Technical Note 1967-11

D. L. Snyder

Navigation with High-Altitude Satellites: A Study of Errors in Position Determination

6 February 1967

Prepared under Electronic Systems Division Contract AF 19(628)-5167 by

Lincoln Laboratory

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Lexington, Massachusetts

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the support of the U.S. Air Force under Contract AF 19(628)-5167.

The work reported in this document was performed at Lincoln Laboratory, a center for research operated by Massachusetts Institute of Technology, with the support of the U.S. Air Force under Contract AF 19(628)-5167.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

LINCOLN LABORATORY

NAVIGATION WITH HIGH-ALTITUDE SATELLITES: A STUDY OF ERRORS IN POSITION DETERMINATION

D. L. SNYDER

Consultant
Group 66

TECHNICAL NOTE 1967-11

6 FEBRUARY 1967

LEXINGTON MASSACHUSETTS

ABSTRACT

In this report we investigate the accuracy with which the position of a receiver can be determined by use of a high-altitude satellite navigation system. The navigation system is modelled as a problem of nonlinear estimation in the presence of random disturbances. Equations are derived for describing positioning errors by using linearization and the Kalman-Bucy filtering equations.

Accepted for the Air Force
Franklin C. Hudson
Chief, Lincoln Laboratory Office

Table of Contents

NAVIGATION WITH HIGH-ALTITUDE SATELLITES: A STUDY OF ERRORS IN POSITION DETERMINATION

1. INTRODUCTION

A possible scheme using high-altitude satellites for very accurate position determination and navigation on the earth's surface has been described by Goblick.[1] With this scheme, several satellites independently transmit timing signals generated from free-running clocks. Control stations monitor the transmission of timing signals from the satellites and compare the observed timing signals to an accurate master clock. Satellite clock timing errors are then transmitted to users by the control stations via the navigation satellites themselves and in the form of digital data superimposed on the navigation signals. The satellite clocks are thereby effectively synchronized by the control stations. Moreover, the control stations monitor the satellite positions and also relay this information to users as digital data. A user with a synchronized clock could determine his position by knowing: (a) his distance from the earth's center; (b) the satellite positions; and (c) the arrival time of the timing signals from two satellites. If the user does not have a synchronized clock, three satellites are required and differences between the arrival times of timing signals from two pairs of satellites are used.

The scheme has the advantage of requiring only passive operations by the user. In addition, position fixes can be accomplished in a relatively short period of time compared to systems employing a low-altitude satellite.

In this report we shall investigate the accuracy with which a user's position can be determined in the presence of uncertainties which inevitably exist; the uncertainties arise both because of observation noise which, for example, results in imprecise time or range measurements and because of inaccurate knowledge of the satellite positions, user's height, and timing. We consider two cases. In the first, we assume the satellites and user are motionless and that only one observation is made.* For the second case, we account for simple satellite and user motion as well as multiple observations. In addition to studying positioning errors, we also indicate the structure of processors for realizing the position fix.

The theory we shall present incorporates a linearized version of the navigation problem so that it is convenient to divide the discussion into three steps. In the first step, we consider the general linear estimation problem formulated in a way that results in the discrete filtering equations of Kalman and Bucy[2-3]. The derivation of these equations has been given by several authors[2-6], so for brevity we present only the broad points leaving certain of the mathematical subtleties to the cited references. In the second step, we consider the general nonlinear estimation problem by a technique employing linearization and the application of Kalman and Bucy's equations. This technique appears to have first been used by Smith et al.[7] More recent discussions are given by Ohap[9] and Cox[10]. Finally, in the third step, we show how the navigation problem fits into the model of the general nonlinear estimation problem and then simply apply the previously derived results.

---

* This portion of the study is based on unpublished studies performed at MIT Lincoln Laboratory. See the acknowledgment on page 35.

2. LINEAR ESTIMATION THEORY

We now postulate a model for a general linear estimation problem. The interpretation of the model is left to Section 4 where the navigation problem is presented. For the present, we simply mention that the model is a state-variable representation which can describe a wide class of discrete linear systems and their observed responses to Gaussian input sequences.

