi^T(k,k-1) + \Gamma\ \hat{Q}(k-1)\ \Gamma^T \tag{2.33}$$

Error variance algorithm

$$P(k|k) = [I - K(k)\ H(k)]\ P(k|k-1) \tag{2.34}$$

Measurement noise mean algorithm

$$\hat{r}(k) = \frac{1}{k-1}\ [(k-2)\ \hat{r}(k-1) + y(k-1) - H(k-1)\ \hat{x}(k|k-1)]. \tag{2.35}$$

Measurement noise variance algorithm

$$\hat{R}(k) = \frac{1}{k-1}\ [(k-2)\ \hat{R}(k-1) + \tilde{y}(k-1)\ \tilde{y}^T(k-1) - H(k-1) \times P(k-1|k-1)\ H^T(k-1)] \tag{2.36}$$

Plant noise mean algorithm

$$\Gamma \hat{q}(k) = \frac{1}{k} [(k-1) \Gamma \hat{q}(k-1) + \hat{x}(k|k) - \Phi(k,k-1) \hat{x}(k-1|k-1)] \tag{2.37}$$

Plant noise variance algorithm

$$\Gamma \hat{Q}(k) \Gamma^T = \frac{1}{k} [(k-1) \Gamma \hat{Q}(k-1) \Gamma^T + K(k) \tilde{y}(k) \tilde{y}^T(k) K^T(k) + P(k|k) - \Phi(k,k-1) P(k-1|k-1) \Phi^T(k,k-1)] \tag{2.38}$$

Measurement error algorithm

$$\tilde{y}(k) = y(k) - H(k) \hat{x}(k|k-1) - \hat{r}(k|k-1) \tag{2.39}$$

Initial conditions

$$\begin{aligned} &\hat{x}(0) = x_o, \quad P(0|0) = P_o, \\ &\hat{r}(0) = r_o, \quad \hat{R}(0) = R_o, \\ &\hat{q}(0) = q_o, \quad \hat{Q}(0) = Q_o \end{aligned} \tag{2.40}$$

The continuous version of the algorithm is given in reference [20] in detail.

This algorithm has been used by the author for orbit determination of communications satellite.

2.5 Innovations Approach to Linear Estimation

Kailath [15,16] has provided considerable insight into the problem of optimum filtering and estimation by proposing and proving the "Innovations Theorem". Using this theorem Kailath has been able not only to derive the filtering equations of Kalman and Bucy in a simpler and more elegant manner, but also to extend the concepts to certain nonlinear problems.

The observations process is given in the form:

$$y(t) = z(t) + v(t), \qquad 0 \leq t \leq T \tag{2.41}$$

where $z(\cdot)$ and $v(\cdot)$ are statistically independent vector processes; the process $v(\cdot)$ is a white Gaussian noise with

$$E\,\{v(t)\;v^T(\tau)] = R(t)\;\delta_D(t-\tau) \tag{2.42}$$

and the signal process $z(t)$, which is not necessarily Gaussian, has the properties

$$E\{z(t)\} = 0, \qquad \int_0^T E\,\{z(t)\;z^T(t)\}\,dt < \infty \tag{2.43}$$

Let $x(\cdot)$ be a random process related to the signal process $z(\cdot)$ and obeying

$$E\{x(t)\} = 0, \qquad \int_0^T E\,\{x(t)\;x^T(t)\}\,dt < \infty \tag{2.44}$$

The usual assumption is that $z(\cdot)$ is a function of past and present $x(\cdot)$, say

$$z(t) = H(t)\ x(t) \tag{2.45}$$

Under the above assumptions we wish to find the least-squares estimate of $x(t)$, $\hat{x}(t|t)$, given the observations $\{y(\tau),\ 0 \leq \tau < t \leq T\}$. It is well known that $\hat{x}(t|t)$ is given by the conditional mean, [37]

$$\hat{x}(t|t) = E\ \{x(t)\ |\ y(\tau),\ 0 \leq \tau < t\} \tag{2.46}$$

In his work Kailath shows that the observation process $y(\cdot)$ can be transformed into a white Gaussian process $\xi(\cdot)$, which is called the innovations process and formulas for the estimate $\hat{x}(t|t)$ are then readily obtained. Under certain assumptions [15], it is shown that the innovations process and the observation process are equivalent in the sense that there exists a causal and causally invertible transformation from $y(\cdot)$ to $\xi(\cdot)$ such that

$$\begin{aligned} \hat{x}(t|t) &\triangleq E\ \{x(t)\ |\ y(\tau)\quad 0 \leq \tau < t\} \\ &= E\ \{x(t)\ |\ \xi(\tau)\quad 0 \leq \tau < t\} \end{aligned} \tag{2.47}$$

Using the innovations theorem [Appendix A], which states that $\xi(t)$ is a white Gaussian noise with the same covariance as the observation error $v(\cdot)$, and the projection theorem for least-squares estimates [37], which states that the estimation error is uncorrelated with the observation, yields

$$\hat{x}(t|t) = \int_0^t h(t,\tau)\ \xi(\tau)\ d\tau \tag{2.48}$$

where

$$\xi(t) = y(t) - \hat{z}(t|t) \tag{2.49}$$

and

$$\hat{z}(t) = H(t)\ \hat{x}(t) \tag{2.50}$$

The innovations process can be considered as the new information in the observations. From the previously mentioned orthogonal-projection theorem the estimation error $\tilde{x}(t)$ and the observation sequence $y(t)$ are orthogonal. Because $\xi(t)$ and $y(t)$ are equivalent, it is clear that $\tilde{x}(t)$ and $\xi(t)$ must also be orthogonal, then

$$\text{cov}\ \{\tilde{x}(t),\ \xi(\sigma)\} = 0, \qquad 0 \leq \sigma < t \tag{2.51}$$

we have, from equation (2.51)

$$\begin{aligned} \text{cov}\ \{\tilde{x}(t),\ \xi(\sigma)\} &= \text{cov}\ \{x(t) - \hat{x}(t|t),\ \xi(\sigma)\} \\ &= \text{cov}\ \{x(t),\ \xi(\sigma)\} - \int_0^t h(t,\tau) \\ &\quad \text{cov}\ \{\xi(\tau),\ \xi(\sigma)\}\ d\tau = 0 \end{aligned} \tag{2.52}$$

so that

$$\mathrm{cov}\ \{x(t),\ \xi(\sigma)\} = \int_0^t h(t,\tau)\ \mathrm{cov}\ \{\xi(\tau),\ \xi(\sigma)\}\ d\tau \tag{2.53}$$

This is the Wiener-Hopf equation (2.3), except that, in this case, because the innovations process is white, with known variance R (same as the observation error v), we find that

$$\mathrm{cov}\ \{x(t),\ \xi(\sigma)\} = \int_0^t h(t,\tau)\ R(\tau)\ \delta\ (\tau-\sigma)\ d\tau \tag{2.54}$$

which, because of the Dirac delta function, becomes

$$\mathrm{cov}\ \{x(t),\ \xi(\sigma)\} = h(t,\tau)\ R(\sigma) \tag{2.55}$$

Hence the Wiener-Hopf integral equation for $E\{\cdot\}$ has been changed to a simple algebraic expression, which gives

$$h(t,\tau) = \mathrm{cov}\ \{x(t),\ \xi(\tau)\}\ R^{-1}(\tau) \tag{2.56}$$

The estimate $\hat{x}(t|t)$ is therefore

$$\hat{x}(t|t) = \int_0^t \mathrm{cov}\ \{x(t),\ \xi(\tau)\}\ R^{-1}(\tau)\ \xi(\tau)\ d\tau \tag{2.57}$$

If we take the derivative with respect to t on both sides of equation (2.57), and after several manipulations, we obtain

$$\hat{x}(t|t) = F(t)\,\hat{x}(t) + \text{cov}\,\{x(t),\,\xi(t)\}\,R^{-1}(t)\,\xi(t) \qquad (2.58)$$

which is the Kalman-Bucy continuous filter if

$$K(t) = \text{cov}\,\{x(t),\,\xi(t)\}\,R^{-1}(t) \qquad (2.59)$$

The above equations show the advantage of the innovations process by giving a simple solution to the Wiener-Hopf equation. Another advantage of the innovations approach is that it permits a more general statement of the estimation problem, and can be extended to nonlinear problems as will be shown in the next chapter.

The innovations approach has also been used to derive discrete sequential linear estimator [38].

CHAPTER 3

NONLINEAR ESTIMATION

3.1 Introduction

The problem of deriving suitable algorithms for the recursive state estimation of nonlinear dynamical systems, either continuous-time or discrete-time, has drawn wide attention in recent years. The first definitive contributions were those of Kalman and Bucy [2,3] which dealt with the optimal estimation of the state variables of a linear dynamical system. These ideas were used soon for the estimation of the states of nonlinear dynamical systems using the so-called first-order, or quasilinear filter, or extended Kalman filter (see Cox [4], Mowery [5], Ho and Lee [6], and many others).

Different techniques have been used to derive the filter equations (e.g., least-squares, maximum-likelihood, etc.). Most of these techniques, at one stage or another, employ a Taylor series expansion, neglect second and higher order terms, and use linearized equations to compute the conditional error variance matrix and the filter gains.

Another approach is based on the determination of the exact equations satisfied by the conditional probability density functions and conditional expectations. This method uses the stochastic Ito calculus and the results indicate that the optimal filter cannot be realized by a finite-dimensional system. However, the exact equations can be approximated to derive suboptimal finite-dimensional filters. This approach has been used by Kushner [7], Bucy [8], Bass et al. [9], Jazwinski [29], and many others. In many of the schemes used for the approximation second-order terms are retained, and second-order suboptimal nonlinear filters are derived.

The innovations approach to estimation theory can be extended to nonlinear least-squares estimation. The first results on the innovations approach to nonlinear estimation are due to Frost [39], and Frost and Kailath [40]. The use of the innovations allows us to obtain formulas and simple derivations that are remarkably similar to those used for the linear case. Despite the advantages of this method, it does not immediately yield any practically usable nonlinear estimators. It only suggests some suboptimal estimators.

The invariant imbedding approach is used to derive some useful results in nonlinear filtering and estimation problems. Detchmendy and Sridhar [13] and Kagiwada et al. [14] have derived filtering algorithms similar to first-order filters for nonlinear problems using the least-squares criterion and the invariant imbedding concept. This approach will be studied in detail in Chapter 4.

3.2 The Extended Kalman Filter

The theory of state estimation had its beginnings in problems of space, and its most recent advances are associated with modern aerospace problems. As we are interested in orbit determination, we will focus our attention on continuous-discrete filtering problems. In this case the system dynamics are continuous in time and the observations are discrete. The orbital dynamics can be effectively discretized at least to first order, but in highly nonlinear problems, the continuous model is more reliable.

Rather than being a nonlinear filter, the extended Kalman filter is considered to be an extension of the linear filtering theory to nonlinear problems.

As will be shown in the derivation below the extended Kalman filter is a first-order approximate nonlinear filter. This filter has been commonly and successfully used in orbit determination problems.

Suppose the nonlinear system is described by the stochastic differential equation

$$\dot{x}(t) = f(x,t) + G(t)\ w(t), \quad t \geq t_o \tag{3.1}$$

It is assumed that the initial state vector $x(0)$ is a Gaussian random variable with known mean and known covariance

$$E[x(0)] = \hat{x}_o, \quad \text{var }[x(0)] = P_o \tag{3.2}$$

and $\{w(t)\}$ is a zero-mean, white Gaussian noise process with

$$\text{cov }\{w(t), w(\tau)\} \triangleq Q(t)\,\delta_D(t-\tau) \tag{3.3}$$

Now suppose that we generate a reference (or nominal) deterministic trajectory $\bar{x}(t)$, with given $\bar{x}_o$, satisfying

$$\dot{\bar{x}}(t) = f(\bar{x},t), \qquad t \geq t_o \tag{3.4}$$

Define

$$\delta x(t) \triangleq x(t) - \bar{x}(t) \tag{3.5}$$

the deviation from the reference trajectory, then we see that $\{\delta x(t)\}$ is a stochastic process satisfying the differential equation

$$\delta\dot{x}(t) = f(x,t) - f(\bar{x},t) + G(t)\,w(t) \tag{3.6}$$

with

$$E\{\delta x(0)\} = \hat{x}_o - \bar{x}_o, \qquad \text{var }\{\delta x(0)\} = P_o \tag{3.7}$$

If the deviations from the reference trajectory are small (say in the mean square sense), then a Taylor series expansion gives

$$f(x,t) - f(\bar{x},t) \approx F(\bar{x}_o,t)\ \delta x(t) \tag{3.8}$$

where

$$F(\bar{x}_o,t) \triangleq \frac{\partial \bar{f}_i(x,t)}{\partial x_j} \tag{3.9}$$

is the matrix of partial derivatives evaluated along the reference trajectory. Thus we obtain the approximate linear equation

$$\delta\dot{x}(t) = F(\bar{x}_o,t)\ \delta x(t) + G(t)\ w(t) \tag{3.10}$$

Now we discretize (3.10) as

$$\delta x(k+1) = \Phi(k+1,k)\ \delta x(k) + w(k+1) \tag{3.11}$$

where Φ is the state transition matrix, and can be computed using the matrix F [41]

$$\phi = e^{Ft} = 1 + Ft + \frac{(Ft)^2}{2!} + \frac{(Ft)^3}{3!} + \ldots + \frac{(Ft)^n}{n!} \tag{3.12}$$

The last term in the expansion should satisfy

$$\frac{||(Ft)^n||}{n!} < a$$

where a is an arbitrary small number.

Now, the measurement model is assumed of the form

$$y(k) = h[x(k),k] + v(k) \tag{3.13}$$

where v(k) is white Gaussian noise with zero-mean and known covariance matrix

$$\text{cov}\ \{v(k)\ v(j)\} = R(k)\ \delta_K(k-j) \tag{3.14}$$

Define the nominal measurement as

$$\bar{y}(k) \triangleq h(\bar{x}(k),k) \tag{3.15}$$

and

$$\delta y(k) \triangleq y(k) - \bar{y}(k) \tag{3.16}$$

Performing a similar linearization, we get the linearized measurement equation

$$\delta y(k) = M\ [\bar{x}(k),k]\ \delta x(k) + v(k) \tag{3.17}$$

where

$$M\ [\bar{x}(k),k) \triangleq \frac{\partial \bar{h}_i\{x(k),k\}}{\partial x_j} \tag{3.18}$$

Now, the theory of the linear filter is directly applicable to the linearized system (3.11) and (3.17). Instead of state and measurement, we speak of the state deviations and measurement deviations.

Given a reference trajectory and measurements $y(k)$, we can compute $\delta y(k)$ via (3.16) and process the measurement deviations through the linear filter [Chapter 2] to estimate the state deviations. Furthermore, with the recursive structure of the linear filter, we can relinearize about each new estimate as new estimates become available. At t_o, linearize about $\hat{x}_o$, once $y(1)$ is processed, relinearize about $\hat{x}(1|1)$, and so on. The point of this is to use a better reference trajectory as soon as one is available. As a consequence of relinearization, large initial estimation errors are not allowed to propagate through time, and, therefore, the linearity assumptions are less likely to be violated.

The filter resulting from this relinearization procedure is the extended Kalman filter and is summarized as follows [26,29].

Filter algorithm

$$\hat{x}(k+1) = \hat{x}(k+1|k) + K(k+1)\ \{y(k+1) - h\ [\hat{x}(k+1|k),k+1]\} \tag{3.19}$$

One-stage prediction algorithm

$$\hat{x}(k+1|k) = \phi\,[\hat{x}(k),k] \tag{3.20}$$

Filter gain algorithm

$$K(k+1) = P(k+1|k)\, M^T\,[\hat{x}(k+1|k),k+1]\ \{M\,[\hat{x}(k+1|k),k+1]\ P(k+1|k)\, M^T\,[\hat{x}(k+1|k),k+1] + R(k+1)\}^{-1} \tag{3.21}$$

Prior error-variance algorithm

$$P(k+1|k) = \frac{\partial\phi[\hat{x}(k,k]}{\partial\hat{x}(k)}\, P(k)\, \frac{\partial\phi^T[\hat{x}(k),k]}{\partial\hat{x}(k)} + Q(k+1) \tag{3.22}$$

Error-variance algorithm

$$P(k+1) = [I - K(k+1)\, M\,\{\hat{x}(k+1|k),k+1\}]\ P(k+1|k) \tag{3.23}$$

Initial conditions

$$\hat{x}_o \text{ and } P_o$$

From the above analysis we can see that the extended Kalman filter theoretically produces an increasingly accurate estimate as additional observation data are processed. The magnitude of estimation errors as measured by the determinant of the estimation error covariance matrix is a monotonically decreasing function of the number of observations. However, in some applications, one finds that the actual estimation errors greatly exceed the values which would be theoretically predicted by the error covariance matrix. In fact, the actual error may become unbounded, even though the error variance is very small. This phenomenon, referred to as divergence, can seriously affect the filter performance. The possibility of such unstable behaviour was first suggested by Kalman [2], and later noted by others in the application of Kalman-filter algorithms to space navigation and orbit determination.

Some of the major causes of divergence are inaccuracies in the modelling process, due to failure of linearization, lack of complete knowledge of the physical problem or the simplifying assumptions necessary for computation. Errors in the statistical modelling of noise variances and mean may also lead to divergence. Another source of divergence is round-off error in digital implementation which may cause the error-covariance matrix to lose its positive definiteness or symmetry.

Different approaches have been suggested to eliminate divergence. Schee et al. [42] discuss several methods of eliminating divergence, especially with respect to orbit determination problems. Another approach which has been proposed for the control of divergence involves computing the square root of the error covariance matrix. This procedure, first suggested by Potter [43], and later extended by Andrews [44], is useful when round-off errors are a cause of divergence. Jazwinski [29] presents a suboptimal procedure to reduce the divergence problem. Adaptive filtering algorithms are also used to eliminate divergence (Smith [31], Jazwinski [32], Sage and Husa [20]). Another approach to improve the filter performance is the use of second-order filters.

3.3 Second-Order Nonlinear Filter

The primary motivation for the development of second-order filters is the divergence problem encountered when using the extended Kalman filter.

Athans et al. [11] and many others solved the problem of divergence by taking into account the state and output nonlinearities by simply retaining second-order terms in the usual Taylor series expansion, hence, deriving second-order filters.

Due to a lack of published papers that give clear comparisons between the performance of first-order and second-order filters, no one can claim that the second-order filter performance is always much better than the first-order filter. In some cases simulation results prove the improvement in the performance of second-order filters [44,45,12].

To derive the second-order filter algorithm, Athans et al. consider the system and measurement models given in equations (3.1) and (3.13). Noise statistics are the same as described in the previous section. In the derivation of the nonlinear filter, taylor series expansions of the two nonlinear functions f(·) and h(·) are employed. The Taylor series expansions are carried out to second-order terms. Certain quantities which will appear in the expansions will be defined below.

First define F(u) to be the Jacobian matrix of f(u) with elements

$$[F(u)]_{\alpha\beta} \triangleq \frac{\partial f_\alpha}{\partial u_\beta} \tag{3.24}$$

and $A_i(u)$ to be the Hessian matrix whose elements involve the various second partial derivatives of $f_i(u)$

$$[A_i(u)]_{\alpha\beta} \triangleq \frac{\partial f_i^2}{\partial u_\alpha \, \partial u_\beta} \tag{3.25}$$

The Taylor series expansion of f(u) about a vector u_o, "near" u, is given by

$$f(u) \approx f(u_o) + F(u_o)\ (u-u_o) + \frac{1}{2} \sum_{i=1}^{n} \pi_i (u-u_o)^T A_i(u_o)\ (u-u_o) \tag{3.26}$$

where π_i denotes the natural basis vector, i = 1, 2, ..., n, and n is the dimension of the state vector $x(t)$.

Similarly the Taylor series expansion of $h(u)$ about u_o is given by

$$h(u) \approx h(u_o) + M(u_o)\,(u-u_o) + \frac{1}{2}\sum_{j=1}^{m} \pi_j (u-u_o)^T C_j(u_o)\,(u-u_o) \tag{3.27}$$

where

$$[M(u)]_{\alpha\beta} \triangleq \frac{\partial h_\alpha}{\partial u_\beta} \tag{3.28}$$

and

$$[C_j(u)]_{\alpha\beta} \triangleq \frac{\partial^2 h_j}{\partial u_\alpha\, \partial u_\beta} \quad , \; j = 1,\, 2,\, \ldots,\, m \tag{3.29}$$

m is the dimension of the observation vector y. The details of the derivations and the assumptions made are given in reference [11]. The filter algorithm is summarized below.

<u>Plant</u>

$$\dot{x}(t) = f(x,t)\,,\; x(t_o) = x_o$$

<u>Observations</u>

$$y(k) = h[x(k),k] + v(k)$$

State estimate at observation time

$$t_k: \quad \hat{x}(k)$$

$$\hat{e}(k) = x(k) - \hat{x}(k) \tag{3.30}$$

$$P(k) = E\ [\hat{e}(k)\ \hat{e}^T(k)] \tag{3.31}$$

State estimate at t

$$t_k < t < t_{k+1}: n(t)$$

$$e(t) = x(t) - n(t) \tag{3.32}$$

$$S(t) = E\ [e(t)\ e^T(t)] \tag{3.33}$$

Basic asumptions

$\hat{e}(k)$ and $e(t)$ Gaussian, zero-mean

Starting conditions

$$\hat{x}_o = n(t_o) = E\ [x_o] = \bar{x}_o \tag{3.34}$$

$$P_o = S(t_o) = E\ [(x_o - \bar{x}_o)\ (x_o - \bar{x}_o)^T] \tag{3.35}$$

Continuous-time filter

$$t_k < t < t_{k+1}$$

$$\dot{n}(t) = f(n,t) + \frac{1}{2} \sum_{i=1}^{n} \pi_i \ \mathrm{tr}\ [A_i(n,t)\ S(t)] \tag{3.36}$$

$$\dot{S}(t) = F(n,t)\, S(t) + S(t)\, F^T(n,t) \tag{3.37}$$

$$n(t_k) = \hat{x}(k)$$

$$S(t_k) = P(k)$$

Update at $t = t_{k+1}$

$$\hat{x}(k+1) = n(k+1) + K(k+1)\, [y(k+1) - h\{n(k+1),k+1\}] - N(k+1) \tag{3.38}$$

$$N(k+1) = \frac{1}{2} K(k+1) \sum_{j=1}^{m} \pi_j \,\mathrm{tr}\, [C_j \{n(k+1)\}\, S(k+1)] \tag{3.39}$$

$$K(k+1) = S(k+1)\, M^T\{n(k+1)\}\, [M\{n(k+1)\}\, S(k+1)\, M^T\{n(k+1)\} + R(k+1) + L(k+1)]^{-1} \tag{3.40}$$

$$P(k+1) = S(k+1) - K(k+1)\, M\{n(k+1)\}\, S(k+1) \tag{3.41}$$

$$[L(k+1)]_{ij} = \frac{1}{2}\,\mathrm{tr}\, [C_i \{n(k+1)\}\, S(k+1)\, C_j \{n(k+1)\}\, S(k+1)\} \tag{3.42}$$

where $[L(k+1)]_{ij}$ is the ijth element of the mxm matrix $L(k+1)$.