2.1 The Linear Estimation Model

Let $\underline{x}(i)$ and $\underline{r}(i)$ for $i = 1, 2, \ldots, m$ be vector sequences of Gaussian random variables defined by:

$$\underline{x}(i) = \underline{\phi}(i, i-1)\underline{x}(i-1) + \underline{G}(i)\underline{w}(i) \tag{1}$$

$$\underline{x}(0) = \underline{x}_0$$

and

$$\underline{r}(i) = \underline{H}(i)\underline{x}(i) + \underline{n}(i) \tag{2}$$

where, for $i = 1, 2, \ldots, m$:

(a) $\underline{x}(i)$ is an k-dimensional column vector. The initial value for the sequence is a random variable denoted by $\underline{x}_o$. We assume that the a priori knowledge about $\underline{x}_o$ consists of a mean, $E[\underline{x}_0] \triangleq \underline{\hat{x}}(0)$, and covariance matrix, $E[\underline{x}_0 - \underline{\hat{x}}(0)][\underline{x}_0 - \underline{\hat{x}}(0)] \triangleq \underline{V}(0)$;

(b) $\underline{\phi}(i, i-1)$ is a known transition matrix relating $\underline{x}(i)$ to $\underline{x}(i-1)$:

(c) $\underline{w}(i)$ is an ℓ-dimensional uncorrelated vector-Gaussian sequence with zero mean and known covariance matrix:

$$E[\underline{w}(i)\underline{w}'(j)] = \underline{W}(i)\,\delta_{ij} \tag{3}$$

where $\underline{W}(i)$ is a symmetric non-negative definite $\ell \times \ell$ matrix and δ_{ij} is unity when $i = j$ and zero otherwise. Deterministic and correlated input sequences can be accommodated by making $\underline{w}(i)$ have a non-zero mean and by increasing the dimension of $\underline{x}(i)$, respectively;

(d) $\underline{G}(i)$ is a known $k \times \ell$ matrix;

(e) $\underline{r}(i)$ is an observed p-dimensional vector Gaussian sequence. We shall denote the collection of observed vectors, $\underline{r}(1), \underline{r}(2), \ldots, \underline{r}(m)$, by $\underline{r}_{1,m}$;

(f) $\underline{H}(i)$ is a known $p \times k$ matrix;

(g) $\underline{n}(i)$ is a p-dimensional vector Gaussian sequence with zero mean and known covariance matrix:

$$E[\underline{n}(i)\underline{n}'(j)] = \underline{N}(i)\,\delta_{ij} \tag{4}$$

where $\underline{N}(i)$ is a symmetric positive definite $p \times p$ matrix; $\underline{n}(i)$ represents noise interferring with the observation of $\underline{H}(i)\underline{x}(i)$. By assuming $\underline{N}(i)$ to be positive definite, we are implying that no disturbance-free observations can be made.

2.2 The Linear Estimation Problem

The problem we wish to consider is that of estimating $\underline{x}(m)$, the state of the linear system at the m^{th} point in the sequence, * given a specified criterion which the estimate is to satisfy and the m observations $[\underline{r}(i): 1 \leq i \leq m] \triangleq \underline{r}_{1,m}$. We shall use a maximum a posteriori probability criterion--the estimate is that value of $\underline{x}$ maximizing $p(\underline{x}|\underline{r}_{1,m})$. However, because of the linearity and Gaussian assumptions, our estimate is identical to that resulting from other criteria such as minimum mean square error.

---

* As new data is observed, m increases in real time.

Before proceeding, it is convenient to give some indication of the approach to be taken by first considering the following two special cases of the general problem of interest.

Case I: For the first problem, we remove all time dependence and let

$$\underline{r} = \underline{H}\,\underline{x} + \underline{n} \tag{5}$$

where $\underline{x}$ and $\underline{n}$ are uncorrelated zero-mean Gaussian random variables with positive definite covariance matrices $\underline{W}$ and $\underline{N}$ respectively. The probability densities of $\underline{x}$ and $\underline{n}$ are then given by

$$p(\underline{x}) = \frac{1}{\sqrt{(2\pi)^m \det(\underline{W})}} \exp[-\tfrac{1}{2}\,\underline{x}'\,\underline{W}^{-1}\,\underline{x}] \tag{6}$$

and

$$p(\underline{n}) = \frac{1}{\sqrt{(2\pi)^p \det(\underline{N})}} \exp[-\tfrac{1}{2}\,\underline{n}'\,\underline{N}^{-1}\,\underline{n}] \tag{7}$$

Case II: For the second problem, we keep $\underline{x}$ fixed and reintroduce multiple time-dependent observations by letting

$$\underline{r}(i) = \underline{H}(i)\,\underline{x} + \underline{n}(i) \quad i = 1, 2, \ldots, m \tag{8}$$

where $\underline{x}$ and $\underline{n}(i)$, $i = 1, 2, \ldots, m$, are uncorrelated zero-mean Gaussian random variables with positive definite covariance matrices $\underline{W}$ and $\underline{N}(i)$, $i = 1, 2, \ldots, m$. The joint probability density of the vectors $\{\underline{n}(i) : 1, 2, \ldots, m\} \triangleq \underline{n}_{i,m}$ is given by:

$$p(\underline{n}_{i,m}) = \frac{1}{\prod_{i=1}^{m} \sqrt{(2\pi)^p \det[\underline{N}(i)]}} \exp\left[-\frac{1}{2} \sum_{i=1}^{m} \underline{n}'(i)\,\underline{N}^{-1}(i)\,\underline{n}(i)\right] \tag{9}$$

The first problem is illustrative of the approach and is also of interest because of its close relation to Schweppe's error analysis. The results from this problem will eventually be applied to the navigation problem when both the satellites and user are fixed and only one observation is used to determine the user's position. The second problem is presented so that the notions and advantages of recursive estimation procedures can be seen.

2.2.1 Special Case I: Time-Independent Observations

We take as the optimum estimate that value of $\underline{x}$ maximizing the a posteriori probability, $p(\underline{x}|\underline{r})$, or equivalently its logarithm, $\ln p(\underline{x}|\underline{r})$. Using Baye's rule, we have

$$\ln p(\underline{x}|\underline{r}) = \ln p(\underline{r}|\underline{x}) + \ln p(\underline{x}) + c_o$$

where c_o is a normalization constant independent of $\underline{x}$. From Eqs. 6 and 7, we then have

$$\ln p(\underline{x}|\underline{r}) = -\frac{1}{2}[\underline{r} - \underline{H}\underline{x}]' N^{-1} [\underline{r} - \underline{H}\underline{x}] - \frac{1}{2}\underline{x}' \underline{W}^{-1} \underline{x} + c_o' \quad (10)$$

We now expand the right side of Eq. 10 and then complete the square, the result being

$$\ln p(\underline{x}|\underline{r}) = -\frac{1}{2}[-\underline{x}' \underline{H}' \underline{N}^{-1} \underline{r} + \underline{x}'\{\underline{W}^{-1} + \underline{H}' \underline{N}^{-1} \underline{H}\} \underline{x} - \underline{r}' \underline{N}^{-1} \underline{H}\underline{x}] + c_o''$$

$$= -\frac{1}{2}[\underline{x} - \underline{V}\underline{H}'\underline{N}^{-1}\underline{r}]' \underline{V}^{-1} [\underline{x} - \underline{V}\underline{H}'\underline{N}^{-1}\underline{r}] + c_o'' \quad (11)$$

where we have set

$$\underline{V} = \{\underline{W}^{-1} + \underline{H}'\underline{N}^{-1}\underline{H}\}^{-1} \quad (12)$$

assuming the inverse exists.