We conclude from the above equations that the one-stage prediction algorithms for the state and the error covariance matrix are the solutions of first order differential equations (3.36) and (3.37). The filter and filter gain algorithms are discrete-time recursive algorithms given by equations (3.38) and (3.40). Real-time estimation is possible using the above algorithm as long as the time required to complete the computation cycle is less than the time interval between successive observations.

In order to compare the differences between the first-order and second-order filter, simulation results for an orbit determination case will be given in Chapter 6.

3.4 Innovations Approach to Nonlinear Estimation

Frost and Kailath [40] show how the innovations approach to estimation theory can be extended to the nonlinear least-squares estimation of non-Gaussian signals in additive white Gaussian noise.

There are several ways of using the innovations concept, the one used by Kailath has the advantage that it leads to derivations similar to the linear case. However, this approach does not immediately yield any dramatically simple and practically usable nonlinear estimators.

The basic results are obtained for the following model, given observation of the form

$$y(t) = z(t) + v(t), \qquad 0 \leq t < T \tag{3.43}$$

The statistical assumptions made are the same as those given in Chapter 2 by equations (2.42), (2.43) and (2.44).

In the case of nonlinear estimation z(t) is given by

$$z(t) = h[x(\tau), \tau \leq t] \tag{3.44}$$

Under the above assumptions we wish to find the least-squares estimate $\hat{x}(t|t)$ of x(t) given $\{y(\tau), 0 \leq \tau < t \leq T\}$. As in Chapter 2 the estimate is given by

$$\hat{x}(t|t) = E\ \{x(t)\ |\ y(\tau),\ 0 \leq \tau < t\} \tag{3.45}$$

The search for a more explicit representation for equation (3.45) is an imposing task challenging researchers in the field of nonlinear estimation.

Define the innovations process

$$\xi(t) = y(t) - \hat{z}(t|t) \tag{3.46}$$

where

$$\hat{z}(t|t) = E\ \{z(t)|y(\tau),\ 0 \leq \tau < t\} \tag{3.47}$$

By using the innovations approach the data process $y(\cdot)$ is transformed into a white noise process $\xi(\cdot)$ and the optimal estimator is determined as a functional of the innovations process.

The basic idea is to construct a causal and causally invertible transformation of the observation process into a white noise process to simplify the problem. However, it is necessary to establish that the observations process can be generated from the innovations process before this method can be applied fruitfully.

The equivalence problem or "causal equivalence" has not been solved for nonlinear problems with Gaussian noise. For nonlinear problems with white noise, therefore, the resulting estimate based on pre-supposed innovations process could best be called "suboptimal".

Based on the projection theorem for nonlinear systems and the causal equivalence theorem [40], the estimator is given by

$$\hat{x}(t|t) = E\ [x(t)\ |\ \xi(\tau),\ 0 \le \tau < t] \tag{3.48}$$

The right-hand expression of equation (3.48) is a functional of a white Gaussian noise process. Hence, it could be expressed as an Itô integral using the following lemma:

Lemma (Doob [37])

Given a zero-mean finite variance functional $G(\cdot)$ of the white Gaussian noise $\{\xi(\sigma),\ 0 \le \sigma < t\}$, there exists a functional $g\ [t,\ \tau,\ (\xi(\sigma),\ 0 \le \sigma < \tau)]$ with the property

$$G\ [\xi(\tau),\ 0 \le \tau < t] = \int_0^t g^T\ (t,\ \tau,\ w)\ \xi(\tau)\ d\tau \tag{3.49}$$

where

$$E\{G^2\} = \int_0^t E\{g^T(t, \tau, w)\, g(t, \tau, w)\}\, d\tau < \infty \quad (3.50)$$

f

The proof is given in reference [46].

Based on equations (3.48) and (3.49), a basic least-squares representation formula is developed

$$\hat{x}(t|t) = \int_0^t E\,[x(t)\,\xi^T(\tau) \mid \xi(\sigma),\ 0 \le \sigma < \tau]\,\xi(\tau)\, d\tau \quad (3.51)$$

The formula given above is very similar to that for the optimal linear estimate given in Chapter 2 by equation (2.48). However, this similarity does not imply that the nonlinear estimate $\hat{x}(t|t)$ is easy to implement as the linear estimate considered previously. In the linear estimation estimates are epxressed in terms of covariance functions that can be calculated analytically. As for the nonlinear estimation problem equation (3.51) for the optimal estimate simply states relationships between conditional expectations.

The nonlinear systems we are studying are given in the form of stochastic differential equations, which are a special case of Itô processes.

In the case of Itô processes Kailath derived a differential structure of $\hat{x}(t|t)$ given by

$$\dot{\hat{x}}(t|t) = \hat{f}(t|t) + K(x,t)\ \xi(t) \tag{3.52}$$

where

$$\dot{x}(t) = f(x,t) + G(t)\ w(t)$$

$$\hat{f}(t|t) = E\ [f(x,t)\ |\ \xi(\sigma),\ 0 \le \sigma < t] \tag{3.53}$$

and

$$K(x,t) = E\ [x(t)\ \xi^T(t)\ |\ \xi(\sigma),\ 0 \le \sigma,\ t] \tag{3.54}$$

Equation (3.52) is simply a representation for the recursive filtering rather than an explicit formula since the functions $\hat{f}(\cdot)$ and $K(\cdot)$ are in general indeterminate.

For computational purposes, it would be desirable to obtain an iterative form for the gain matrix K. This is the motivation for the adaptive scheme proposed in Chapter 4. A stochastic approximation algorithm for numerical computations of the innovations process and the system state is combined with the invariant imbedding algorithm to derive a new nonlinear filtering algorithm.

CHAPTER 4

NEW NONLINEAR ESTIMATION ALGORITHM BASED ON INVARIANT IMBEDDING AND STOCHASTIC APPROXIMATION

4.1 Introduction

The main objective of this research is to develop an efficient algorithm for optimal nonlinear estimation. The extended Kalman filter and the second-order nonlinear filter algorithms presented in Chapter 3 require prior knowledge of the message-generating and observation noise covariance matrices, as well as the covariance matrix of the initial estimation error. In practice, however, such extensive a priori information is seldom available, with the result that the optimal filter gain matrix cannot be calculated.

On the other hand the innovations approach provides an abstract mathematical model which is not practically usable for nonlinear estimation.

To overcome these difficulties, the invariant imbedding concept is used to obtain a recursive estimator which does not depend on a priori noise statistical assumptions. Stochastic approximation is shown to be similar to the invariant imbedding method; both the estimators do not require prior knowledge of the system statistics, and they are of a sequential nature.

It is proposed to develop an algorithm combining the method of invariant imbedding with stochastic approximation in order to obtain an adaptive approach for estimating the optimal gain matrix and also to improve the rate of convergence.

4.2 Invariant Imbedding and Nonlinear Estimation

The invariance principle, now known as invariant imbedding, has been introduced in the study of transport phenomena, radiative transfer and wave propagation. In general this approach has proved useful in treating boundary values problems, eigenvalue problems and nonlinear estimation theory [47,13,14].

Since the invariant imbedding approach is different from the usual classical approach, several advantages have been gained. First, the present approach is applicable to a wide variety of nonlinear problems. Second, a sequential estimator is obtained, which can be implemented in real time. Another advantage of the invariant imbedding filtering algorithm is that no statistical assumptions will be made concerning the noise or disturbances, because for most practical problems the determination of valid statistical data is itself a difficult problem.

The generally used least squares criterion will be employed to obtain the optimal estimates. The numerical aspects of this approach are given in detail in reference [47].

Consider the nonlinear vector equation

$$\dot{x}(t) = f(x,t), \qquad 0 \leq t \leq \tau \tag{4.1}$$

where x and f are n-dimensional vectors.

The measurement model is given by the m-dimensional vector equation

$$y(t) = h(x,t) + (\text{measurement errors}) \tag{4.2}$$

The above estimation problem is essentially the same as that considered in the previous chapter. Although the signal, y(t), has been expressed as a continuously measured signal in time, it can be extended directly to the discrete-time case.

On the basis of the measurements or observations y(t), $0 \leq t \leq \tau$, estimate the n conditions

$$x(\tau) = c \tag{4.3}$$

for equation (4.1) such that the integral

$$J = \int_0^{\tau} \sum_{j=1}^{m} \{y_j(t) - h_j(x,t)\}^2 \, dt \tag{4.4}$$

is minimized. The functions h_j are evaluated by using the values of x obtained from (4.1).

The estimator equation for this problem can be obtained as follows. Define a new variable z(t):

$$z(t) = \int_0^t \sum_{j=1}^{m} \{y_j(t) - h_j(x,t)\}^2 \, dt \tag{4.5}$$

Differentiating z(t), we get

$$\dot{z}(t) = \sum_{j=1}^{m} \{y_j(t) - h_j(x,t)\}^2 \tag{4.6}$$

From (4.4) and (4.5)

$$z(\tau) = J \tag{4.7}$$

The differential equations to be considered are (4.1) and (4.6). Although the original problem is to minimize $z(\tau)$, we shall ignore the minimization first and obtain $z(\tau)$ for the above system by invariant imbedding. Instead of a two-point boundary value problem, the invariant imbedding approach treats the problem as a family of problems with different final points, τ. So, consider the family of problems with final points a

$$x(a) = c, \qquad 0 \leq t \leq a \tag{4.8}$$

If we define

$r(c,a) \triangleq$ the missing final condition for the system represented by (4.1) and (4.8) where the process ends at $t = a$ with $x(a) = c$

then

$$z(a) = r(c,a) \tag{4.9}$$

We shall consider r as the dependent variable, c and a as the independent variables. An expression for r in terms of c and a will be obtained. Considering the neighbouring process with starting value $a+\Delta$, the missing initial condition of this neighbouring process can be related to $z(a)$ by the use of Taylor's series expansion

$$z(a+\Delta) = z(a) + \dot{z}(a)\ \Delta + 0(\Delta) \tag{4.10}$$

At the starting value a, equations (4.1) and (4.6) become

$$\dot{x}(a) = f(c,a) \tag{4.11}$$

$$\dot{z}(a) = \sum_{j=1}^{m} \{y_j(a) - h_j(c,a)\}^2 \tag{4.12}$$

Substituting (4.12) and (4.9) into (4.10), we obtain

$$z(a+\Delta) = r(c,a) + \sum_{j=1}^{m} \{y_j(a) - h_j(c,a)\}^2\ \Delta + 0(\Delta) \tag{4.13}$$

On the other hand from (4.9)

$$z(a+\Delta) = r\{x(a+\Delta), a+\Delta\} \tag{4.14}$$

Again, the expression $x(a+\Delta)$ can be related to its neighbouring process $x(a) = c$ by Taylor's series expansion

$$
\begin{aligned}
x(a+\Delta) &= x(a) + \dot{x}(a)\,\Delta + O(\Delta) \\
&= c + f(c,a)\,\Delta + O(\Delta) \qquad (4.15)
\end{aligned}
$$

Equating equations (4.13) and (4.14), we obtain,

$$r(c,a) + \sum_{j=1}^{m} \{y_j(a) - h_j(c,a)\}^2 \Delta = r\{c+f(c,a)\Delta, a+\Delta\} \qquad (4.16)$$

Omitting the terms involving powers of Δ higher than the first, expanding the right-hand side of (4.16) by Taylor's series expansion, we obtain

$$
\begin{aligned}
r\{c+f(c,a)\Delta,\ a+\Delta\} &= r(c,a) + \sum_{i=1}^{n} f\ (c,a)\ \Delta \frac{\partial r(c,a)}{\partial c_i} \\
&\quad + \Delta \frac{\partial r(c,a)}{\partial a} + O(\Delta) \qquad (4.17)
\end{aligned}
$$

In the limit as Δ tends to zero, the following first-order quasilinear partial differential equation is obtained from (4.16) and (4.17)

$$\frac{\partial r(c,a)}{\partial a} + \sum_{i=1}^{n} f_i\ (c,a) \frac{\partial r(c,a)}{\partial c_i} = \sum_{j=1}^{m} \{y_j(a) - h_j(c,a)\}^2 \qquad (4.18)$$

This is the invariant imbedding equation for the missing final condition.

If $e(a)$ is the optimal estimate of c, then

$$\frac{\partial r(e,a)}{\partial c_i} = r_{c_i}(e,a) = 0, \; i = 1, 2, \ldots, n \qquad (4.19)$$

or

$$\sum_{k=1}^{n} \frac{\partial}{\partial c_k} [r_{c_i}(e,a)] \, de_k + \frac{\partial}{\partial a} [r_{c_i}(e,a)] \, da = 0 \qquad (4.20)$$

In matrix notation, the above equation becomes

$$\frac{de}{da} = - [r_{cc}(e,a)]^{-1} \, r_{ca}(e,a) \qquad (4.21)$$

where the symbol $[r_{cc}]^{-1}$ denotes the inverse of the matrix r_{cc} given by

$$r_{cc} = \begin{bmatrix} r_{c_1c_1} & \cdots & r_{c_1c_2} & \cdots & r_{c_1c_n} \\ \vdots & & \vdots & & \vdots \\ r_{c_nc_1} & \cdots & \cdots & \cdots & r_{c_nc_n} \end{bmatrix} \qquad (4.22)$$

and

$$r_{ca} = \frac{\partial}{\partial a} \begin{bmatrix} r_{c_1} \\ r_{c_2} \\ \vdots \\ r_{c_n} \end{bmatrix} \qquad (4.23)$$

To find an expression for the right-hand side of (4.21), equation (4.18) can be differentiated with respect to c_1,

$c_2, \ldots, c_n$, we obtain

$$r_{ca}(c,a) + r_{cc}(c,a)\ f(c,a) + f_c^T(c,a)\ r_c(c,a)$$
$$= -2\ h_c^T(c,a)\ \{y(a) - h(c,a)\} \tag{4.24}$$

At the optimal estimate of c, equation (4.19) can be substituted into (4.24)

$$r_{ca}(e,a) + r_{cc}(e,a)\ f(e,a) = -2\ h_c^T(e,a)\ \{y(a) - h(e,a)\} \tag{4.25}$$

Combining (4.25) and (4.21) we obtain

$$\frac{de}{da} = f(e,a) + 2\ [r_{cc}(e,a)]^{-1}\ h_c^T(e,a)\ \{y(a) - h(e,a)\} \tag{4.26}$$

If we let

$$q(a) = 2\ [r_{cc}(e,a)]^{-1} \tag{4.27}$$

then

$$\frac{de}{da} = f(e,a) + q(a)\ h_c^T(e,a)\ \{y(a) - h(e,a)\} \tag{4.28}$$

where q is an nxn matrix.

A set of differential equations for q can be obtained. Rewrite (4.27) in the form

$$\frac{1}{2}\ q(a)\ r_{cc}(e,a) = 1 \tag{4.29}$$

Differentiating (4.29), we obtain

$$\frac{dq}{da} r_{cc}(e,a) + q(a) \frac{d}{da} \{r_{cc}(e,a)\} = 0 \tag{4.30}$$

or

$$\frac{dq}{da} = - \frac{1}{2} q(a) \frac{d}{da} \{r_{cc}(e,a)\} q(a) \tag{4.31}$$

where

$$\frac{d}{da} \{r_{cc}(e,a)\} = r_{cca}(e,a) + (\text{terms involving } r_{ccc}) \tag{4.32}$$

Assume that the terms involving r_{ccc} are negligible, then

$$\frac{dq}{da} = - \frac{1}{2} q(a) \, r_{cca}(e,a) \, q(a) \tag{4.33}$$

where r_{cca} is the $\partial/\partial a$ of the matrix (4.22).

To find an expression for r_{cca}, consider the partial derivative of (4.24) with respect to c_1, c_2, ..., c_n, we obtain

$$r_{cca}(c,a) + r_{cc}(c,a) \, f_c(c,a) + f_c^T(c,a) \, r_{cc}(c,a)$$

$$+\xi = -2 \, [h_{cc}(c,a) \, \{y(a) - h(c,a)\}] + 2h_c^T(c,a) \, h_c(c,a) \tag{4.34}$$

The elements of the matrix represented by the first term on the right-hand side are scalar or inner products of the vector $h_{c_i c_j}$ and $\{y-h\}$. Thus

$$h_{cc}[y-h] = \begin{bmatrix} h^T_{c_1c_1}\,[y-h] & h^T_{c_1c_2}\,[y-h] & \cdots & h^T_{c_1c_n}\,[y-h] \\ h^T_{c_2c_1}\,[y-h] & \vdots & & \vdots \\ \vdots & & & \\ h^T_{c_nc_1}\,[y-h] & \cdots & \cdots & h^T_{c_nc_n}\,[y-h] \end{bmatrix} \qquad (4.35)$$

where

$$h_{c_ic_j} = \frac{\partial^2}{\partial c_i\ \partial c_j} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{bmatrix} \qquad (4.36)$$

The term ξ in (4.34) represents terms that consist of terms of the form of r_c and r_{ccc}. When c takes on its optimal estimate e, $r_c = 0$, and r_{ccc} are negligible. Consequently ξ is negligible.

Combining (4.29), (4.33) and (4.34), the desired differential equation for q is obtained:

$$\begin{aligned} \frac{dq}{da} = f_c(e,a)\ q(a) + q(a)\ f^T_c(e,a) + q(a) \\ \{h_{cc}(e,a)\ [y(a) - h(e,a)]\}\ q(a) \\ - q(a)\ h^T_c(e,a)\ q(a) \end{aligned} \qquad (4.37)$$

Finally the desired estimator equations are summarized as follows

$$\frac{de}{da} = f(e,a) + q(a)\, h_e^T(e,a)\, \{y(a) - h(e,a)\} \qquad (4.37)$$

$$\frac{dq}{da} = f_e(e,a)\, q(a) + q(a)\, f_e^T(e,a) + q(a)$$

$$\{h_{ee}(e,a)\, [y(a) - h(e,a)]\}\, q(a) = q(a)\, h_e^T(e,a)$$

$$h_e(e,a)\, q(a) \qquad (4.38)$$

To express e and q in a practical form, similar to the previous filters, consider $x(t_o)$ to be a Gaussian random variable, then it is appropriate that the values of e_o and q_o are chosen as the mean x_o and covariance P_o of the distribution respectively.

So the estimator equations (4.37) and (4.38) provide the solution for the optimal state estimate $\hat{x}(t)$ and the corresponding state error covariance matrix P(t).

4.3 Stochastic Approximation

4.3.1 Stochastic Approximation Algorithms

Stochastic approximation methods may be considered as recursive estimation methods, updated by an appropriately weighted, arbitrarily chosen error corrective term, with the only requirement that, in the limit, the estimate converges to the true parameter sought. Applications of stochastic approximation algorithms have been proposed in adaptive and learning systems [48], adaptive communication [49] and systems identification [50].

Historically, stochastic approximation was first treated by Robbins and Monro [51] and Keifer and Wolfowitz [52], who were concerned with solution to two specific problems; finding the root of a regression function, and finding the value that minimizes a regression function given only pertinent random observations. It was Dvoretzky [53] who generalized stochastic approximation to any sort of iterative solution algorithm, which is convergent, when direct observations of a regression function can be adopted successfully. Excellent surveys of stochastic approximation can be found in papers by Sakrison [49] and Saridis [54].

For most practical applications stochastic approximation search algorithms are point estimators of the form

$$\alpha(k) = \alpha(k-1) + \gamma_k \, [F(k) + v(k)] \tag{4.39}$$

where

γ_k = gain sequence of suitable chosen smoothing values

$[F(k) + v(k)]$ = error correction sequence generated at every time instant k by measuring the deviation from an appropriate goal.

The iterative scheme (4.39) approaches the optimal parameter value, α_{op}, where $E[F(k)+v(k)] = 0$, in the mean square sense provided the following assumptions are satisfied [54]

$$\lim_{k\to\infty} \gamma_k = 0 \quad , \qquad \sum_{k=1}^{\infty} \gamma_k = \infty, \qquad \sum_{k=1}^{\infty} \gamma_k^2 < \infty \qquad (4.40)$$

and the error correction sequence satisfies

$$a||\alpha-\alpha_{op}||^2 \leq \langle\alpha-\alpha_{op}, E[F(k)] + v(k)\rangle \leq b||\alpha-\alpha_{op}||^2$$
$$0 < a \leq b < \infty \qquad (4.41)$$

where $\langle\cdot,\cdot\rangle$ denotes the matrix inner product operator.

The conditions (4.40) on the gains may be interpreted heuristically as follows. The first one provides the smoothing effect on the random correction term, the second provides unlimited correction effort, and the third guarantees mutual cancellation of individual errors for a large number of iterations. The harmonic sequence [1/k] as well as any sequence of the form [a/k+b] satisfy the above conditions.

Conditions (4.41) imply that the regression function $[F(k)+v(k)]$ is bounded on all sides of a true solution by a rectangular set in the solution space such that it is not possible to overshoot the solution, α_{op}, which cannot be corrected by a γ_k satisfying (4.40).

Stochastic approximation techniques are used in adaptive filtering problems to adaptively estimate the filter gain.

4.3.2 Adaptive Filtering and Stochastic Approximation

In general, stochastic approximation algorithms of the Robbins and Monro type used in adaptive filtering are of the form

$$K(k+1) = K(k) + \gamma(k)\ [f(y(k),\ K(k))-m_o] \tag{4.42}$$

The iterative scheme (4.42) approaches the optimal gain, K_{op}, where $E[f(y(k),\ K_{op})] = m_o$, in the mean square sense provided the assumptions (4.40) and (4.41) are satisfied (replacing α with K).