Inspection of Eq. 11 reveals that the a posteriori density of $\underline{x}$ is normal and maximized by the conditional mean, $\underline{V}\,\underline{H}'\,\underline{N}^{-1}\,\underline{r}$. Consequently, the estimate we seek is given by

$$\hat{\underline{x}} = \underline{V}\,\underline{H}'\,\underline{N}^{-1}\,\underline{r} \tag{13}$$

Furthermore, the error covariance matrix, defined as the conditional expectation of $[\underline{x} - \hat{\underline{x}}]\,[\underline{x} - \hat{\underline{x}}]'$, is just the conditional covariance, $\underline{V}$.

It is observed from Eq. 12 that the evaluation of $\underline{V}$ requires the inversion of three matrices. An alternative expression requiring only one inversion can be given by applying the matrix relation proved in Appendix 1. The result is

$$\underline{V} = \underline{W} - \underline{W}\,\underline{H}'\,[\underline{H}\,\underline{W}\,\underline{H}' + \underline{N}]^{-1}\,\underline{H}\,\underline{W} \tag{14}$$

Since $\underline{V}$ does not depend on $\underline{r}$, it can be determined before making an observation. Consequently, the error performance can be studied without actually carrying out experiments. Furthermore, we see from Eq. 13 that, if $\underline{V}$ is determined beforehand, the computation of $\hat{\underline{x}}$ involves only a matrix multiplication of the observed data.

Collecting all results, we have

Model: $\underline{r} = \underline{H}\,\underline{x} + \underline{n}$

Estimation Equations:

$$\hat{\underline{x}} = \underline{V}\,\underline{H}'\,\underline{N}^{-1}\,\underline{r}$$

$$\underline{V} = [\underline{W}^{-1} + \underline{H}'\,\underline{N}^{-1}\,\underline{H}]^{-1}$$

$$= \underline{W} - \underline{W}\,\underline{H}'\,[\underline{H}\,\underline{W}\,\underline{H}' + \underline{N}]^{-1}\,\underline{H}\,\underline{W}$$

Let

$$\underline{e} = \underline{x} - \hat{\underline{x}}$$

be the error vector associated with $\hat{\underline{x}}$. Then the mean-square values of the components of $\underline{e}$ are the diagonal elements of $\underline{V}$; that is, $E[e_i^2] = v_{ii}$. The error vector $\underline{e}$ is a zero-mean vector Gaussian random variable with multidimensional ellipsoids as constant probability contours--the "error" ellipsoids are defined by

$$\underline{e}' \underline{V} \underline{e} = \text{constant.}$$

The squared lengths of the semiaxes of the particular ellipsoid,

$$\underline{e}' \underline{V} \underline{e} = 1,$$

are equal to the eigenvalues of $\underline{V}$ and provide a convenient coordinate-free measure of the accuracy in estimating $\underline{x}$. In some instances, the overall shape and orientation of the error ellipsoid is of interest and can be examined by determining all the eigenvalues and eigenvectors of $\underline{V}$, respectively. On the other hand, simpler measures of the error are often sufficient. One such measure is the squared length of the semimajor axis which equals the maximum eigenvalue of $\underline{V}$; this measure upper bounds $\underset{i}{\text{Max}}\,[v_{ii}]$. Another simple measure is the sum of the squared lengths of the semiaxes and equals the trace of $\underline{V}$, $\sum_{i=1}^{m} v_{ii}$.

2.2.2 Special Case II: Time-Dependent Observations

We again take as optimum that value of $\underline{x}$ maximizing the a posteriori probability, $p(\underline{x}|\underline{r}_{1,m})$, or equivalently its logarithm, $\ln p(\underline{x}|\underline{r}_{1,m})$. Using Baye's rule and Eqs. 6 and 9, we have

$$\ln p(\underline{x}|\underline{r}_{1,m}) = -\frac{1}{2} \sum_{i=1}^{m} [\underline{r}(i) - \underline{H}(i)\underline{x}]' \underline{N}^{-1}(i)[\underline{r}(i) - \underline{H}(i)\underline{x}] - \frac{1}{2}\underline{x}' \underline{W}^{-1}\underline{x} + c_o \tag{15}$$

where c_o is a normalization constant independent of $\underline{x}$. Proceeding once again through the steps leading from Eq. 10 to Eqs. 11 and 12, we obtain from Eq. 15

$$\underline{x}(m) = \underline{V}(m) \sum_{i=1}^{m} \underline{H}'(i)\underline{N}^{-1}(i)\underline{r}(i) \tag{16}$$

where

$$\underline{V}(m) = [\underline{W}^{-1} + \sum_{i=1}^{m} \underline{H}'(i)\underline{N}^{-1}(i)\underline{H}(i)]^{-1} \tag{17}$$