The successful application of stochastic approximation to search for the optimum gain matrix, K_{op}, requires some suitable method for testing if the value presently being used is optimum. Hampton and Schultz [55] have proposed a Robbins and Monro type algorithm which uses the orthogonal condition

$$E\ [\{x(k) - \hat{x}(k|k-1)\}\ y^T(j)] = 0 \tag{4.43}$$

for this test. Since the actual x(k) is not known, equation (4.42) can only be approximated in a rather involved manner. More recently, Sinha and Mukherjee [36] have also proposed a Robbins and Monro type algorithm which makes use of the property that the innovations process is white [26]. They used for the test of optimality

$$E\,[\xi(k)\ \xi^T(j)] = 0 \tag{4.44}$$

Although their idea is conceptually more direct and works quite well for the scalar case, it is unsuitable for the multivariable case since the error correction term is restricted to a subspace of the solution set in most instances.

Following the above concept equation (4.42) is written in the form

$$K(k+1) = K(k) + \gamma(k)\ f\,[\xi(k)] \tag{4.45}$$

The function $f[\xi(k)]$ should be a null matrix when $K(k)$ is identical with K_{op}.

An algorithm proposed by Scharf and Alspach [34] uses the orthogonality between the innovations process and the estimated state, that is,

$$E\,[\hat{x}(k|k-1)\ \xi^T(k)] = 0 \tag{4.46}$$

Note that the product $\hat{x}(k|k-1)\ \xi^T(k)$ is contained in the same space as that of the Kalman gain, making it applicable for multivariable problems. The stochastic approximation algorithm thus can be written as

$$K(k+1) = K(k) + \gamma(k) \frac{\hat{x}(k|k-1) \ \xi^T(k)}{||x(k|k-1) \ \xi^T(k)||^2} \tag{4.47}$$

where $\gamma(k)$ is chosen so as to satisfy assumption (4.40). Scharf and Alspach have shown, in the scalar case of (4.47), that the regression function, $\hat{x}(k|k-1) \ \xi^T(k)$, satisfies assumption (4.41), thereby showing that (4.47) converges in the mean-square sense to K_{op}. The proof of mean-square convergence of (4.47) in the multivariable case is given in reference [56].

The above algorithm was used by the author for adaptive state estimation of a synchronous-orbit satellite [57].

4.4 Combined Invariant Imbedding and Stochastic Approximation Algorithm

4.4.1 The Invariant Imbedding Algorithm

The problem of orbit determination, considered in the thesis, is a combination of continuous and discrete-time problems, since the observations are available only at discrete instants of time, but the state model is of the continuous time-type.

Therefore, emphasis will be placed on a discrete-estimation case using invariant imbedding. The algorithm developed will then be extended to the special "continuous-discrete" trajectory estimation problem.

Sage and Melsa [58] developed an algorithm for maximum a posteriori discrete filtering by using the method of invariant imbedding to determine approximate filtering solutions.

The discrete message and observation models are given by

$$x(k+1) = \phi\,[x(k),k] + \Gamma(k)\; w(k) \tag{4.48}$$

$$y(k) = h\,[x(k),k] + v(k) \tag{4.49}$$

$w(k)$ and $v(k)$ are assumed to be independent zero-mean Gaussian white sequences with covariances $Q(k)$ and $R(k)$ respectively.

The best estimate of x throughout an interval will in general depend on the criteria used to determine the best estimate. In their analysis, Sage and Melsa derived the best estimate from maximizing the conditional probability function $p[X|Y]$, where X is the sequence $x(k_o)$, $x(k_1)$, ..., $x(k_f)$ and Y is the sequence $y(k_1)$, $y(k_2)$, ..., $y(k_f)$.

Applying Baye's rule to $p[X|Y]$ results in

$$p[X|Y] = \frac{p[Y|X]\; p[X]}{p[Y]} \tag{4.50}$$

With a considerable amount of algebraic manipulation it is shown that maximizing (4.50) is equivalent to minimizing

$$J = \frac{1}{2} \| x(k_o) - \hat{x}(k_o) \|^2_{P^{-1}(k_o)} + \frac{1}{2} \sum_{k=k_o}^{k_f-1} \| y(k+1) - h[x(k+1), k+1] \|^2_{R^{-1}(k+1)} + \frac{1}{2} \sum_{k=k_o}^{k_f-1} \| w(k) \|^2_{Q^{-1}(k)} \tag{4.51}$$

subject to the equality constraint of equation (4.48).

Applying the discrete Euler-Lagrange equations [58], the Hamiltonian is defined as

$$H\, [x(k),\, w(k),\, \lambda(k+1),\, k] = \frac{1}{2} \| y(k+1) - h\, [\bar{x}(k+1),\, k+1] \|^2_{R^{-1}(k+1)} + \frac{1}{2} \| w(k) \|_{Q^{-1}(k)} + \lambda^T(k+1)\; x(k+1) \tag{4.52}$$

Based on equation (4.52) and its associated boundary conditions, a nonlinear two-point boundary-value problem is derived in the form

$$\hat{x}(k+1) = f\, [\hat{x}(k),\, \lambda(k),\, k] \tag{4.53}$$

$$\lambda(k+1) = g\, [\hat{x}(k),\, \lambda(k),\, k] \tag{4.54}$$

with boundary conditions given by

$$\lambda(k_o) = a \qquad \lambda(k_f) = 0$$

To derive a sequential approximate filtering equations, the invariant imbedding concept is used. The state $\hat{x}$ and the final conditions λ are imbedded so that

$$\hat{x}(k_f) = r(c,k_f) \tag{4.55}$$

$$\lambda(k_f) = c \tag{4.56}$$

Based on the analysis given in section 4.2, a discrete invariant-imbedding equation is obtained in the form

$$\frac{\partial r}{\partial k} + \frac{\partial r}{\partial c} + \frac{\partial^2 r}{\partial c\,\partial k}\{g(r,c,k_f) - c\} = f(r,c,k_f) - r \tag{4.57}$$

A solution of (4.57) is assumed to have the form

$$r(c,k) = \hat{x}(k) - P(k)\;c \tag{4.58}$$

After several manipulations the following approximate-sequential-estimation equations are obtained

$$\hat{x}(k+1) - P(k+1)\;g\;[\hat{x}(k),\;0,\;k] = f\;[\hat{x}(k),\;0,\;k] \tag{4.59}$$

$$P(k+1)\;\left.\frac{\partial g\;[r(c,k),c,k]}{\partial c}\right|_{c=0} = \left.\frac{-\partial f\;[r(c,k),c,k]}{\partial c}\right|_{c=0} \tag{4.60}$$

with given initial conditions

$$\hat{x}_o \text{ and } P_o$$

To proceed to the development of the final filter algorithms, the problem of minimizing J is considered and the following filter equations are obtained.

Filter algorithm

$$\hat{x}(k+1) = \hat{x}(k+1|k) + K(k+1)\ \{y(k+1) - h[\hat{x}(k+1|k),k+1]\} \quad (4.61)$$

One stage prediction algorithm

$$\hat{x}(k+1|k) = \phi\ [\hat{x}(k),k] \quad (4.62)$$

Filter gain algorithm

$$K(k+1) = P(k+1)\ \frac{\partial h^T\ [\hat{x}(k+1|k),k+1]}{\partial \hat{x}(k+1|k)}\ R^{-1}(k+1) \quad (4.63)$$

Prior error-variance algorithm

$$P(k+1|k) = \frac{\partial \phi[\hat{x}(k),k]}{\partial \hat{x}(k)}\ P(k)\ \frac{\partial \phi^T[\hat{x}(k),k]}{\partial \hat{x}(k)} + \Gamma(k)\ Q(k)\ \Gamma^T(k) \quad (4.64)$$

Error-variance algorithm

$$P(k+1) = [P^{-1}(k+1|k) - \frac{\partial}{\partial \hat{x}(k+1|k)}\ (\frac{\partial h^T[\hat{x}(k+1|k),k+1]}{\partial \hat{x}(k+1|k)}$$

$$R^{-1}(k+1)\ \{y(k+1) - h\ [\hat{x}(k+1|k),\ k+1]\}))]^{-1} \quad (4.65)$$

In the special case of continuous message model and discrete observations for orbital trajectory estimation the discrete invariant imbedding algorithm given above can be used with the following modifications.

One-stage prediction algorithm

$$\hat{x}(k+1|k) \triangleq x(t_{k+1}^{-})$$

which is the solution of

$$\dot{x}(t) = f\,[x(t),\, t]\,, \qquad t_k < t < t_{k+1} \tag{4.66}$$

Prior-error-variance algorithm

$$P(k+1|k) = \Phi(k+1|k)\, P(k)\, \Phi^T(k+1|k) \tag{4.67}$$

where ϕ is the state transition matrix and is obtained by solving

$$\begin{aligned} \dot{\phi}(t) &= \frac{\partial f}{\partial x}\,(x,t)\ \phi(t) \\ \phi(k+1|k) &= \phi(t_{k+1}^{-}) \end{aligned} \tag{4.68}$$

The sequential filter algorithm used in the thesis consists of the computations outlined in equations (4.59) to (4.63) with (4.60) replaced by (4.64) and (4.62) by (4.65) and (4.66).

4.4.2 Proposed Algorithm

The approximate nonlinear filter based on the invariant imbedding concept and outlined above has the following advantages.

(1) No statistical assumptions are required concerning the nature of the input disturbances or the measurement errors. The system and measurement error covariance matrices appear as weighting functions R^{-1} and Q^{-1} in equation (4.51) allowing us to place more emphasis on the most reliable measurements. Otherwise, we assume $Q(k) = 0$ and $R(k) = [I]$.

(2) A sequential estimation scheme is obtained, which makes it possible to implement in real time.

(3) Convergence of the algorithm is theoretically guaranteed since the estimator equations include terms of the form $[y-h(\hat{x})]$, which is the innovations process discussed in Chapters 2 and 3.

On the other hand, the disadvantages of this method are

(1) More computations are needed compared to quasilinear or first-order filters. The invariant imbedding algorithm is a second-order filter as second derivatives of the nonlinear function $h(\cdot)$ are used in the computations.

(2) In real-time simulations, it is necessary that the time required to complete the computation cycle be less than the time interval between successive observations. This is why, in some applications, second-order filters as the invariant imbedding filter are not suitable.

(3) The initial values x_o and P_o affect the rate of convergence.

Stochastic approximation is a scheme for successive approximation of a sought quantity (i.e. the filter gain), when the observations involve random errors due to the stochastic nature of the problem. It has the following advantages:

(1) Only a small amount of data needs processing.

(2) Only simple computations are required. In the algorithm proposed by Scharf and Alspach (Section 4.3) the computations needed are those of the innovations process and the state estimates.

(3) A priori knowledge of the process statistics is not necessary. The only requirements are that the regression function satisfy certain regularity conditions and that the regression problem have a unique solution.

The disadvantages of stochastic approximation algorithms are

(1) Slow convergence
(2) Convergence properties depend on the starting values.

From the above analysis, it would appear logical to combine the two methods in such a manner as to retain their relative advantages, while disposing with their basic drawbacks. It is proposed, therefore, to use the invariant imbedding algorithm to obtain the initial value of the gain matrix and then apply stochastic approximation to track any changes in the gain matrix which may bring further improvement.

When treated separately, the convergence of each of the two estimators is theoretically justified. However, the convergence of the combined filter algorithm is not obvious. This is because the behaviour of one affects the operation of the other. A theoretical proof of convergence could not be obtained, so the algorithm is tested using real data and simulation results of orbit determination study prove the convergence and efficiency of the proposed estimator.

The proposed algorithm is summarized as follows.

(1) Set $k = 0$. Solve the suboptimal nonlinear estimation problem using invariant imbedding for $k = 0, 1, \ldots,$

p, where p is a certain number of data points.

(2) Compute the suboptimal gain matrix.

(3) Starting at $k = p+1$ use stochastic approximation for adaptive gain estimation for the remainder of the observations as given by equation (4.47).

Therefore, using the invariant imbedding algorithm a "good" starting value for the gain matrix is provided for the stochastic approximation algorithm. Since the stochastic approximation algorithm requires very little computation per iteration the proposed algorithm is computationally more efficient relative to the invariant imbedding algorithm if used alone.

Simulation results in Chapter 6 demonstrate the convergence and the efficiency of the proposed algorithm with comparison to other approaches in the case of satellite orbit determination considered.

CHAPTER 5

MODELLING OF DYNAMICS AND OBSERVATIONS FOR COMMUNICATIONS SATELLITES

5.1 Introduction

With the increasing application of man-made satellites for communication purposes, there has developed a considerable amount of technology for their control. Such control must be in the form of corrective action in order to maintain the satellite in a desired orbit. This requires obtaining proper estimates of the state of the satellite from measurements contaminated with noise. The states of the satellite are the fundamental elements or parameters that define the orbit, in this case the estimation problem is an orbit determination problem.

A realistic dynamical model that describes both the attitude and trajectory dynamics for the satellite is developed based on the model equations given by Altman [18]. In order to employ the state model in estimation or control, a measurement model should be provided.

5.2 A Unified State Model for Orbital Trajectory and Attitude Dynamics

Altman [18] has proposed a unified state model of the orbital trajectory and attitude dynamics of an orbital spacecraft. The synthesis and analysis of the orbital trajectories of spacecraft are usually conducted separately from the attitude dynamics. This is due to the difficulty in finding dynamical models valid in the analytical sense and efficient for machine computation. However, the interdependence between the trajectory dynamics of point-mass motion and the attitude dynamics of spacecraft body orientation about that point-mass makes the use of a unified dynamical model essential. This interdependence is due to natural forces such as gravity gradient and atmospheric drag. It is also due to artificial forces and energy interchange between the two classes of motion.

To develop the model a set of dynamical variables is necessary. Many sets are available for use in aerospace systems and trajectory analysis. The fundamental set is the six elements vector $(x, y, z, \dot{x}, \dot{y}, \dot{z})$, that gives the components of the position and velocity vectors of the spacecraft. These are called the "cartesian variables" compared to the unified state model elements used in this work.

In the case of the unified state model the elements of the state vector should define the trajectory and attitude dynamics in a common form, therefore, this model employs seven variables for trajectory dynamics instead of the minimum definitive number of six. The first three state variables are parametric forms of the orbital momenta. The other four are the Euler parameters, a four-dimensional representation of the rotation transformation for a coordinate frame-triad.

The main advantage of this model is that strong interpendence or coupling between the first three state variables and the other four does not exist. So, filtering and estimation can be accomplished more rapidly and accurately by exploitation of this functional independence between momenta and coordinate variables. In his work Altman states that the unified state model is extremely accurate for reentry dynamics computation and unified applications as in-orbit and reentry operations.

Moreover, the orbital velocity state and the position state can be obtained by algebraic transformations from the unified state model variables.

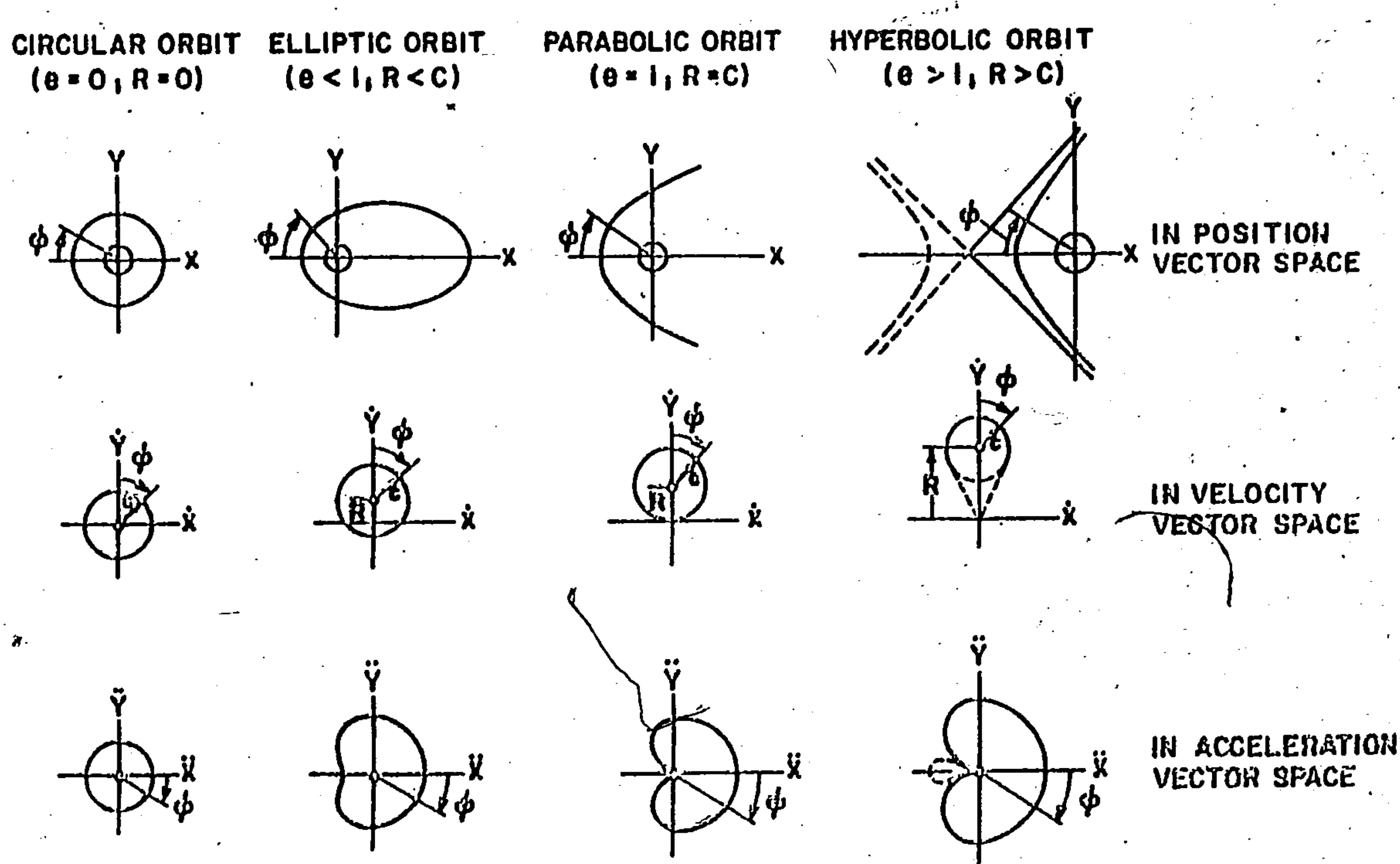
5.2.1 The Dynamical Variables

A definitive set of dynamical variables will comprise the state variables and coordinate variables. The state variables may be position, velocity, or other variables which relate to the system energy. The coordinate variables may be three-parameter sets (Euler angles), four-parameter sets (Euler parameters), or nine-parameter sets (direction cosines of a rotating triad). In the unified state model the state variables are parametric forms of the momenta, and the coordinate variables are the Euler parameters.

An unperturbed orbital trajectory, occurring in the presence of the simple spherical harmonic function of gravity field due to one celestial body, is represented by cyclic figures (lying in an orbital plane) in position, velocity and acceleration vector spaces as shown in Fig. 5.1. The orbital figure is a conic in position space, a circle in velocity space, and a limacon-like figure in acceleration space. However, as the orbital energy level changes, only the velocity space map remains invariant in geometric figure. Consequently, a differential formulation of the orbital trajectory dynamics - expressing the orbital state resulting from impressed or perturbing forces - will not encounter singularities in the state variables; that is, these velocity parameter variables are regularized, with consequent advantages in computation.

Figure 5.1

State maps of an orbit



The orbital trajectory state variables are the velocity state parameters (C,R). They are defined as functions of the radial momentum ($P_r = mv_1$) and angular momentum ($p_\lambda = mrv_2$) as follows

$$C = \mu m/p_\lambda \tag{5.1}$$

$$R = [(p_r/m)^2 + (p_\lambda/mr - C)^2]^{1/2}$$

$$= [2S+C^2]^{1/2} \tag{5.2}$$

where

μ = planetary gravitational constant

m = mass of orbital body

S = orbital energy per unit mass.

That is, the state parameters (C,R) are implicit forms of the orbital momenta. The velocity components v_1 and v_2 are shown in Fig. 5.2.

The complete three-dimensional vector equations of position, velocity and acceleration are defined as functions of (C,R) as follows

$$r = [\frac{\mu}{C(C+R\cos\phi)}]\, e^{i\phi} \tag{5.3}$$

$$v = R + Cr/|r| \tag{5.4}$$

$$a = \dot{R} + Cr/|r| + a_b$$

$$|a_b| = -C\,\dot{\lambda} \tag{5.5}$$

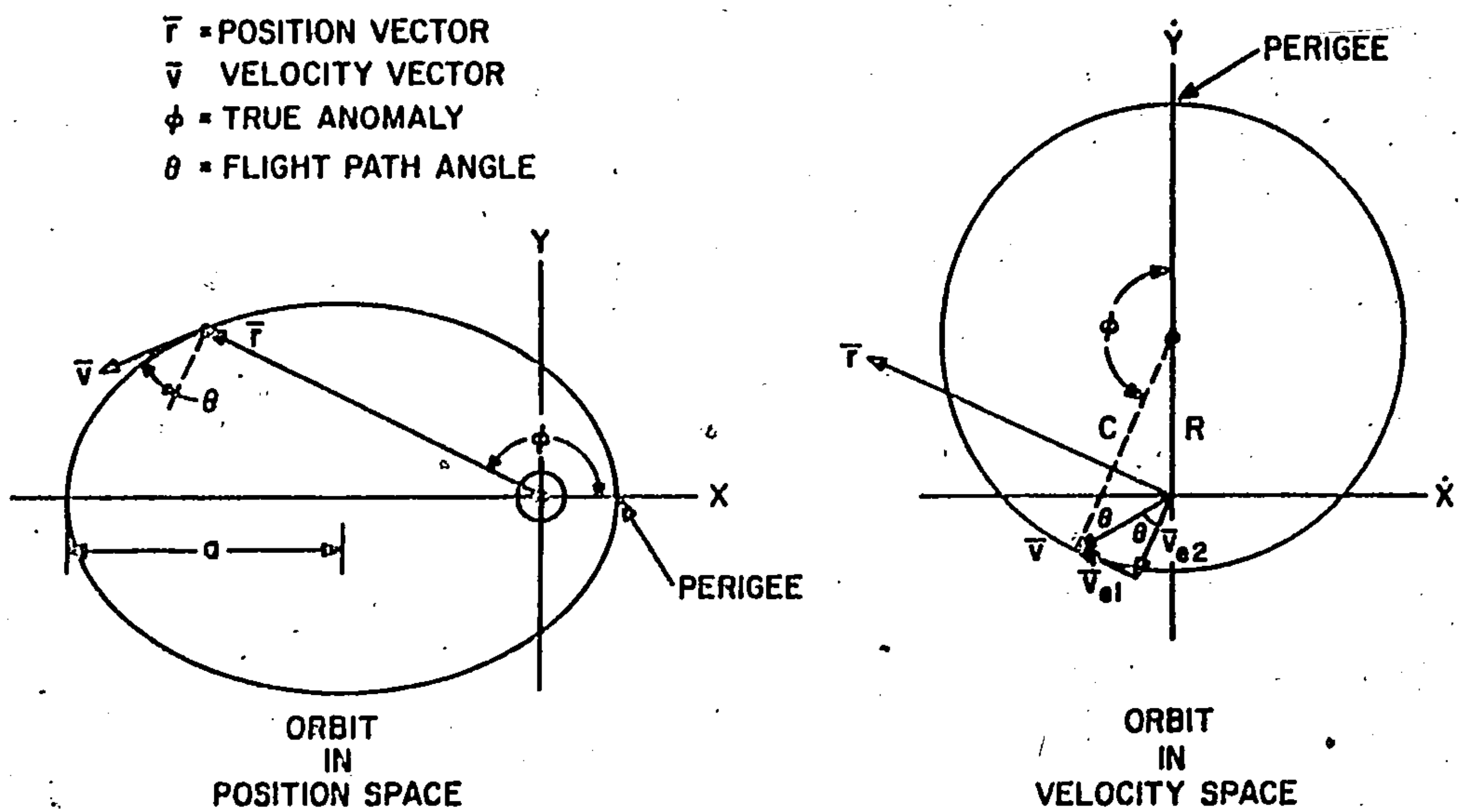


Figure 5.2

Velocity space mapping of an orbit

where C, R, r and λ are defined as shown in Fig. 5.3. In this figure a sphere is taken as the model of the Earth. The fundamental plane is the Earth's equator, the x axis points at the vernal equinox. Three coordinate sets are selected for this model. The planetocentric inertial set (g_1, g_2, g_3) defines the coordinate axes (X, Y, Z) fixed in inertial space about the origin. The intermediate set (f_1, f_2, f_3) defines the coordinate axis in the instantaneous orbital plane, with f_1 directed along the intermediate axis X', f_3 along the orbital angular momentum vector and f_2 completing the right-handed set. The axis X' is defined by rotation of the planetocentric inertial axis X into the instantaneous orbital plane about the lines-of-nodes LN. The rotating polar set (e_1, e_2, e_3) defines the coordinate axis rotating with the orbital body, with e_1 along the position vector r, e_2 normal to e_1 in the instantaneous plane and e_3 completing the right-handed set. Consequently e_1 defines the direction of $\dot{r}$ and e_3 (or f_3) defines the direction of C.

The orbital state variables of the unified state model are the three parameters (C, R_1, R_2) which define the velocity state. The component pair (R_1, R_2) defines both R and the direction angle of perigee apsis from the rotated axis X'. These three variables represent the state

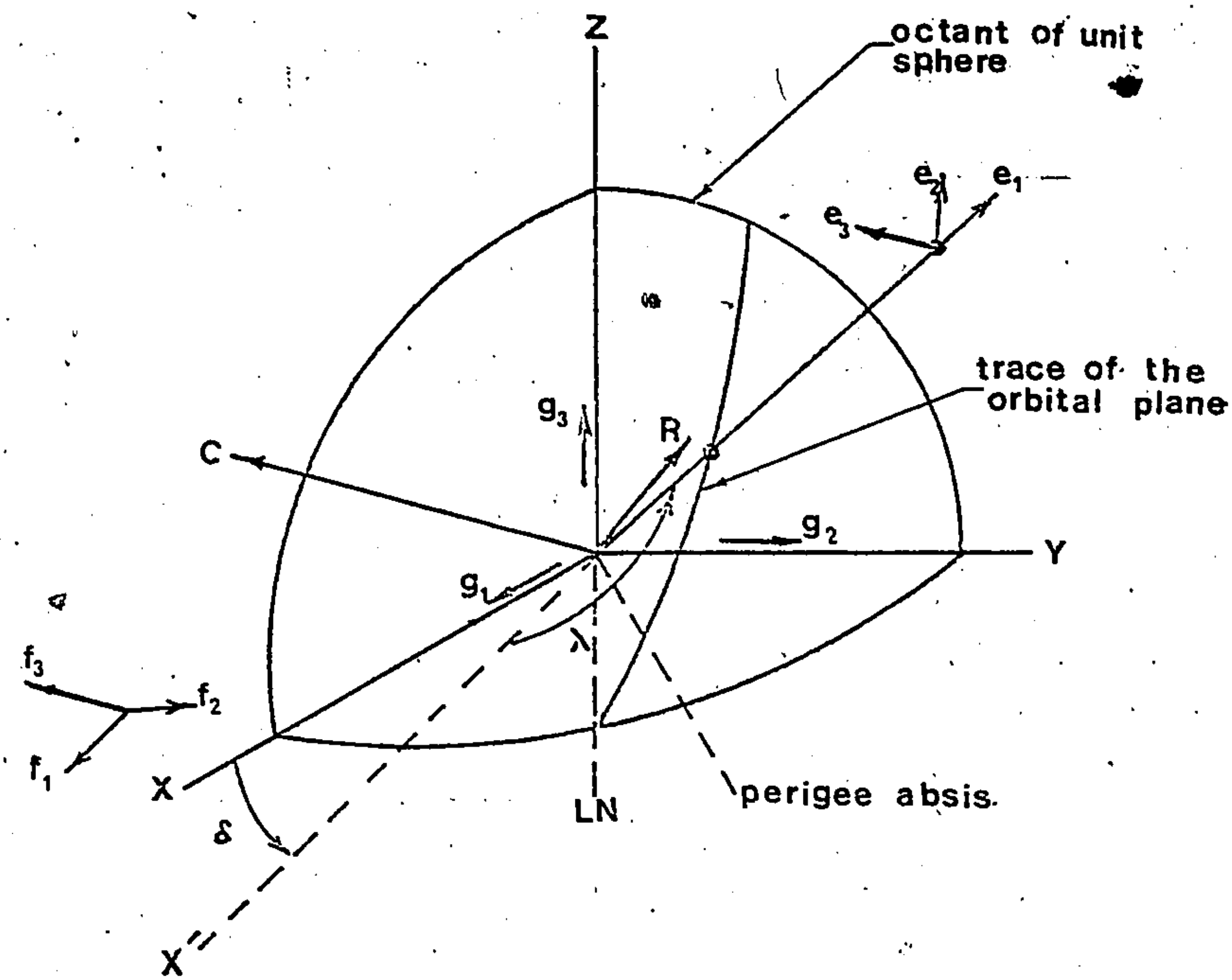


Figure 5.3

State model geometry of an orbital trajectory variables in the dynamical model, the other four are the coordinate variables or the Euler parameters.

The Euler parameters and the consequent constraint equations offer many advantages in computation [59]. The four Euler parameters (ℓ_1, ℓ_2, ℓ_3, ℓ_4) define a unitary quaternion which can be employed to describe the rotation of a coordinate triad [Appendix B]. Consequently, the coordinate rotation of orbital motion and of body motion about its point-mass can be expressed by one matrix form of rotation transformation. That is, a unified form for both attitude and trajectory dynamics.

5.2.2 The Dynamical Model

Our main concern in this work is the problem of orbit determination, so only trajectory dynamics are considered in the following analysis.

The constraint equations for the orbital trajectory dynamics are given as first-order differential equations, functions of the perturbing accelerations a_i [18]. Only the primary gravity harmonic forces in the trajectory dynamics are modeled in this thesis.

The orbital trajectory dynamics are defined by the orbital state:

$$\frac{d}{dt}\begin{bmatrix} C \\ R_1 \\ R_2 \end{bmatrix} = \begin{bmatrix} 0 & -p & 0 \\ \cos\lambda & -(1+p)\sin\lambda & -R(\gamma/v_2) \\ \sin\lambda & (1+p)\cos\lambda & R(\gamma/v_2) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} \qquad (5.6)^*$$

and the orbital coordinates

$$\frac{d}{dt}\begin{bmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 0 & w_3 & 0 & w_1 \\ -w_3 & 0 & w_1 & 0 \\ 0 & -w_1 & 0 & w_3 \\ -w_1 & 0 & -w_3 & 0 \end{bmatrix} \begin{bmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{bmatrix} \qquad (5.7)$$

where

$$p = C/v_2$$

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ C \end{bmatrix} + \begin{bmatrix} \cos\lambda & \sin\lambda \\ -\sin\lambda & \cos\lambda \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$

$$\begin{bmatrix} \sin\lambda \\ \cos\lambda \end{bmatrix} = \frac{1}{(\ell_3^2 + \ell_4^2)} \begin{bmatrix} 2\,\ell_3\,\ell_4 \\ \ell_4^2 - \ell_3^2 \end{bmatrix}$$

$$w_1 = a_3/v_2$$

$$w_3 = C\,v_2^2/\mu$$

$$r = \frac{\ell_1\ell_3 - \ell_2\ell_4}{\ell_3^2+\ell_4^2} = \frac{\xi_{13}}{1+\xi_{33}}$$

---

* Equation (5.6) is the corrected version of Altman's equations, as pointed out recently by Paul Chodas of the University of Toronto Institute for Astronautic Studies. Unfortunately, these corrections came too late to be incorporated in this work, but it is also believed that their effect on this work is rather insignificant.

$$[a] = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \text{sum of perturbing accelerations}$$

$$\ell_1^2 + \ell_2^2 + \ell_3^2 + \ell_4^2 = 1 \qquad (5.8)$$

The perturbing accelerations may be due to drag and wind forces, zonal harmonics, tesseral harmonics and other effects. One of the major sources of perturbations is the nonsymmetry of the Earth, which has the tendency to produce a noncentral force field. The departures, due to asymmetry from a concentrically homogeneous sphere, are of two kinds, zonal departures and tesseral departures. Zonal harmonics are due to meridian ellipticity, while tesseral harmonics are due to longitudinal variations in the shape of the Earth. In the development of the model equations only zonal harmonics will be considered, which is usually the case in general perturbation techniques [60].

From equation (5.6) we note that the acceleration state is a direct function of in-plane perturbations (a_1, a_2, a_3). We also note that the reference from rotation defined by equation (5.7) is a direct function of out-of-plane perturbation (a_3) alone. Consequently, strong interdependence or coupling between these two matrix equations is not present. As mentioned previously this functional independence is advantageous for efficient data processing.

The zonal harmonics are described by

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = -\frac{3}{2}\mu R_e^2 J_2 \left(\frac{Cv_2}{\mu}\right)^4 \begin{bmatrix} 1-3\,\xi_{13}^2 \\ 2\,\xi_{13}\,\xi_{23} \\ 2\,\xi_{13}\,\xi_{33} \end{bmatrix} \tag{5.9}$$

where

R_e = mean equatorial radius of the Earth.

J_2 = mean zonal harmonic of Earth's gravity field

$$\xi_{13} = 2\,(\ell_1\,\ell_3 - \ell_2\,\ell_4)$$
$$\xi_{23} = 2\,(\ell_2\,\ell_3 + \ell_1\,\ell_4)$$
$$\xi_{33} = 1 - 2\,(\ell_1^2 + \ell_2^2) \tag{5.10}$$

From equations (5.7) and (5.8) we have

$$\frac{dC}{dt} = -p\,a_2$$

$$\frac{dR_1}{dt} = a_1 \cos\lambda - (1+p)\,a_2 \sin\lambda$$

$$\frac{dR_2}{dt} = a_1 \sin\lambda + (1+p)\,a_2 \cos\lambda$$

$$\frac{d\ell_1}{dt} = \frac{1}{2}\,(w_3\,\ell_2 + w_1\,\ell_4)$$

$$\frac{d\ell_2}{dt} = \frac{1}{2}\,(-w_3\,\ell_1 + w_1\,\ell_3)$$

$$\frac{d\ell_3}{dt} = \frac{1}{2}\,(-w_1\,\ell_2 + w_3\,\ell_4)$$

$$\frac{d\ell_4}{dt} = \frac{1}{2}\,(-w_1\,\ell_1 - w_3\,\ell_3) \qquad (5.11)$$

From equations (5.9) and (5.10) we have

$$a_1 = K\,C^4\,v_2^4\,[1 - 12\,(\ell_1\,\ell_3 - \ell_2\,\ell_4)^2]$$

$$a_2 = K\,C^4\,v_2^4\,[8\,(\ell_1\,\ell_3 - \ell_2\,\ell_4)\,(\ell_2\,\ell_3 + \ell_1\,\ell_4)]$$

$$a_3 = K\,C^4\,v_2^4\,4\,(\ell_1\,\ell_3 - \ell_2\,\ell_4)\,[1 - 2\,(\ell_1^2 + \ell_2^2)] \qquad (5.12)$$

where

$$K = -\frac{3}{2}\,Re^2\,J_2/\mu^3 = \text{constant}$$

By substituting equations (5.8) and (5.9) into equation (5.11) we can derive the nonlinear system model in the general form

$$\dot{x}(t) = f(x,t)$$

where the seven-dimensional vector x is defined as

$$[x] = [C, R_1, R_2, \ell_1, \ell_2, \ell_3, \ell_4]$$
$$= [x_1, x_2, \ldots, x_7] \qquad (5.13)$$

After a considerable amount of algebraic manipulations the following nonlinear differential equations are obtained

$$\dot{x}_1 = F\ x_1$$

$$\dot{x}_2 = F_1 \cos\lambda + F_2 \sin\lambda$$

$$\dot{x}_3 = F_1 \sin\lambda - F_2 \cos\lambda$$

$$\dot{x}_4 = F_3\ x_5 + F_4\ x_7$$

$$\dot{x}_5 = -F_3\ x_4 + F_4\ x_6$$

$$\dot{x}_6 = F_3\ x_7 - F_4\ x_5$$

$$\dot{x}_7 = -F_3\ x_6 - F_4\ x_4 \tag{5.14}$$

where

$$F = -8\ K\ x_1^4\ v_2^3\ (x_4\ x_6 - x_5\ x_7)\ (x_5\ x_6 + x_4\ x_7)$$

$$F_1 = K\ x_1^4\ v_2^4\ [1 - 12\ (x_4\ x_6 - x_5\ x_7)^2]$$

$$F_2 = F\ (v_2 + x_1)$$

$$F_3 = x_1\ v_2^2/2\mu$$

$$F_4 = 2\ K\ x_1^4\ v_2^3\ (x_4\ x_6 - x_5\ x_7)\ (1 - 2\ x_4^2 - 2x_5^2)$$

and

$$v_2 = x_1 - x_2 \sin \lambda + x_3 \cos \lambda$$

$$\sin \lambda = 2\, x_6\, x_7 \,/\, (x_6^2 + x_7^2)$$

$$\cos \lambda = x_7^2 - x_6^2 \,/\, (x_6^2 + x_7^2) \qquad (5.15)$$

and the constant K is defined previously.

In reference [18] Altman gives the equations for perturbations due to higher order zonal harmonics (J_3 and J_4) as well as tesseral harmonics and atmospheric drag and wind perturbations. In the model derivation outlined above only J_2 zonal harmonics are considered.

5.3 Measurement (Observation) Model

In order to use the state model, previously developed, in aerospace applications, a measurement model will be developed in this section.

The introduction of radar into the implements of modern science has produced a great variation from established techniques in orbit determination schemes. Now, the orbit determiner could measure the distance or extension between the point of observation and the satellite, that is, the slant range. Furthermore, angular data are provided, such as Azimuth and Elevation.

Usually, observations are made from a coordinate system that is rotating or is different from the preferred or inertial system in which analysis is to be performed. Thus a transformation of coordinate is required. The satellite position and velocity, or in general the state variables, must be referred to the geocenter and its inertial coordinate set as reference.

Figure 5.4 shows the Azimuth-Elevation coordinate system. In this system, the observer is at the origin of the coordinate system and the fundamental plane is the local horizon, that is a topocentric system. In this figure, the two angles needed to define the location of an object along some ray from the origin are defined as follows

γ (Elevation): Angular elevation of an object above a tangent plane to the observer's position

A (azimuth): The angle from the North to the object's meridian, measured in the tangent plane to the observer's meridian, at the observer's position.

The distance from the observer to the object is

h (slant range): Distance between the origin of the coordinate system and the location of a point (object) within the coordinate system

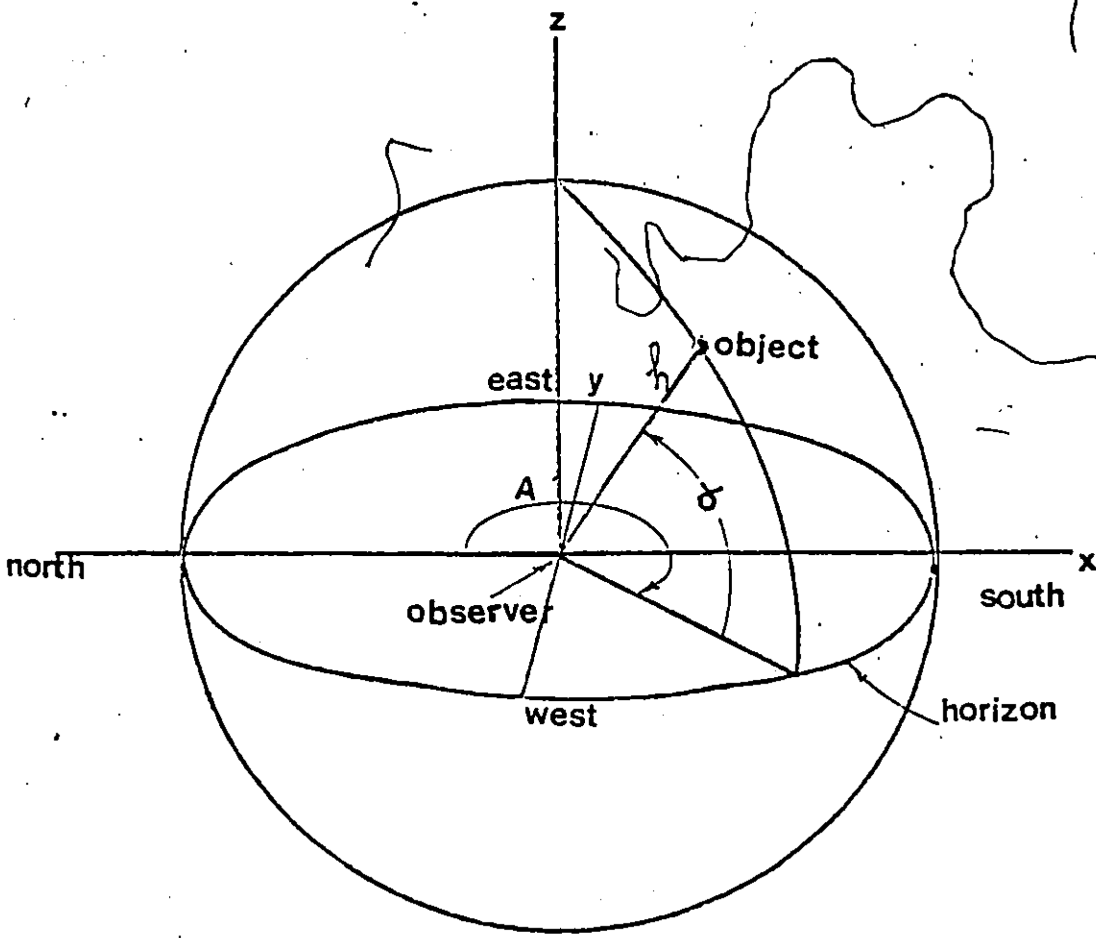


Figure 5.4

Azimuth-Elevation coordinate system

The following convention is usually adopted [59]

Variable	Sense	Range
γ	Positive above horizon	$-90^\circ \leq \gamma \leq 90^\circ$
A	Positive to East from North	$0^\circ \leq A \leq 360^\circ$
h	Positive length	$h \geq 0$

To convert polar coordinates (γ, A, h) into rectilinear coordinates we use the conventional transformations, given in reference [60], as follows:

$$x = -h \cos A \cos \gamma$$
$$y = h \sin A \cos \gamma$$
$$z = h \sin \gamma \qquad (5.16)$$

where (x, y, z) define the observer's coordinate system and is shown in Fig. 5.5. The inertial coordinate system is defined by (X, Y, Z).