It can be easily shown that $\hat{\underline{x}}$ is an unbiased estimate of $\underline{x}$. $\underline{V}(m)$ is the associated error covariance matrix; that is, $\underline{V}(m)$ is the conditional expectation of $[\underline{x} - \hat{\underline{x}}(m)][\underline{x} - \hat{\underline{x}}(m)]'$.

Equations 16 and 17 are the estimation equations. Note that all the observations $\{\underline{r}(i) : 1 \le i \le m\}$ are required to compute $\hat{\underline{x}}(m)$. Data-storage requirements, therefore, increase in direct proportion to the number of observations taken. Recursive equations for $\hat{\underline{x}}(m)$ can be obtained from Eqs. 16 and 17 by simple matrix manipulation. These equations provide an updating procedure by which only the most current observation is used in conjunction with the preceding estimate, $\hat{\underline{x}}(m-1)$, and covariance matrix, $\underline{V}(m-1)$, to determine $\hat{\underline{x}}(m)$ and $\underline{V}(m)$. They have the advantage that storage requirements do not increase with each new observation. We shall derive the recursive equation for $\underline{V}(m)$ first. From Eq. 17,

$$\underline{V}(m) = [\underline{W}^{-1} + \sum_{i=1}^{m-1} \underline{H}'(i)\underline{N}^{-1}(i)\underline{H}(i) + \underline{H}'(m)\underline{N}^{-1}(m)\underline{H}(m)]^{-1}$$

$$= [\underline{V}^{-1}(m-1) + \underline{H}'(m)\underline{N}^{-1}(m)\underline{H}(m)]^{-1} \qquad (18)$$

Equation 18 is the desired relation. An alternative form can be obtained by use of the lemma of Appendix 1, the result being

$$\underline{V}(m) = \underline{V}(m-1) - \underline{V}(m-1)\underline{H}'(m)[\underline{H}(m)\underline{V}(m-1)\underline{H}'(m) + \underline{N}(m)]^{-1}\underline{H}(m)\underline{V}(m-1) \qquad (19)$$

The recursive equation for $\hat{\underline{x}}(m)$ is derived starting with Eq. 16 from which we obtain

$$\hat{\underline{x}}(m) = \underline{V}(m)\left[\sum_{i=1}^{m-1} \underline{H}'(i)\underline{N}^{-1}(i)\underline{r}(i) + \underline{H}'(m)\underline{N}^{-1}(m)\underline{r}(m)\right]$$

$$= \underline{V}(m)\underline{V}^{-1}(m-1)\hat{\underline{x}}(m-1) + \underline{V}(m)\underline{H}'(m)\underline{N}^{-1}(m)\underline{r}(m) \qquad (20)$$

From Eq. 18 it can be verified that

$$\underline{V}(m)\underline{V}^{-1}(m-1) = \underline{I} - \underline{V}(m)\underline{H}'(m)\underline{N}^{-1}(m)\underline{H}(m) \qquad (21)$$

Therefore,

$$\hat{\underline{x}}(m) = \hat{\underline{x}}(m-1) + \underline{V}(m)\underline{H}'(m)\underline{N}^{-1}(m)[\underline{r}(m) - \underline{H}(m)\hat{\underline{x}}(m-1)] \qquad (22)$$

which is the desired result.