Introduce the unit vectors, N pointing to North, E pointing to East, and U pointing vertically outward along the normal from the observer's meridian, and by projective

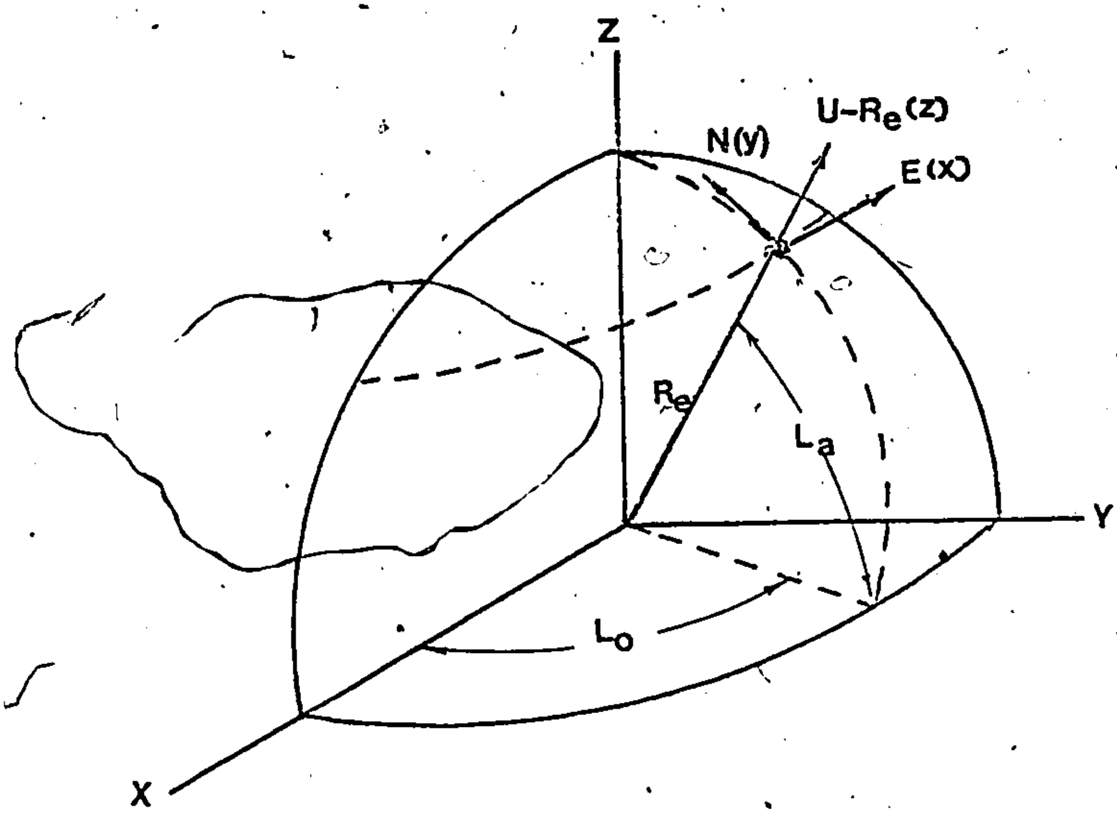


Figure 5.5

Coordinate set for a ground-based site

principles, we obtain

$$N = \begin{bmatrix} N_x \\ N_y \\ N_z \end{bmatrix} = \begin{bmatrix} \sin L_a \sin L_o \\ -\sin L_a \cos L_o \\ \cos L_o \end{bmatrix}$$

$$E = \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} = \begin{bmatrix} \cos L_o \\ \sin L_o \\ 0 \end{bmatrix}$$

$$U = \begin{bmatrix} U_x \\ U_y \\ U_z \end{bmatrix} = \begin{bmatrix} -\sin L_o \cos L_a \\ \cos L_o \cos L_a \\ \sin L_a \end{bmatrix} \tag{5.17}$$

The above equations give the direction cosines of a vector in the Azimuth-Elevation coordinate system with respect to the inertial system. Hence, if L is a unit vector along the slant range we have

$$L = \begin{bmatrix} L_x \\ L_y \\ L_z \end{bmatrix} = \begin{bmatrix} N_x & E_x & U_x & L_x \\ N_y & E_y & U_y & L_y \\ N_z & E_z & U_z & L_z \end{bmatrix} \tag{5.18}$$

From equations (5.17) and (5.18) we have

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \sin L_a \sin L_o & -\sin L_a \cos L_o & \cos L_a \\ \cos L_o & \sin L_o & 0 \\ -\sin L_o \cos L_a & \cos L_o \cos L_a & \sin L_a \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (5.19)$$

where L_a = Geodetic Latitude

L_o = Longitude

To proceed with the model derivation a transformation from position vector components (X, Y, Z) into the unified state model variables (C, R_1, R_2, ℓ_1, ℓ_2, ℓ_3, ℓ_4) is necessary. The following transformation is used [18]

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [E]^T \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} \quad (5.20)$$

where [E] is the transformation matrix used when a coordinate triad is rotated from an inertial reference triad [Appendix B]

$$[E] = \begin{bmatrix} \xi_{11} & \xi_{12} & \xi_{13} \\ \xi_{21} & \xi_{22} & \xi_{23} \\ \xi_{31} & \xi_{32} & \xi_{33} \end{bmatrix} \quad (5.21)$$

where ξ_{ij} is given as a function of the Euler parameters (see Appendix B).

Substituting for [E] we get

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = r\ [D] \begin{bmatrix} 1 - 2\ (\ell_2^2 + \ell_3^2) \\ 2\ (\ell_1\ \ell_2 + \ell_3\ \ell_4) \\ 2\ (\ell_1\ \ell_3 - \ell_2\ \ell_4) \end{bmatrix} \tag{5.22}$$

where

$r = \mu/C\ v_2$

[D] = constant matrix for known values of angles L_a and L_o, that is, independent of the states and observations.

Finally from the transformations (5.16) we have

$$\tan A = x/y$$

$$\sin \gamma = z/h$$

$$h^2 = x^2 + y^2 + z^2 \tag{5.24}$$

Define the observation vector ν as the three-dimensional vector

$$[\nu] = [\gamma,\ h,\ A] \tag{5.25}$$

For given latitude and longitude and using equations (5.23) and (5.24) we can express the observations as nonlinear functions of the unified state model variables as follows

$$\begin{bmatrix} \gamma \\ h \\ A \end{bmatrix} = \begin{bmatrix} H_1 \ (x_1, x_2, \ldots, x_7) \\ H_2 \ (x_1, x_2, \ldots, x_7) \\ H_3 \ (x_1, x_2, \ldots, x_7) \end{bmatrix} \tag{5.26}$$

which corresponds to the discrete-time nonlinear measurement model used for state estimation with the message model and given by

$$\nu(k) = H\ [x(k), k] + \text{measurement errors}$$

The observations may be perturbed by several different sources of noise such as calibration errors, readout errors, tracking errors, etc.

When the Kalman filter solution is applied to a practical orbit determination problem, a knowledge of the a priori statistics of measurement noise is necessary. So, it is possible to guess a priori statistics for use in the Kalman filter, then examine the residuals $\{\nu(k)-H[\hat{x}(k),k]\}$, from the resulting solution, which in a sense constitute estimates of the observation errors, to produce a better guess as to the unknown statistics, and then iterate the entire process. But, this reprocessing of all the observation data is quite often impractical, especially in real-time simulations.

Therefore, a more practical solution to this problem is the use of adaptive filtering techniques. The adaptive filter algorithm suggested by Sage and Husa [Chapter 2] was applied by the author with the unified state model, using actual records of observations, supplied by the Communications Research Centre in Ottawa. Simulation results prove the fast convergence of the filter.

Another adaptive scheme suggested by Sinha and Tom [61] was applied to the orbit determination problem when the measurement are contaminated with unknown coloured measurement noise [57].

The unified state model and the derived observation model are used in the simulation study given in Chapter 6 for nonlinear state estimation.

CHAPTER 6

COMPARISON OF FOUR NONLINEAR ESTIMATION SCHEMES WITH APPLICATIONS TO A COMMUNICATIONS SATELLITE

6.1 Introduction

The problem of estimating the orbital states of a synchronous-orbit communications satellite from ground-based measurements, which are contaminated with noise, is complicated by the nonlinearity of the dynamic equations as well as lack of prior knowledge of the noise statistics.

The purpose of this chapter is to compare four distinct schemes for recursive estimation of the state variables of a continuous-time nonlinear system on the basis of measuring the discrete-time outputs of the system in the presence of noise. The numerical characteristics of the four nonlinear filtering algorithms are illustrated with a realistic orbit determination study.

The schemes considered for nonlinear state estimation are:

(1) The extended Kalman filter
(2) The second-order suboptimal nonlinear filter
(3) The nonlinear filter based on invariant imbedding
(4) The combined invariant imbedding and stochastic approximation algorithm.

These algorithms are given in detail in Chapters 3 and 4 and the related references.

To represent the orbital trajectory dynamics the unified state model, proposed by Altman [Chapter 5], is utilized. The model consists of seven state variables, three of them are the components of the orbital velocity and the other four are coordinate variables. For simulation study a conversion routine is used to perform conversion of state vector formats from unified state model (USM) to classical inertial coordinates, that is position and velocity components.

The measurement model has been derived in Chapter 5. The observations, which are slant range, Azimuth and Elevation, are nonlinearily related to the states of the unified state model.

The simulation study is carried out using the data supplied by the Spacecraft Mechanics Division of the Communications Research Centre in Ottawa, which contain actual records of observations. Radar measurements for azimuth elevation and range are contained in data records covering 24 hours.

The results using the proposed invariant imbedding-stochastic approximation algorithm are compared with those obtained using the extended Kalman filter, the second-order filter and the invariant imbedding approach. In the case studied the new algorithm is proven to be more efficient than the other techniques; it requires less computation time to achieve convergence, which makes it practical for real-time operations.

6.2 Problem Formulation

The problem of trajectory determination may be summarized as follows: Given the orbital trajectory dynamics of the satellite in the form,

$$\dot{x}(t) = f(x,t) \tag{6.1}$$

the trajectory of the spacecraft can be specified uniquely from a knowledge at any time of the three components of its position vector and three components of its velocity vector in an orthogonal reference frame. In the case of the USM these six components are replaced by the seven USM variables. These variables define the state vector which is a continuous-time function generated by integration of equation (6.1) with appropriate initial conditions. Because the initial conditions are not known precisely, the present state is also not known, and it is the function of the trajectory determination system to estimate the state on the basis of observations. The system is then regarded as a multidimensional filter, its input being a time sequence of observations of variables related to the state, corrupted by additive errors. Its output is the estimate of the state at present time, and the filter is to be designed to make this an optimal estimate in the sense of minimizing some function of the estimation error.

The initial conditions are not known, but it is assumed that they can be described probabilistically at least up to second-order statistics. Thus, the initial conditions are regarded as a vector valued random variable with mean x_o and

$$\text{cov } \{x_o, x_o\} = P \tag{6.2}$$

The observation model is given by:

$$y(k) = h\ [x(k), k] + v(k) \tag{6.3}$$

The measurement noise v(k) is assumed to be zero mean white noise with

$$\text{cov } [v(k), v(k)] = R(k)\ S_K(k-j) \tag{6.4}$$

In some problems a plant noise is assumed to be associated with the system model in the form

$$x(k) = x(k-1) + u(k) \tag{6.5}$$

where $x(k-1) = x(t_k)$, which is the solution at time t_k to the nonlinear differential equation (6.1). The noise sequence u(k) is assumed to be a zero-mean white Gaussian noise with covariance Q(k).

In any nonlinear filter algorithm the computation of the gain matrix K and the error covariance matrix P utilizes derivatives of $f(\cdot)$ and $h(\cdot)$ with respect to the state $\hat{x}$. The estimation algorithms considered in this study are either first-order or second-order filters depending on the order of linearization.

The linearization will always be around the estimated rather than the reference trajectory. This is clearly the correct procedure since P has to do with the difference between the estimate and the true state, and the estimate is on the average closer to the true state than the reference. Errors arising from the linearization assumptions are thereby minimized.

Covergence of the filter algorithm depends on nonlinearity effects, assumed noise statistics and initial assumptions. So, one has to resort to experimental studies to obtain the best performance. In this simulation study we compare the performance of different filters on the basis of the error covariance matrix P or the estimation errors.

6.3 Description of the Computer Program

The four nonlinear estimation algorithms mentioned previously were programmed for computer simulation using Fortran IV and have been run on the CDC 6400 computer system. Results of state estimation of the satellite are obtained for different initial conditions and are given in the next section.

A flowchart of the general program is shown in Fig. 6.1. The program is started by setting the initial parameters which include the initial state vector x_o and the initial error covariance matrix P_o. When an observation has been made and is to be processed, the input data are read, then integration of the model equations is initiated.

In the case of the extended Kalman filter and the invariant imbedding nonlinear filter it is the state transition matrix which is integrated forward and then used to compute the extrapolated covariance matrix as follows

$$P(k|k-1) = \phi(k|k-1)\ P(k-1|k-1)\ \phi^T(k|k-1) + Q(k) \tag{6.7}$$

where $\phi(k|k-1)$ satisfies the matrix differential equation

$$\frac{d}{dt}\ \phi(t,\ t_{k-1}) = \left.\frac{\partial f(x,t)}{\partial x(t)}\right|_{\hat{x}(t)} \phi(t,\ t_{k-1}) \tag{6.8}$$

with initial condition $\phi(t_{k-1},\ t_{k-1}) = [I]$.

In the case of the second order filter, the state equation and covariance equation are coupled and are both integrated forward as follows

$$P(k|k-1) = P(t_k^-) + Q(k)$$

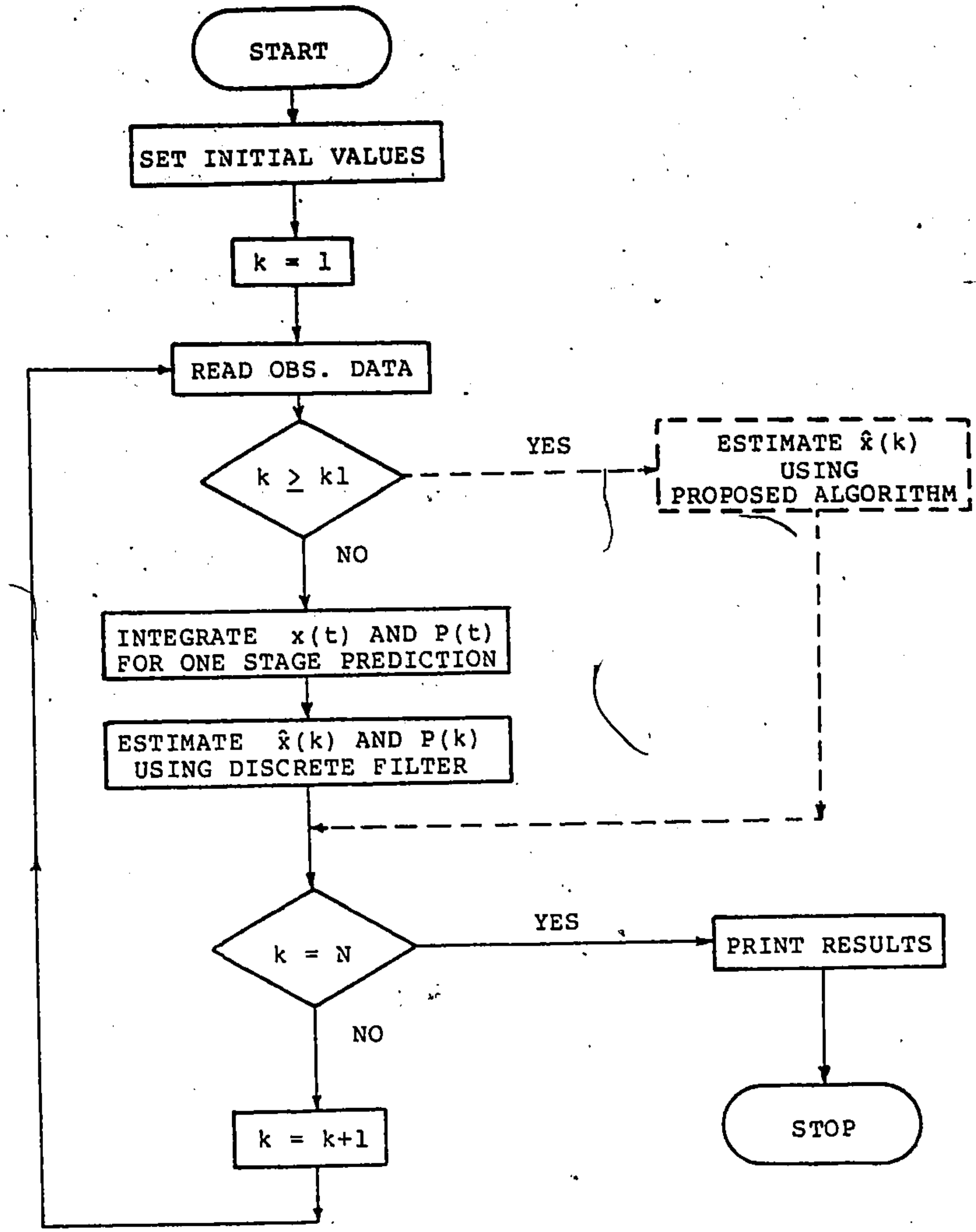


Figure 6.1

Program flowchart

where $P(t_k^-)$ is the solution of the differential equation

$$\dot{P}(t) = \left.\frac{\partial f(x,t)}{\partial x(t)}\right|_{\hat{x}(t)} P(t) + P(t) \left.\frac{\partial f^T(x,t)}{\partial x(t)}\right|_{\hat{x}(t)} \tag{6.9}$$

and $\hat{x}(k|k-1)$ is the solution at time t_k of

$$\dot{x}(t) = f(x,t) + \frac{1}{2} \sum_{i=1}^{n} \pi_i \text{ tr } [A(t)\ P(t)] \tag{6.10}$$

These equations are given in detail with the filter algorithms in Chapter 4.

The routine used for numerical integration is DVOGER which uses a predictor corrector method.

The estimation routine provides the discrete updating of the state estimate and the error covariance matrix. A priori noise statistics, i.e., noise covariance matrices Q and R are necessary for state estimation using the extended Kalman filter and the second-order filter. A block diagram of the extended Kalman filter is shown in Fig. 6.2.

To achieve increased accuracy in the computations, the nonlinear functions $f(\cdot)$ and $h(\cdot)$ are approximated to second-order in the Taylor's series expansion and the second-order filter is derived. The block-diagram of the second-order filter is similar to that of the first-order filter with two more routines to compute the second-order derivatives of the two nonlinear functions $f(\cdot)$ and $h(\cdot)$.

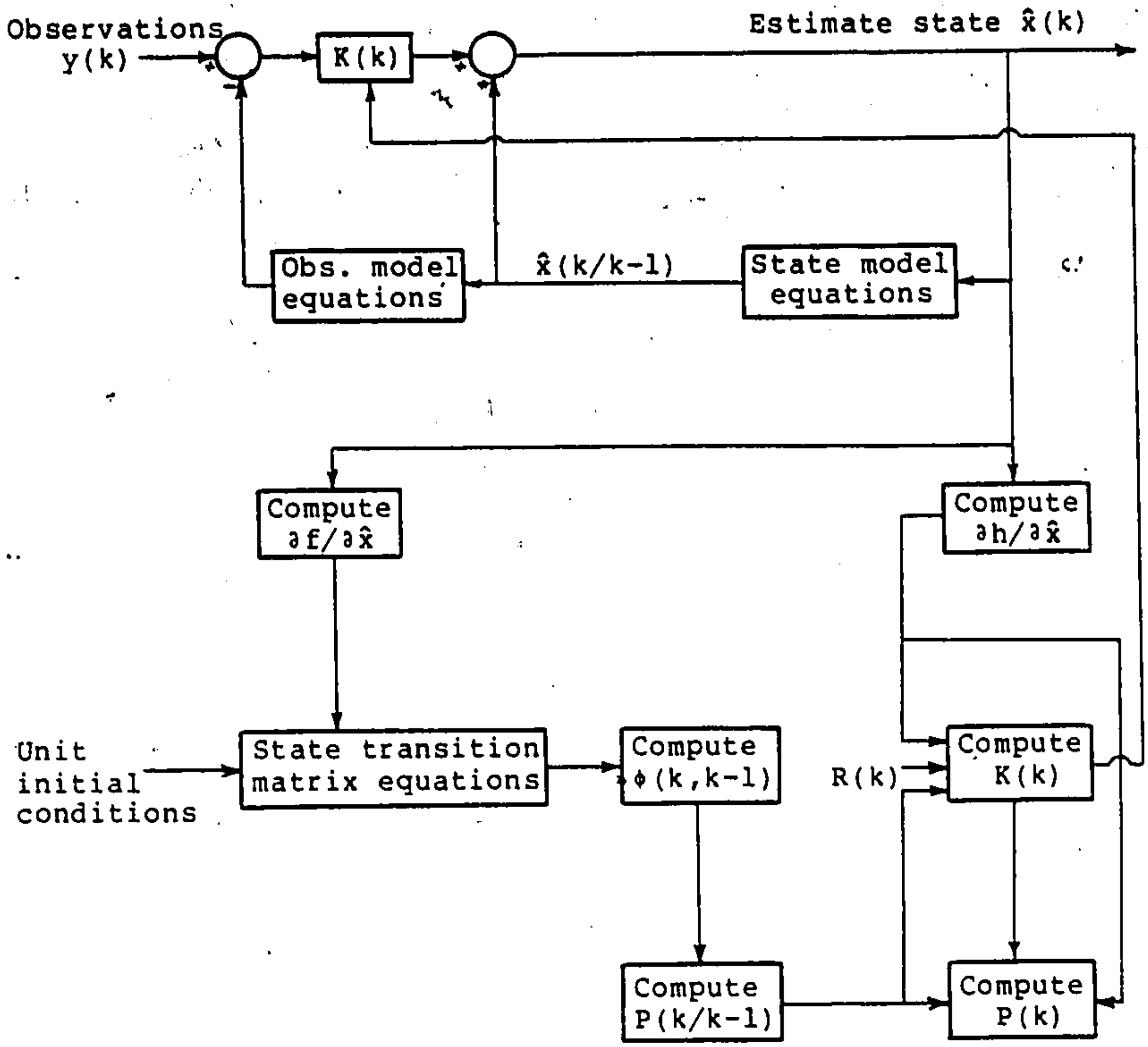


Figure 6.2

Block diagram of the

extended Kalman filter algorithm

The invariant imbedding algorithm computations do not require a specific knowledge of Q and R. Although the noise covariance matrices appear in the filter equations [Chapter 4] as weighting functions, it is proved experimentally that the filter performance is not sensitive to the values of Q and R. A block diagram similar to Fig. 6.2 can represent the filter with an additional routine to compute the second derivative of the observation function h(•).

At the end of each iteration results of the state estimates and the error covariance matrix are printed. A conversion routine is used to give the estimates of the satellite position and velocity at each iteration using

$$r = \mu/C \, v_2$$

$$v = \sqrt{v_1^2 + v_2^2} \tag{6.11}$$

defined in Chapter 5.

The above process is repeated with each new observation until k = N, where N is the total number of observations.

The proposed algorithm follows a different sequence of computations compared to the other three algorithms, as shown in Fig. 6.1. The invariant imbedding algorithm is utilized for the first k_1 observations, where $k_1 << N$ and is determined experimentally. Then, the stochastic approximation gain algorithm given by

$$K(k+1) = K(k) + \frac{a}{k+b} \frac{\hat{x}(k|k-1)\ \xi^T(k)}{||x(k|k-1)\ \xi^T(k)||^2} \tag{6.12}$$

is applied. The constants a and b are determined experimentally to give the best performance.

6.4 Simulation Results

The simulations utilizing the four nonlinear estimation schemes described previously were performed in conjunction with the mathematical system and observations models given in Chapter 5.

Observations of range, azimuth and elevation are obtained from real tracking data supplied by the Communications Research Centre. The sampling rate of the observations is 10 sec. Successive observations covering 470 sec are chosen for the simulation.

The following physical constants are used for all the computations

μ = gravitational constant = 398600 km^3/sec^2

R_e = mean equatorial radius = 6378.14 km

J_2 = zonal harmonic of earth's gravity field = 0.108265×10^{-2}

$$\left.\begin{matrix} C_{22} \\ \\ S_{22} \end{matrix}\right\} = \text{Tesseral harmonic of earth's gravity field} = \begin{cases} 0.1566511 \times 10^{-5} \\ \\ -0.8869932 \times 10^{-6} \end{cases}$$

To initiate the estimation procedure, the initial guess for $\hat{x}_o$ and P_o is required. The initial error covariance is assumed to be a diagonal matrix with equal variances and zero covariances. Results will be given for two values of error variances

(1) $\sigma_o^2 = 10^{-4}$

(2) $\sigma_o^2 = 10^{-2}$

The initial state vector $\hat{x}_o$ is defined by the initial values of the USM variables (C, R_1, R_2, ℓ_1, ℓ_2, ℓ_3, ℓ_4). Two cases of initial orbit trajectory are considered.

a. Initial circular orbit

The initial orbit is determined by the following values:

$$C = 3.074656 \text{ km/sec}$$

$$R_1 = 0 \text{ km/sec}$$

$$R_2 = 0 \text{ km/sec}$$

which means that the initial orbit is circular as the velocity components are (see Fig. 5.2)

$$v_1 = 0 \text{ km/sec}$$

$$v_2 = C = 3.074656 \text{ km/sec}$$

The initial Euler parameters are [Appendix B]

$$\ell_1 = \sin i/2 \cos (r-\nu)/2$$

$$\ell_2 = \sin i/2 \sin (r-\nu)/2$$

$$\ell_3 = \cos i/2 \sin (r+\nu)/2$$

$$\ell_4 = \cos i/2 \cos (r+\nu)/2$$

where

$$i = 0.9^\circ$$

$$r = 240^\circ$$

$$\nu = -\Omega$$

To determine the true orbit the system model given by equation (6.1) is integrated numerically using a fourth-order Runge-Kutta integration algorithm.

We choose the magnitude of the position and velocity estimation errors as a measure of estimation accuracy. The trace of the error covariance matrix gives an indication of the rate of convergence for each filter.

For initial error variances, $\sigma_o^2 = 10^{-4}$, the magnitude of the actual estimation errors in the satellite position $|\tilde{r}|$, for the various algorithms, is plotted in Fig. 6.3. The corresponding magnitude of the velocity estimation errors $|\tilde{v}|$ is plotted in Fig. 6.4.

The extended Kalman filter gives a biased estimate, the error decreases rapidly in the first 30 sec and then remains constant, for both $|\tilde{r}|$ and $|\tilde{v}|$. This problem arises when the noise inputs to the system are small or when the

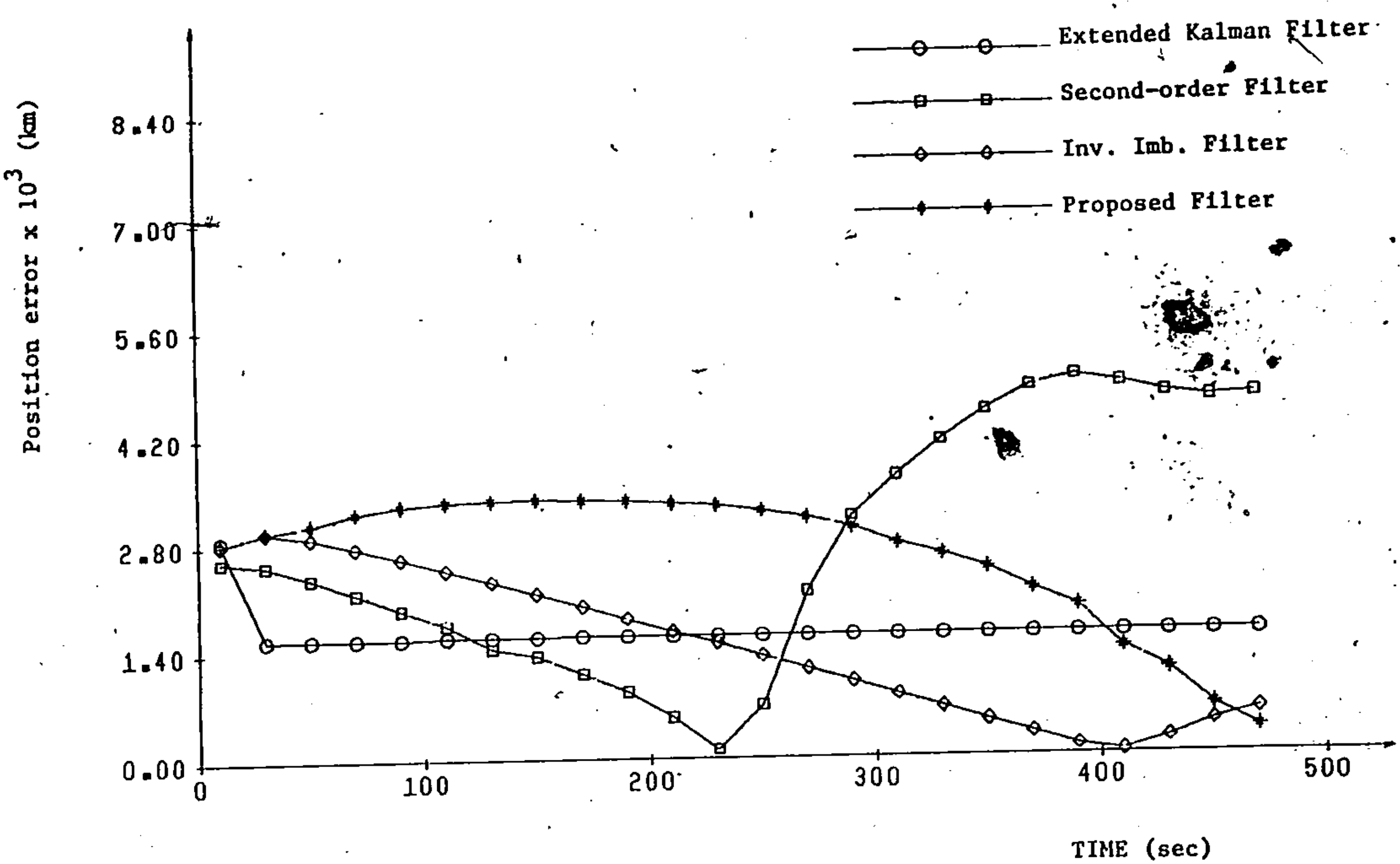


Figure 6.3

Position estimation error (var = 10^{-4})

(Circular orbit)

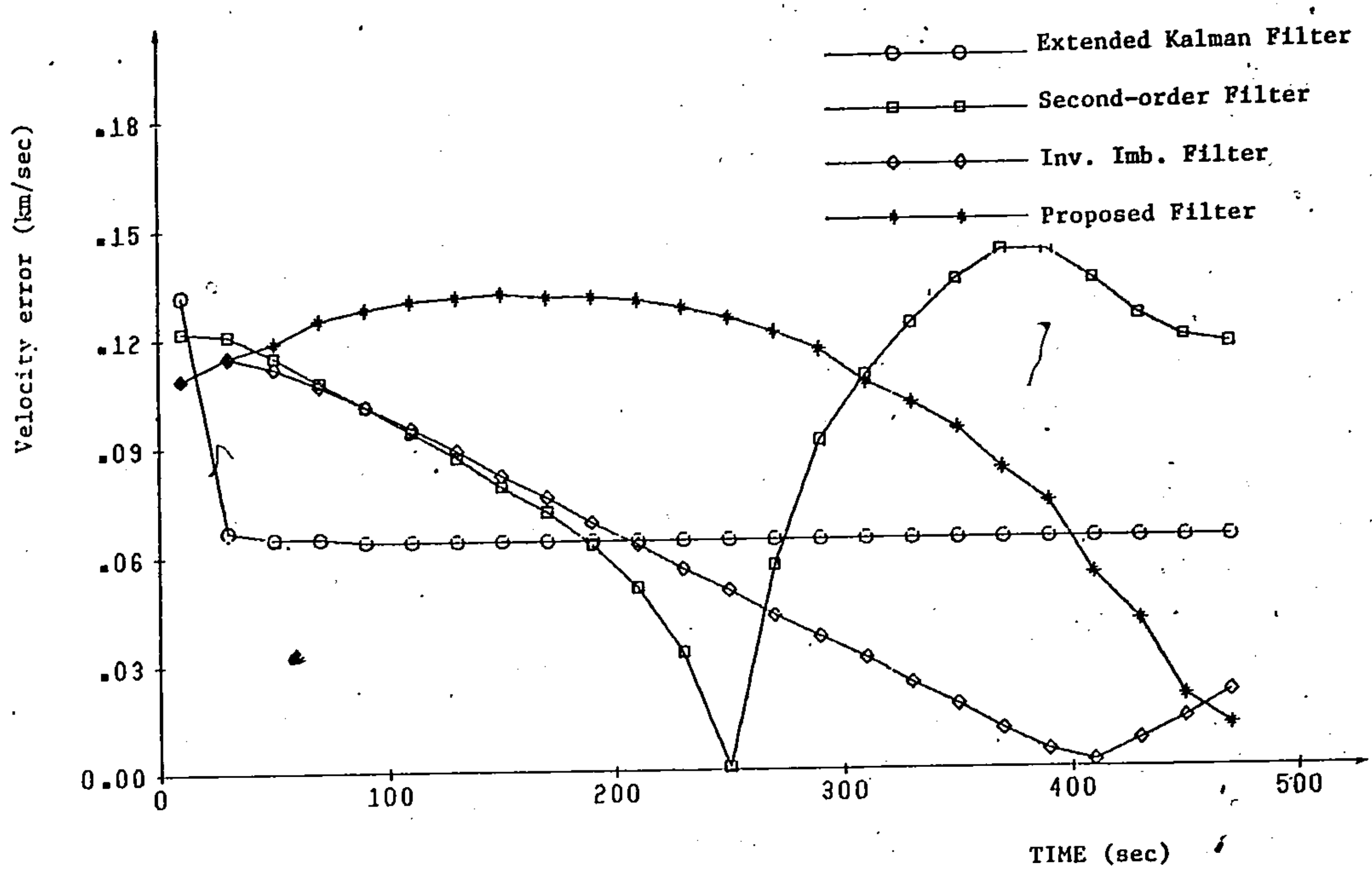


Figure 6.4

Velocity estimation error (var = 10^{-4})

(Circular orbit)

measurement noise is small. The extended Kalman filter algorithm was tested experimentally for different noise statistics to determine the best performance. However, the bias always existed and the best results, shown in the figures above, are obtained when the measurement noise standard deviations are

$$\sigma \text{ (noise in range)} = 1 \text{ km}$$

$$\sigma \text{ (noise in Azimuth and Elevation)} = 1^{\circ}$$

and the input noise is zero.

For the same values of Q, R and P_0 the second-order filter converges and the error is minimum after 230 sec than the filter starts to diverge. In this case it was experimentally proven that the filter performance is very sensitive to the value of R. The best results are obtained for

$$\sigma \text{ (noise in range)} = 100 \text{ km}$$

$$\sigma \text{ (noise in Azimuth and Elevation)} = 1^{\circ}$$

and zero input noise and are shown above.

The invariant imbedding algorithm was tested experimentally for different Q and R, and was found insensitive to noise statistics values. The plots given corresponds to zero input noise and R = [I]. The filter converges less rapidly than the second-order filter in the first 250 sec but with no bias or divergence afterwards.

The proposed algorithm utilizes the invariant imbedding approach for the first 70 sec then the stochastic approximation gain algorithm is applied for the rest of the time. We note the slow convergence compared to the other schemes, this is mainly due to the choice of the gain factor in the stochastic approximation algorithm. Different gain factors were tested and the best convergence was obtained for $\gamma = 2/k+10$, where k is the number of iterations. The efficiency of this algorithm is mainly due to the speed of computation as will be shown later.

The trace of the error covariance matrix is given in Table 6.1. The sum of the error variances decreases with time in all four cases. The best results are obtained when utilizing the invariant imbedding approach which is clear from the plots of the errors. The proposed algorithm gives a slower convergence rate, but the magnitude of the error variances are almost one tenth of that obtained using the first and second-order filters.

The computation time for each algorithm for the same number of observations is given in Table 6.2.

The results given above prove that the combined invariant imbedding and stochastic approximation algorithm is more efficient and practical compared to the other approaches. Convergence is achieved with much less computations.

Time (sec)	Tr [P] x 10^{-4}			
	Extended Kalman	Second Order Filter	Invariant imbedding	Proposed Algorithm
10	5.99	6.08	0.310	0.310
30	4.01	6.02	0.263	0.263
50	3.57	6.00	0.232	0.247
70	3.43	6.00	0.204	0.247
90	3.33	5.99	0.178	0.247
110	2.99	5.99	0.154	0.247
130	2.79	5.99	0.134	0.247
150	2.68	5.99	0.116	0.247
170	2.60	5.98	0.100	0.247
190	2.54	5.98	0.087	0.247
210	2.50	5.98	0.076	0.247
230	2.48	5.97	0.067	0.247
250	2.46	5.91	0.059	0.247
270	2.45	5.61	0.052	0.247
290	2.44	5.21	0.046	0.247
310	2.44	5.09	0.041	0.247
330	2.43	5.05	0.037	0.247
350	2.43	5.01	0.033	0.247
370	2.42	4.97	0.030	0.247
390	2.42	4.87	0.027	0.247
410	2.42	4.70	0.026	0.247
430	2.41	4.50	0.022	0.247
450	2.41	4.36	0.020	0.246
470	2.41	4.27	0.018	0.242

Table 6.1

Trace [P] (var = 10^{-4}) (Circular orbit)

Method	CPU Time (sec)
Extended Kalman	14.214
Second-order filter	176.025
Invariant imbedding	90.084
Proposed algorithm	11.244

Table 6.2

Comparison between execution times of various algorithms

For the same initial circular orbit a variance $\sigma_o^2 = 10^{-2}$ is assumed. The position estimation errors are plotted in Fig. 6.5 and the corresponding velocity estimation errors are shown in Fig. 6.6 and Fig. 6.7.

Table 6.3 gives the trace [P] and the computations time is similar to that in Table 6.2.

In this case the extended Kalman filter gives biased estimates with divergence in the first 70 sec. The trace [P] is rapidly decreasing while the estimation errors remain constant. Again, this is due to the choice of Q and R. The

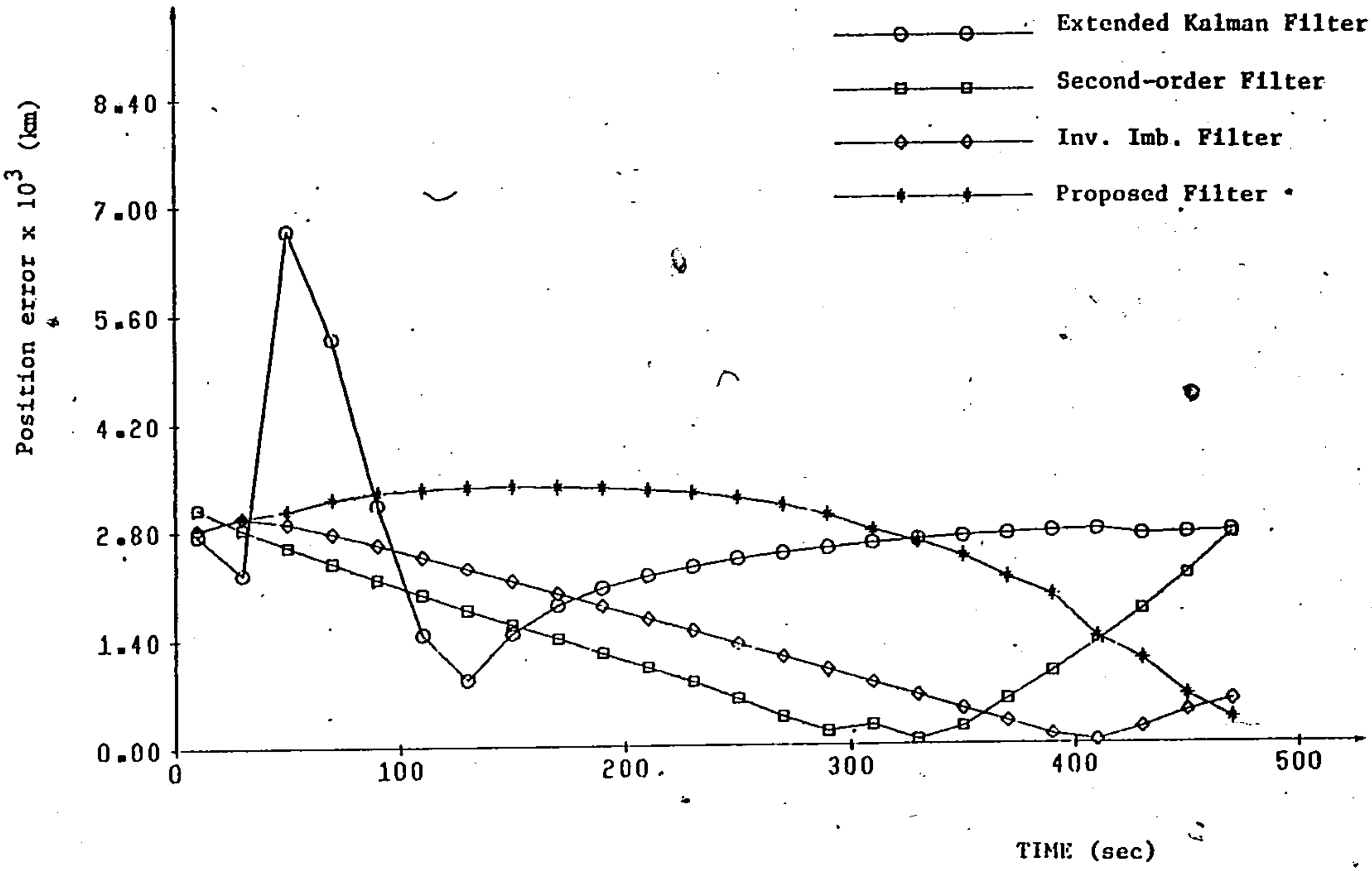


Figure 6.5

Position estimation error (var = 10^{-2})

(Circular orbit)

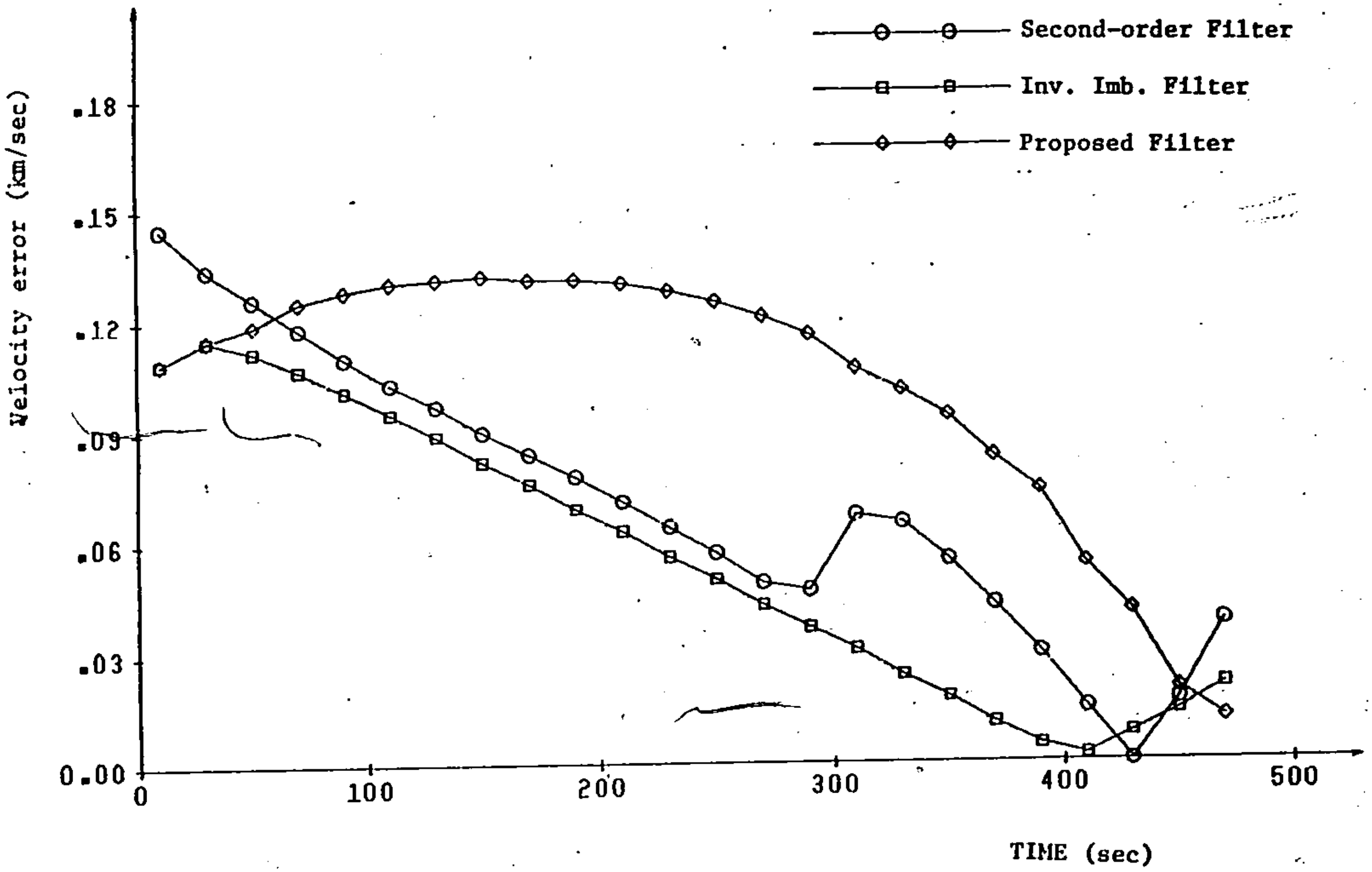


Figure 6.6

Velocity estimation error (var = 10^{-2})

(Circular orbit)

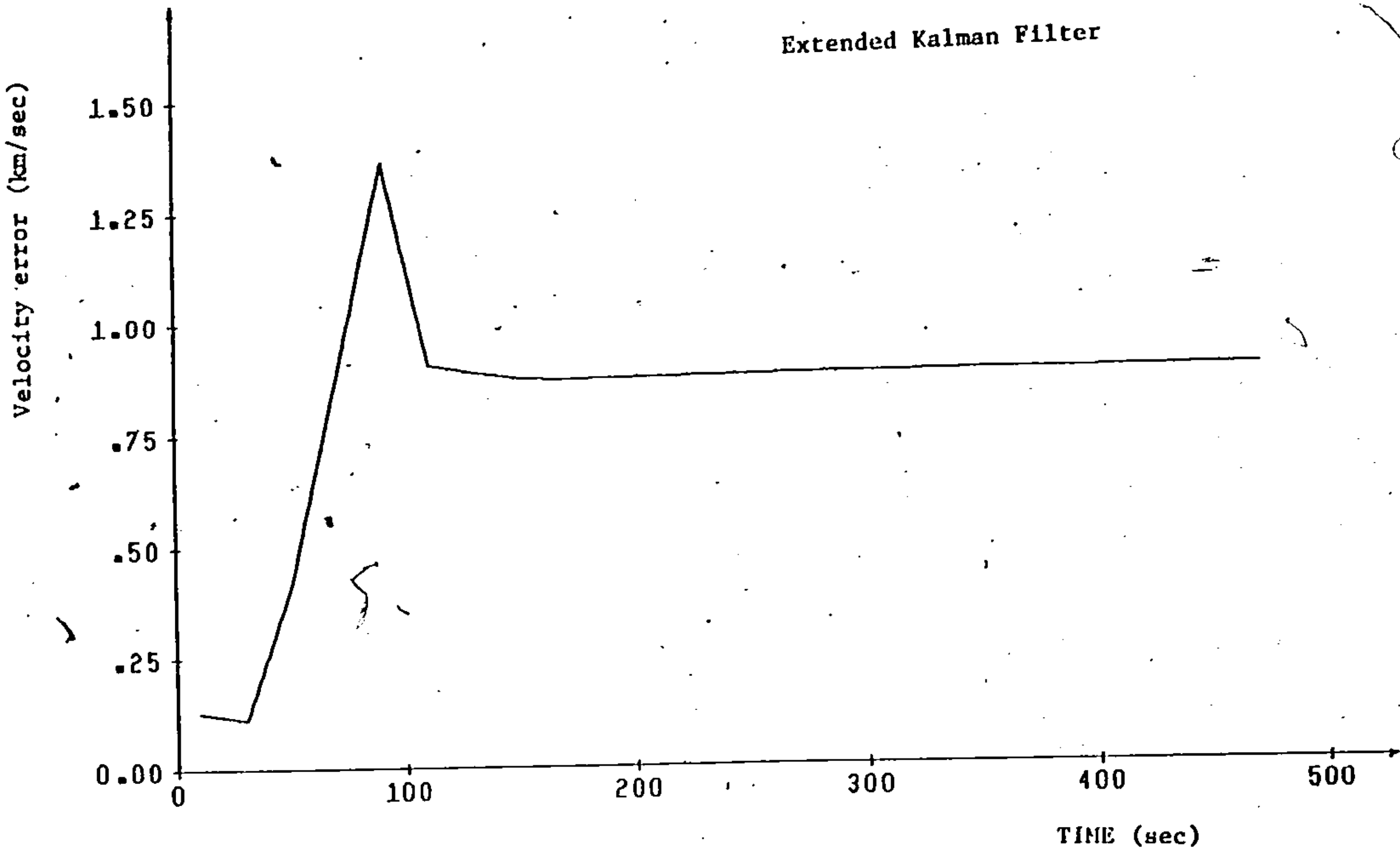


Figure 6.7

Velocity estimation error (var = 10^{-2})

(Circular orbit)

Time (sec)	Tr [P]			
	Extended Kalman	Second Order Filter	Invariant imbedding	Proposed Algorithm
10	5.92×10^{-2}	6.09×10^{-2}	4.45×10^{-5}	4.45×10^{-5}
30	3.84×10^{-2}	6.02×10^{-2}	3.54×10^{-5}	3.54×10^{-5}
50	1.98×10^{-2}	6.00×10^{-2}	2.99×10^{-5}	3.54×10^{-5}
70	0.41×10^{-2}	6.00×10^{-2}	2.53×10^{-5}	3.54×10^{-5}
90	7.37×10^{-6}	5.99×10^{-2}	2.14×10^{-5}	3.54×10^{-5}
110	4.031×10^{-7}	5.99×10^{-2}	1.81×10^{-5}	3.54×10^{-5}
130	2.447×10^{-7}	5.99×10^{-2}	1.53×10^{-5}	3.54×10^{-5}
150	2.429×10^{-7}	5.99×10^{-2}	1.30×10^{-5}	3.54×10^{-5}
170	2.424×10^{-7}	5.99×10^{-2}	1.10×10^{-5}	3.54×10^{-5}
190	2.421×10^{-7}	5.98×10^{-2}	0.95×10^{-5}	3.54×10^{-5}
210	2.420×10^{-7}	5.98×10^{-2}	0.82×10^{-5}	3.54×10^{-5}
230	2.418×10^{-7}	5.98×10^{-2}	0.71×10^{-5}	3.54×10^{-5}
250	2.417×10^{-7}	5.98×10^{-2}	0.62×10^{-5}	3.54×10^{-5}
270	2.416×10^{-7}	5.97×10^{-2}	0.54×10^{-5}	3.54×10^{-5}
290	2.415×10^{-7}	5.87×10^{-2}	0.48×10^{-5}	3.54×10^{-5}
310	2.415×10^{-7}	5.42×10^{-2}	0.42×10^{-5}	3.54×10^{-5}
330	2.414×10^{-7}	5.13×10^{-2}	0.38×10^{-5}	3.54×10^{-5}
350	2.414×10^{-7}	5.05×10^{-2}	0.34×10^{-5}	3.54×10^{-5}
370	2.413×10^{-7}	5.01×10^{-2}	0.30×10^{-5}	3.54×10^{-5}
390	2.413×10^{-7}	4.99×10^{-2}	0.27×10^{-5}	3.54×10^{-5}
410	2.412×10^{-7}	4.97×10^{-2}	0.25×10^{-5}	3.54×10^{-5}
430	2.412×10^{-7}	4.96×10^{-2}	0.22×10^{-5}	3.54×10^{-5}
450	2.412×10^{-7}	4.95×10^{-2}	0.20×10^{-5}	3.54×10^{-5}
470	2.412×10^{-7}	4.94×10^{-2}	0.19×10^{-5}	3.54×10^{-5}

Table 6.3

Trace [P] (var = 10^{-2}) (Circular orbit) given plots are obtained for zero input noise and R = [I].

The second-order filter performance is better than the extended Kalman filter. The algorithm converges more rapidly in the first 300 sec, however, the errors increase again toward the end of the period under investigation.

The invariant imbedding algorithm has a similar performance to the first case, which proves its insensitivity to P_o.

In this case the invariant imbedding algorithm gives the best convergence rate. Both the extended Kalman filter and the second-order filter diverge. The proposed algorithm is again the best in terms of computation time but with slower convergence compared to the invariant imbedding algorithm.

b. <u>Initial Elliptic Orbit</u>

An initial elliptic orbit is considered and the same algorithms are applied. The initial states are given as follows

$$C = 3.074609 \text{ km/sec}$$

$$R_1 = 1.332 \times 10^{-4} \text{ km/sec}$$

$$R_2 = 2.39 \times 10^{-4} \text{ km/sec}$$

$$\ell_1 = 1.747 \times 10^{-4}$$

$$\ell_2 = 1.297 \times 10^{-3}$$

$$\ell_3 = 0.1807$$

$$\ell_4 = 0.9835$$

These values are obtained from results of an actual orbit determination performed at the Communications Research Centre, Ottawa.

Results of simulation are shown in Fig. 6.8 and Fig. 6.9, corresponding to initial error variance $\sigma_0^2 = 10^{-2}$.

In this case the extended Kalman filter diverges and the magnitude of the estimation errors is higher than that predicted by the error covariance matrix, as the trace [P] is decreasing (Table 6.4).

The performance of the other three filters are similar to the circular orbit case, which proves that they are not sensitive to initial values of the states x_0.

The computation time is almost the same as that given in Table 6.2.

The invariant imbedding algorithm is rapidly converging and gives the best performance compared to the first and second-order filters.

The combined algorithm is coverging less rapidly than the invariant imbedding but takes almost 1/10 of the computer time required for the first algorithm.

Based on these results the best convergence rate is obtained using invariant imbedding, but for real time applications the proposed algorithm should be used as it requires the least computer time although the convergence rate is slower than the invariant imbedding scheme.

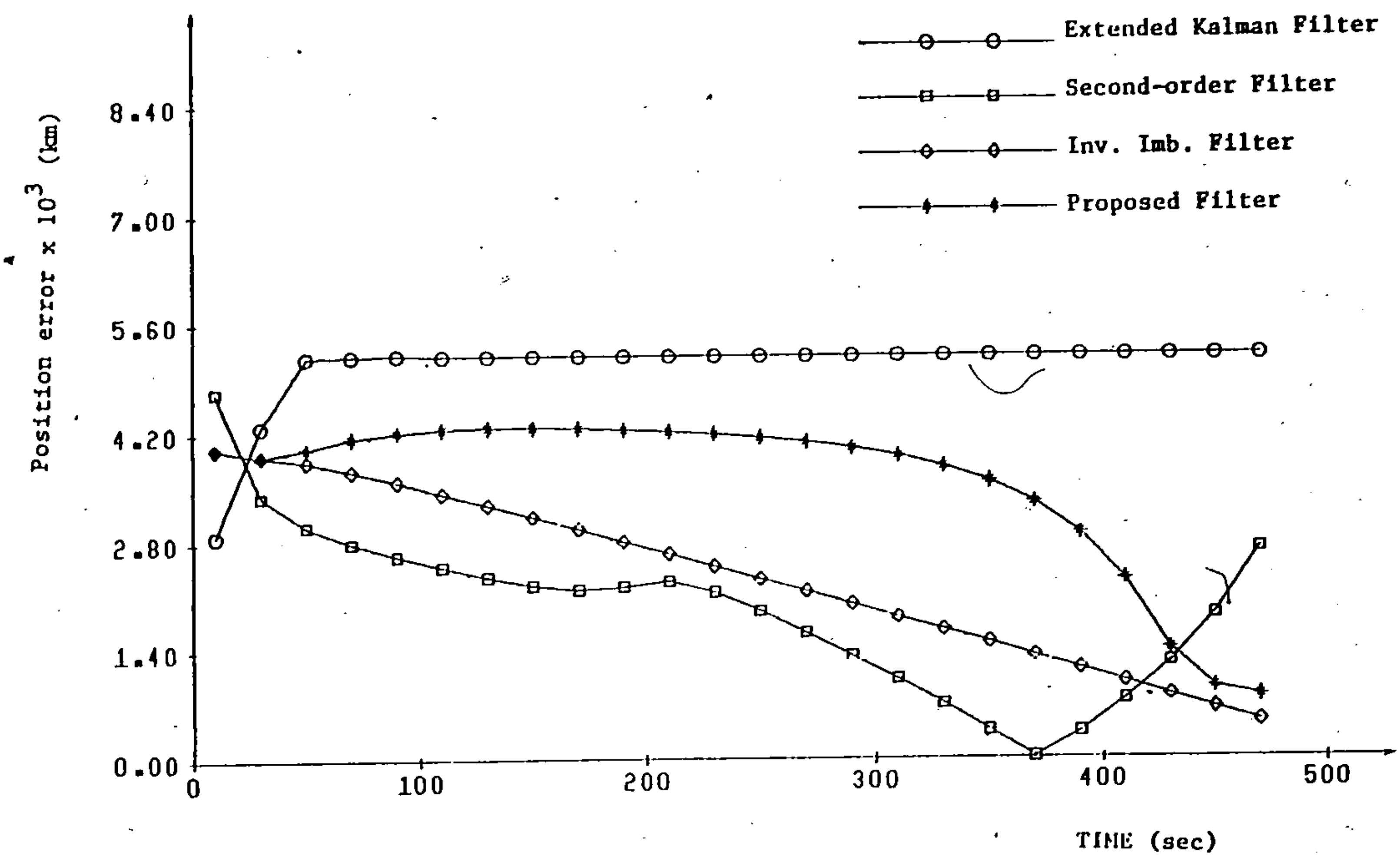


Figure 6.8

Position estimation error (var = 10^{-2})

(Elliptic orbit)

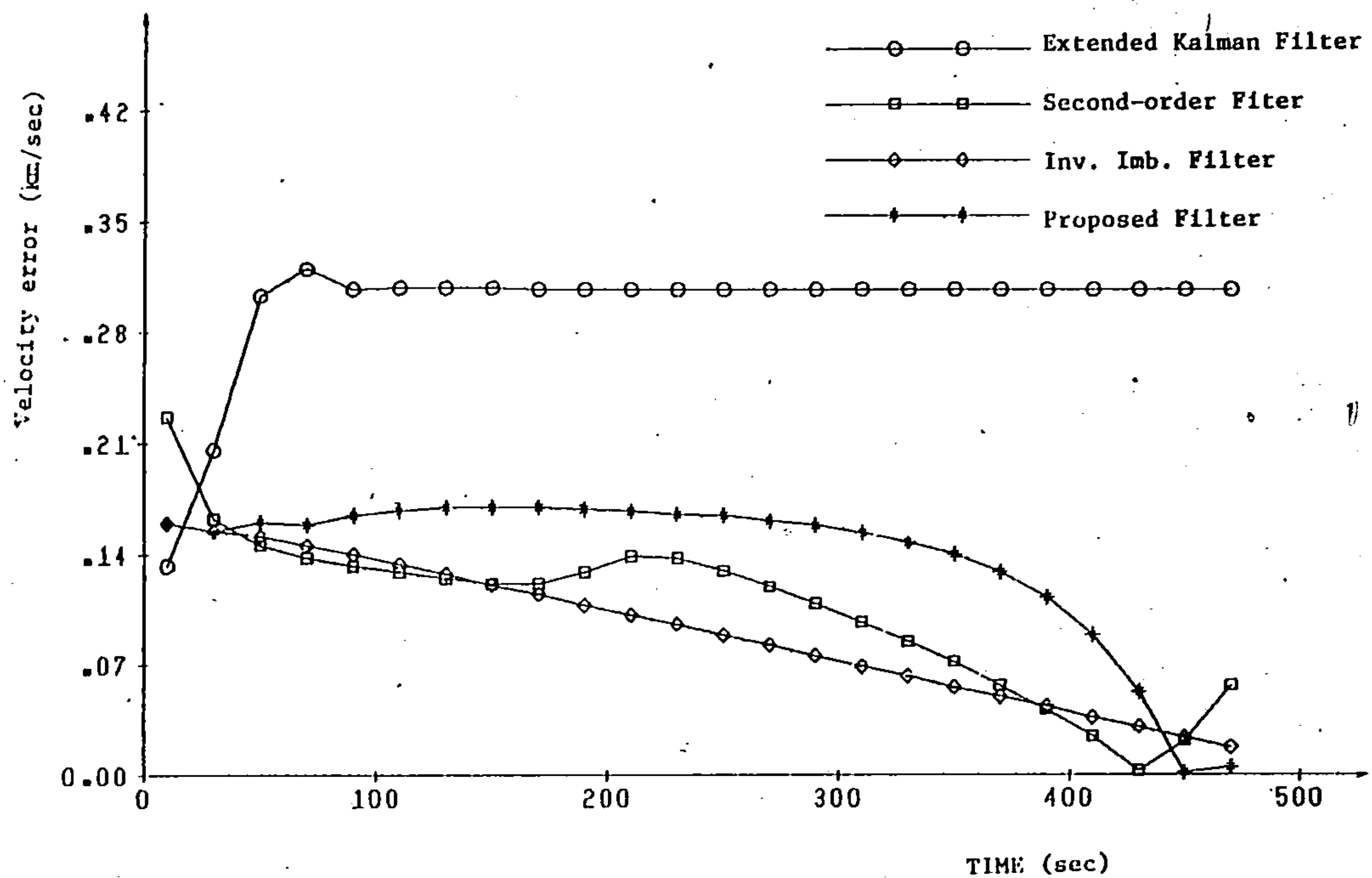


Figure 6.9

Velocity estimation error (var = 10^{-2})

(Elliptic orbit)

Time (sec)	Tr [P]			
	Extended Kalman	Second Order Filter	Invariant imbedding	Proposed Algorithm
10	5.92×10^{-2}	6.12×10^{-2}	1.03×10^{-7}	1.03×10^{-7}
30	3.90×10^{-2}	5.96×10^{-2}	1.05×10^{-7}	1.05×10^{-7}
50	2.01×10^{-2}	5.91×10^{-2}	1.03×10^{-7}	1.073 x "
70	0.166 x "	5.88 x "	0.92 x "	1.072 x "
90	0.039 x "	5.87 x "	0.80 x "	1.072 x "
110	0.039 x "	5.86 x "	0.69 x "	1.071 x "
130	0.039 x "	5.86 x "	0.59 x "	1.071 x "
150	0.039 x "	5.85 x "	0.51 x "	1.070 x "
170	0.039 x "	5.82 x "	0.44 x "	1.069 x "
190	0.039 x "	5.70 x "	0.38 x "	1.068 x "
210	0.039 x "	5.38 x "	0.33 x "	1.067 x "
230	0.039 x "	5.12 x "	0.29 x "	1.066 x "
250	0.039 x "	5.01 x "	0.25 x "	1.065 x "
270	0.039 x "	4.96 x "	0.22 x "	1.064 x "
290	0.039 x "	4.94 x "	0.20 x "	1.063 x "
310	0.039 x "	4.92 x "	0.18 x "	1.062 x "
330	0.039 x "	4.91 x "	0.16 x "	1.060 x "
350	0.039 x "	4.90 x "	0.16 x "	1.058 x "
370	0.039 x "	4.89 x "	0.13 x "	1.054 x "
390	0.039 x "	4.88 x "	0.12 x "	1.047 x "
410	0.039 x "	4.87 x "	0.11 x "	1.035 x "
430	0.039 x "	4.86 x "	0.10 x "	1.004 x "
450	0.039 x "	4.84 x "	0.09 x "	0.902 x "
470	0.039 x "	4.79 x "	0.08 x "	0.804 x "

Table 6.4

Trace [P] (var = 10^{-2}) (Elliptic orbit)

6.5 Concluding Remarks

Divergence problems have been particularly acute in orbit determination. This is because measurement devices are very accurate, many data are available, and the filter is required to operate over many revolutions of the satellite orbit.

The main objective of this research is to implement an algorithm for state estimation that converges independently of the a priori assumptions.

On the other hand it is, of course, necessary that the time required to complete the computation cycle be less, on the average, than the time interval between successive observations.

The above simulation study shows that the extended Kalman filter is very sensitive to the noise statistics and the initial error covariance matrix P_o.

The second-order filter improves the performance by including second-order derivatives in the filter computations. The second-order terms appear as Ph_{xx} in the case of measurement nonlinearity, and as Pf_{xx} in the case of system nonlinearity (where the suffix xx denotes second derivatives with respect to x). If these nonlinear terms are large the nonlinearity effects are significant. Now, the second-order filter is useful and effective when nonlinearities are large because h_{xx} and f_{xx} are large.

However, the performance of the filter is not as excepted when P is large. Significant errors can be made in prediction over long time intervals which have a biasing effect on the estimate. It was demonstrated experimentally that measurement nonlinearities are significant when the noise variance R is small while the error variance P is relatively large. Because, the choice of R and P_o is based on ad hoc assumptions the second-order estimation scheme has its limitations.

The invariant imbedding algorithm performance is as expected theoretically, convergence is obtained independent of noise statistics and a priori assumptions.

The proposed algorithm has the above advantages of the invariant imbedding approach and requires much less computer time. The computation time for each sampling period of 10 sec was found to be ≈ 1/4 sec which makes it applicable for real-time operations.

Some of the results outlined in this chapter are published in reference [62].

CHAPTER 7

CONCLUSION

The motivation for this work centres on the necessity to determine, or estimate, present and future values of system state variables from noisy observations of the output. This problem is of central importance in aerospace applications. Although filtering theory has been applied in several fields, it is probably fair to say that most of the applications have been made by the aerospace community, which contributed significantly to applying the theory to practice.

The problem of estimating the orbital states of a communications satellite from ground-based measurements, which are contaminated with noise, is complicated by the nonlinearity of the dynamic equations as well as lack of prior knowledge of the noise statistics.

The extended Kalman filter has been successfully applied in several aerospace applications, especially orbit determination problems. The filter algorithm does not involve second partial derivatives, and is therefore recommended on the basis of simplicity and smallest computational requirements. Since the system is usually an approximation to the physical situation, the model parameters and noise statistics are seldom exact. It is clear that an inexact filter model will degrade the filter performance and cause divergence.

If nonlinearities are significant, however, filter performance can be improved by inclusion of second-order effects. The price of this improvement is that the second-order filter is more complicated. On the basis of the simulation results given in this thesis, one can question the practical usefulness of higher-order approximations, especially in problems of high dimension (for example 7). Furthermore, it is questionable whether higher-order approximations would improve performance in cases where the extended Kalman filter diverges.

The problems that might occur when using the extended Kalman filter or the second-order filter are mainly due to the lack of knowledge of noise statistics. The superiority of the invariant imbedding approach lies in the fact that it provides a sequential nonlinear estimator which does not depend on a priori noise statistics. Again, in this case, performance is improved at the price of more computations when compared to the extended Kalman filter.

In the final analysis, the performance of the previously mentioned filters must be tested by simulation. The so called "best performance" means fast convergence with simple computations. So, according to the simulations results obtained no filter can be recommended on the basis of superior performance.

In view of the above observations, confirmed by the results of simulation, a new algorithm is presented which combines invariant imbedding with stochastic approximation. It is proposed to use invariant imbedding to determine the initial gain matrix and then apply stochastic approximation to track changes in the gain matrix which will bring further improvement. This algorithm is found to be computationally more efficient relative to the three other schemes. It requires less computing time than the extended Kalman filter and is independent of initial assumptions and noise statistics. The only limitation to the use of this algorithm is the slow convergence.

One of the main advantages of the proposed algorithm is its feasibility for real-time operations. So far, the program has been tested with actual tracking data of the communications satellite to evaluate the time it takes per iteration relative to the time interval between successive observations. According to the results obtained the algorithm can be used for real-time tracking.

Suggestions for Future Research

1. The covergence of the algorithm combining invariant imbedding and stochastic approximation has not been theoretically justified. When these two estimators are treated separately, the convergence of each is theoretically proved. However, the convergence of the overall estimator is not obvious because of their interdependence. This aspect has to be further investigated.
2. At present there is no rational for the choice of the gain factor $\gamma(k)$ in the stochastic approximation. Instead of ad hoc assumptions a computational scheme should be devised for on-line prediction of this gain.
3. An in-depth study of the second-order approximation algorithms and their advantages over first-order filters is a good area for investigation.
4. Adaptive schemes for noise statistics computations can be combined with any general filtering algorithm to avoid the sensitivity to assumed a priori statistics. However, a time consuming adaptive scheme is not desirable for real-time operations. This computational aspect could be further investigated.

REFERENCES

[1] Wiener, N., Extrapolation, interpolation and smoothing of stationary time series, with engineering applications. John Wiley: New York, 1949.

[2] Kalman, R.E., "A new approach to linear filtering and prediction problems", J. Basic Engineering, vol. 82, pp. 34-45, 1960.

[3] Kalman, R.E. and R.S. Bucy, "New results in linear filtering and prediction theory", J. Basic Engineering, vol. 83, pp. 95-107, 1961.

[4] Cox, H., "On the estimation of state variables and parameters for noisy dynamic systems", IEEE Trans. Autom. Control, vol. AC-9, pp. 5-12, 1964.

[5] Mowery, V.O., "Least squares recursive differential correction estimation in nonlinear problems", IEEE Trans. Autom. Control, vol. AC-9, pp. 399-407, 1965.

[6] Ho, Y.C. and R.C.K. Lee, "A Bayesian approach to problems in stochastic estimation and control", IEEE Trans. Autom. Control, vol. AC-9, pp. 333-339, 1964.

[7] Kushner, H.J., "Dynamical equations for optimum nonlinear filtering", J. Differential equations, 3, pp. 179-190, 1967.

[8] Bucy, R.S., "Nonlinear filtering theory", IEEE Trans. Autom. Control, vol. AC-10, pp. 198-199, 1965.

[9] Bass, R.W., V.D. Norum, and L. Schwartz, "Optimal multi-channel nonlinear filtering", J. Math. Analysis and Applications, vol. 16, pp. 152-164, 1966.

[10] Denham, W.F., and S. Pines, "Sequential estimation when measurement function nonlinearity is comparable to measurement error", AIAA J., vol. 4, no. 6, pp. 1071-1076, 1966.

[11] Athans, M., R.P. Wishner, and A. Bertolini, "Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements", IEEE Trans. Autom. Control, vol. AC-13, pp. 504-514, 1968.

[12] Wishner, R.P., J.A. Tabaczynski, and M. Athans, "A comparison of three nonlinear filters", Automatica, vol. 5, no. 4, pp. 487-496, 1969.

[13] Detchmendy, D.M., and R. Sridhar, "Sequential estimation of states and parameters in noisy nonlinear dynamical systems", Trans. ASME, J. Basic Eng., vol. 88D, pp. 362-366, 1966.

[14] Kagiwada, H.H., R.E. Kalaba, A. Schumitzky, and R. Sridhar, "Invariant imbedding and sequential interpolating filters for nonlinear processes", Trans. ASME, J. Basic Eng., vol. 91D, no. 2, pp. 195-200, 1969.

[15] Kailath, T., "An innovations approach to least squares estimation - Part I: Linear filtering in additive white noise", IEEE Trans. Autom. Control, vol. AC-13, pp. 646-655, 1968.

[16] Kailath, T., and P. Frost, "An innovations approach to least squares estimation - Part II: Linear smoothing in additive white noise", IEEE Trans. Autom. Control, vol. AC-13, pp. 655-660, 1968.

[17] Kolmogorov, A.N., "Interpolation and extrapolation of stationary random sequences", Memorandum RM-3090, RP, Rand. Corp., Santa Monica, Calif., 1962.

[18] Altman, S.P., "A unified state model of orbital trajectory and altitude dynamics", AAS/AIAA Astrodynamics Specialists Conference, pp. 425-446, 1971.

[19] Kalman, R.E., "New methods of Wiener filtering theory", Proc. First Symp. Eng. Appl. Random Functions Theory Probability, John Wiley, New York, 1963.

[20] Sage, A.P., and G.W. Husa, "Adaptive filtering with unknown prior statistics", Proc. Joint Autom. Control Conf., pp. 760-769, 1969.

[21] Mehra, R.K., "On the identification of variances and adaptive Kalman filtering", IEEE Trans. Autom. Control, vol. AC-15, pp. 175-184, 1970.

[22] Astrom, K.J., Introduction to stochastic control theory. Academic Press: New York, 1970.

[23] Bode, H.W. and C.E. Shannon, "A simplified derivation of linear least-squares smoothing and prediction theory", Proc. IRE, vol. 38, pp. 417-425, 1950.

[24] Meditch, J.S., Stochastic optimal linear estimation and control. McGraw-Hill: New York, 1969.

[25] Deutsh, R., Estimation theory. Prentice-Hall: N.J., 1965.

[26] Sage, A.P., and J.L. Melsa, Estimation theory with applications to communications and control. McGraw-Hill: New York, 1971.

[27] Swerling, P., "Note on a new computational data smoothing procedure suggested by minimum mean square error estimation", IEEE Trans. Inform. Theory, vol. 12, pp. 9-12, 1966.

[28] Ho, Y.C., "On the stochastic approximation method and optimal filter theory", J. Math. Anal. Appl., vol. 6, pp. 152-154, 1962.

[29] Jazwinski, A.H., Stochastic processes and filtering theory. Academic Press: New York, 1970.

[30] Chen, C.T., Introduction to linear system theory. Holt, Rinehart and Winston Inc., 1970.

[31] Smith, G.L., "Sequential estimation of observation error variances in a trajectory estimation problem", AIAA J., vol. 5, pp. 1964-1970, 1967.

[32] Jazwinski, A.H.,, "Adaptive filtering", Automatica, vol. 5, pp. 475-485, 1969.

[33] Carew, B. and P.R. Belanger, "Identification of optimum filter steady state gain for system with unknown noise covariance", IEEE Trans. Autom. Control, vol. AC-18, pp. 581-587, 1973.

[34] Scharf, L.L. and D.L. Alspach, "On stochastic approximation and an adaptive Kalman filter", Proc. IEEE Decision and Control Conference, pp. 253-257, 1972.

[35] Sinha, N.K., "Adaptive Kalman filtering using stochastic approximation", Electronics Letters, vol. 9, pp. 177-178, 1973.

[36] Sinha, N.K., and S.K. Mukherjee, "An adaptive scheme for Kalman filtering", Proc. IFAC Symp. on Stochastic Control, pp. 363-368, 1974.

[37] Doob, J.L., Stochastic processes. John Wiley: New York, 1953.

[38] Gevers, M.R., and T. Kailath, "An innovations approach to least-squares estimation - Part VI: Discrete-time innovations representations and recursive estimation", IEEE Trans. Autom. Control, vol. AC-18, pp. 588-600, Dec. 1973.

[39] Frost, P.A., "The innovations process and its applications to nonlinear estimation and detection of signals in additive white noise", Proc. UMR-Mervin J. Kelly Communications Conf., Oct., 1970.

[40] Frost, P.A., and T. Kailath, "An innovations approach to least squares estimation - Part III: Nonlinear estimation in W.G.N.", IEEE Trans. Autom. Control, vol. AC-16, 1971.

[41] Azim, S.A., "Modelling and state estimation for synchronous-orbit satellites", M.Eng. Thesis, McMaster University, 1975.

[42] Schee, F.H., C.J. Standish, and N.F. Toda, "Divergence in the Kalman filter", AIAA J., vol. 5, pp. 1114-1120, 1967.

[43] Potter, J.E., and D.C. Fraser, "A formula for updating the determinant of the covariance matrix", AIAA J., vol. 5, pp. 1352-1354, 1967.

[44] Andrews, A., "A square root formulation of the Kalman covariance equations", AIAA J., vol. 6, pp. 1165-1166, 1968.

[45] Tapley, B.D., and C.Y. Choe, "Nonlinear estimation theory applied to the interplanetary orbit determination problem", pp. 214-217.

[46] Clark, J.M.C., "The representation of functionals of Brownian motion by stochastic integrals", Ann. Math. Statist., vol. 4, pp. 1282-1295, 1970.

[47] Lee, E.S., Quasilinearization and invariant imbedding. Academic Press Inc.: New York, 1968.

[48] Mendel, J.M., and K.S. Fu, Adaptive learning and pattern recognition systems. Academic Press Inc.: New York, 1970.

[49] Sakrison, D.J., "Stochastic approximation: Recursive method for solving regression problems", Advances in Communications Systems, vol. 2, pp. 51-106, 1966.

[50] Sage, A.P., and J.L. Melsa, System identification. Academic Press Inc.: New York, 1970.

[51] Robbins, H., and S. Monro, "A stochastic approximation method", Ann. Math. Statistics, vol. 22, pp. 400-407, 1951.

[52] Kiefer, J., and J. Wolfowitz, "Statistical estimation of the maximum of a regression function", Ann. Math. Statistics, vol. 23, pp. 426-466, 1952.

[53] Dvoretzky, A., "On stochastic approximation",, Proc. 3rd Berkeley Symposium Math. Stat. and Probability, vol. 1, pp. 39-55, 1956.

[54] Saridis, G.N., "Stochastic approximation methods for identification and control - A survey", IEEE Trans. Autom. Control, vol. AC-19, pp. 798-809, 1974.

[55] Hampton, R.L.T., and D.G. Schultz, "Stochastic algorithm for self-adaptive filtering and prediction", Proc. Int. IEEE Conf. in Systems, Computers and Networks, pp. 463-467, 1971.

[56] Tom, A.F., "Adaptive state estimation for systems with white and coloured noise", M.Eng. Thesis, McMaster University, 1975.

[57] Sinha, N.K., and S.A. Azim, "Adaptive estimation of the states of a synchronous-orbit satellite", International J. of Systems Science, vol. 9, pp. 121-127, 1978.

[58] Sage, A.P., Optimum systems control. Prentice-Hall: N.J., 1968.

[59] Hamilton, W.R., Elements of Quaternions. Chelsea Publishing Company: New York, vol I, 1969.

[60] Escobal, P.R., Methods of orbit determination. Wiley: New York, 1965.

[61] Sinha, N.K., and A. Tom, "Adaptive state estimation for systems with unknown noise covariances", Int. J. System Science, vol. 8, No. 4, pp. 377-384, 1977.

[62] Azim, S.A., and N.K. Sinha, "Estimation of the orbital states of a satellite from ground-based measurements", Proc. Sixteenth Annual Allerton Conference on Communication, Control and Computing, pp. 202-212, 1978.

APPENDIX A

THE INNOVATIONS THEOREM

The process $\xi(\cdot)$ defined by

$$\xi_t = y_t - \hat{z}_t = \tilde{z}_t + v_t, \qquad 0 \leq t \leq T, \tag{A.1}$$

where $\hat{z}_t = H_t\hat{x}_t$, is the innovations process of y_t.

The innovations process $\xi(\cdot)$ is a white noise process with the same covariance as the observation error $v(\cdot)$, i.e.,

$$E\{\xi_t\xi_s'\} = E\{v_tv_s'\}, \qquad 0 \leq t < s < T. \tag{A.2}$$

Also $y(\cdot)$, $\xi(\cdot)$ can each be obtained from each other through a causal and causally invertible linear operation.

<u>Proof</u>

The proof is divided into two parts. First we show that $\xi(\cdot)$ is white and has the same covariance as $v(\cdot)$ and secondly the existence of a causal and causally invertible linear operator between $\xi(\cdot)$ and $v(\cdot)$ is proved.

I. Let $t > s$. Then

$$E\{\xi_t\xi_s'\} = E\{(\tilde{z}_t+v_t)(\tilde{z}_s+v_s)'\}$$

$$= E\{\tilde{z}_t\tilde{z}_s'\} + E\{\tilde{z}_t v_s'\} + E\{v_t\tilde{z}_s'\} + E\{v_t v_s'\}$$

$$= E\{\tilde{z}_t(z_s-\hat{z}_s)'\} + E\{\tilde{z}_t v_s'\} + E\{v_t v_s'\}$$

$$= E\{\tilde{z}_t z_s'\} - E\{\tilde{z}_t\hat{z}_s'\} + E\{\tilde{z}_t v_s'\} + E\{v_t v_s'\}$$

$$= E\{\tilde{z}_t(z_s+v_s)'\} + E\{v_t v_s'\}$$

$$= 0$$

Therefore,

$$E\{\xi_t\xi_s'\} = E\{v_t v_s'\} = R_t\delta(t-s)$$

It can similarly be proved for $t < s$. For $t = s$, however, since $E\{v_t v_t'\}$ is infinite, $E\{\nu_t\nu_t'\}$ must also be infinite.

Hence, $\xi(\cdot)$ is a white noise process.

II. Let (t,s) denote the optimum causal filter that operates on $\{y_s,\ s \le t\}$ to give $\hat{z}(\cdot)$, i.e., let

$$\hat{z}_t = \int_0^t h(t,s)\ y_s\ ds = Hy \tag{A.3}$$

where H denotes the integral operator with kernel $h(\cdot,\cdot)$ (volterra kernel).

Now,

$$\xi_t = y_t - \hat{z}_t = [I - H]\, y$$

where I is the identity operator with kernel $\delta(t-s)$.

The problem is to show that [I-H] is an invertible operator. Since (I-H) is causal, therefore its inverse (if its exists) is also causal.

Now, $(I-H)^{-1} = I + H + H^2 + \ldots$ (Newman Series) exists whenever H has a square integrable kernel function.

The proof is thus complete.

APPENDIX B

DEFINITION OF THE EULER PARAMETERS

The rotation of a triad set can be generated by a scalar rotation about a directed line from the origin of the inertial space. This rotation is defined by the unitary quaternion

$$q = \ell_4 + (g_1\ell_1 + g_2\ell_2 + g_3\ell_3) \qquad \text{(B.1)}$$

where the set of four Euler parameters

$$\begin{bmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{bmatrix} = \begin{bmatrix} \cos\alpha \sin u/2 \\ \cos\beta \sin u/2 \\ \cos\gamma \sin u/2 \\ \cos u/2 \end{bmatrix} \qquad \text{(B.2)}$$

consists of real scalars such that

$$\ell_1^2 + \ell_2^2 + \ell_3^2 + \ell_4^2 = 1. \qquad \text{(B.3)}$$

The spherical angles (α, β, γ) in equation (B.2) are shown in Fig. B.1. As an alternative, the Euler parameters may be defined in terms of the Euler angles $(\Omega, i, u1)$ shown in Fig. B.2, in accordance with

$$\begin{bmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{bmatrix} = \begin{bmatrix} \sin i/2 \cos(\Omega-ul)/2 \\ \sin i/2 \sin(\Omega-ul)/2 \\ \cos i/2 \sin(\Omega+ul)/2 \\ \cos i/2 \cos(\Omega+ul)/2 \end{bmatrix} \tag{B.4}$$

The variation of the Euler parameters with angular velocity of rotation (W) of the triad set is defined as follows:

$$\frac{d}{dt}\begin{bmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \\ \ell_4 \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 0 & W_z & -W_y & W_x & \ell_1 \\ -W_z & 0 & W_x & W_y & \ell_2 \\ W_y & -W_x & 0 & W_z & \ell_3 \\ -W_x & -W_y & -W_z & 0 & \ell_4 \end{bmatrix} \tag{B.5}$$

In matrix notation, a coordiante triad $[X]_{Body}$ is rotated from an inertial reference triad $[X]_{Inertial}$ by a transformation [E] such that

$$[X]_{Body} = [E]\ [X]_{Inertial} \tag{B.6}$$

where

$$[E] = \begin{bmatrix} \xi_{11} & \xi_{12} & \xi_{13} \\ \xi_{21} & \xi_{22} & \xi_{23} \\ \xi_{31} & \xi_{32} & \xi_{33} \end{bmatrix} \tag{B.7}$$

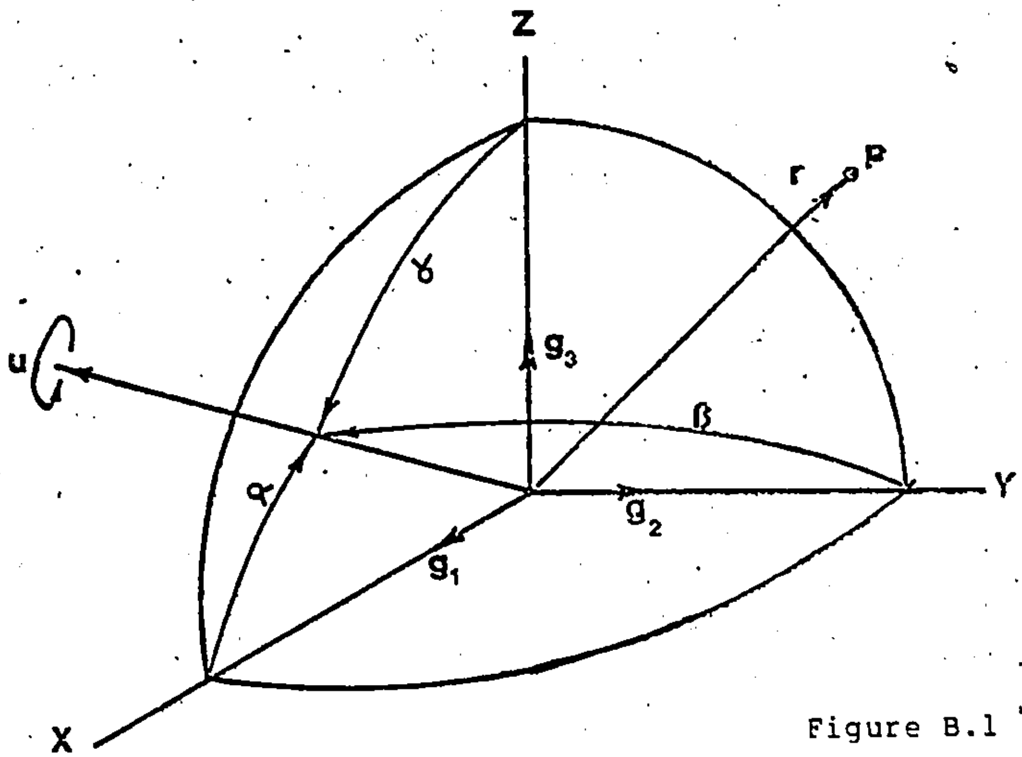


Figure B.1

Geometry for the Euler parameters

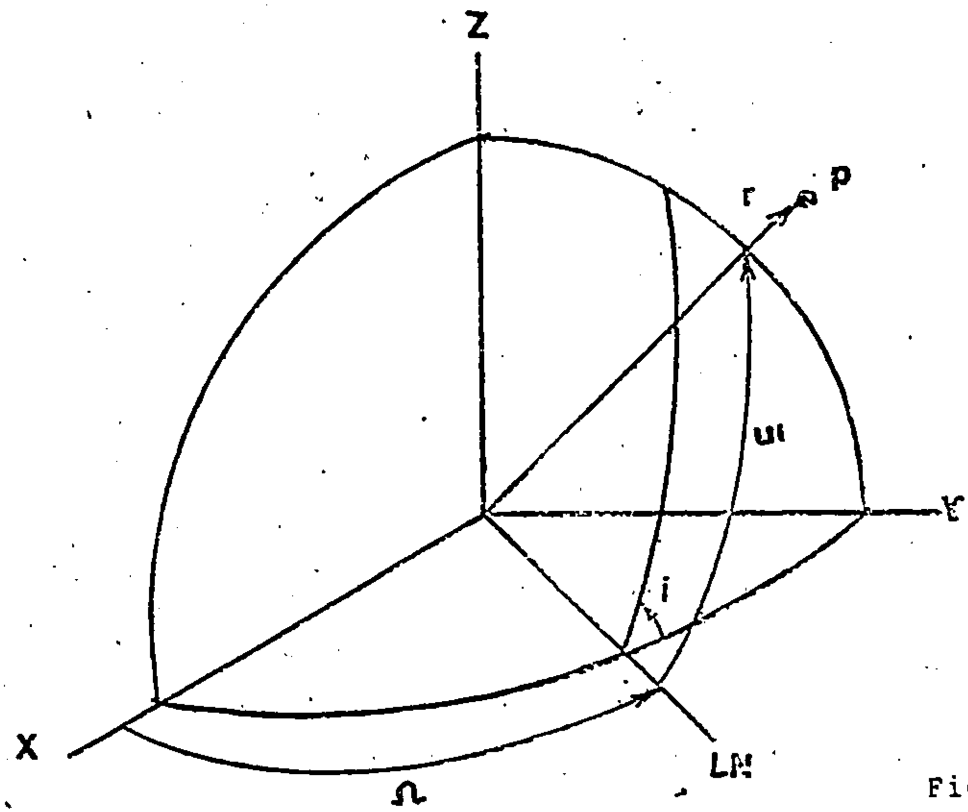


Figure B.2

Euler angles

and

$$\xi_{11} = 1 - 2(\ell_2^2 + \ell_3^2)$$

$$\xi_{12} = 2(\ell_1\ell_2 + \ell_3\ell_4)$$

$$\xi_{13} = 2(\ell_1\ell_3 - \ell_2\ell_4)$$

$$\xi_{21} = 2(\ell_1\ell_2 - \ell_3\ell_4)$$

$$\xi_{22} = 1 - 2(\ell_1^2 + \ell_3^2)$$

$$\xi_{23} = 2(\ell_2\ell_3 + \ell_1\ell_4)$$

$$\xi_{31} = 2(\ell_1\ell_3 + \ell_2\ell_4)$$

$$\xi_{32} = 2(\ell_2\ell_3 - \ell_1\ell_4)$$

$$\xi_{33} = 1 - 2(\ell_1^2 + \ell_2^2) \quad .$$

Note that the transformation matrix in equation (B.5) is skew-symmetric.