Collecting all results, we have

Model: $r(i) = \underline{H}(i)\underline{x} + \underline{n}(i) \quad i = 1, 2, \ldots, m$

Estimation Equations:

$$\hat{\underline{x}}(m) = \hat{\underline{x}}(m-1) + \underline{V}(m)\underline{H}'(m)N^{-1}(m)[\underline{r}(m) - \underline{H}(m)\hat{\underline{x}}(m-1)]$$

$$V(m) = [\underline{V}^{-1}(m-1) + \underline{H}'(m)\underline{N}^{-1}(m)\underline{H}(m)]^{-1}$$

$$= \underline{V}(m-1) - \underline{V}(m-1)\underline{H}'(m)[\underline{H}(m)\underline{V}(m-1)\underline{H}'(m) + \underline{N}(m)]^{-1}\underline{H}(m)\underline{V}(m-1)$$

2.2.3 General Case: Time-Dependent Observations and Parameters

We now return to the original problem and derive the discrete Kalman and Bucy equations. In the following procedure, we shall derive the recursive equations directly. We begin by observing that

$$p[\underline{x}(m)|\underline{r}_{1,m}] = \frac{p[\underline{r}(m), \underline{x}(m), \underline{r}_{1,m-1}]}{p[\underline{r}(m), \underline{r}_{1,m-1}]}$$

$$= \frac{p[\underline{r}(m)|\underline{x}(m), \underline{r}_{1,m-1}]\, p[\underline{x}(m)|\underline{r}_{1,m-1}]}{p[\underline{r}(m)|\underline{r}_{1,m-1}]}$$

The first factor in the numerator can be simplified when it is noted that $\underline{r}(m)$ is only a function of $\underline{n}(m)$ when $\underline{x}(m)$ is specified. Then, because of the independence of $\underline{n}(m)$, $\underline{x}(i)$, and $\underline{n}(i)$ for $i < m$, it follows that $\underline{r}(m)$ is independent of $\underline{r}_{1,m-1}$ when $\underline{x}(m)$ is specified. Consequently,

$$p[\underline{r}(m)|\underline{x}(m), \underline{r}_{1,m-1}] = p[\underline{r}(m)|\underline{x}(m)]$$

and

$$p[\underline{x}(m)|\underline{r}_{1,m}] = \frac{p[\underline{r}(m)|\underline{x}(m)]\, p[\underline{x}(m)|\underline{r}_{1,m-1}]}{p[\underline{r}(m)|\underline{r}_{1,m-1}]} \qquad (23)$$

We, therefore, have the result

$$\ln p[\underline{x}(m)|\underline{r}_{1,m}] = \ln p[\underline{r}(m)|\underline{x}(m)] + \ln p[\underline{x}(m)|\underline{r}_{1,m-1}] + c_o \qquad (24)$$

where c_o is a normalization constant. The first term on the right is given by:

$$\ln p[\underline{r}(m)|\underline{x}(m)] = -\frac{1}{2}[\underline{r}(m) - \underline{H}(m)\underline{x}(m)]'\,\underline{N}^{-1}(m)[\underline{r}(m) - \underline{H}(m)\underline{x}(m)] + c_o' \qquad (25)$$

where again c_o' is a normalization constant. We now need to investigate the second term on the right of Eq. 24. We observe first that $\underline{x}(m)$ conditioned on $\underline{r}_{1, m-1}$ is normal with mean

$$E[\underline{x}(m) | \underline{r}_{1, m-1}] \stackrel{\Delta}{=} \hat{\underline{x}}(m|m-1) \tag{26a}$$

$$= E[\underline{\phi}(m, m-1) \underline{x}(m-1) + \underline{G}(m) \underline{w}(m) | \underline{r}_{1, m-1}]$$

$$= \underline{\phi}(m, m-1) \hat{\underline{x}}(m-1|m-1) \tag{26b}$$

where $\hat{\underline{x}}(m|m-1)$ and $\hat{\underline{x}}(m-1|m-1)$ denote the maximum a posteriori estimates of $\underline{x}(m)$ and $\underline{x}(m-1)$, respectively, based on the m-1 observations, $\underline{r}_{1, m-1}$. In deriving Eq. 26, we have used the independence of $\underline{w}(m)$ and $\underline{r}_{1, m-1}$ and the identity of the conditional mean and maximum a posteriori estimate. The conditional covariance matrix associated with $\underline{x}(m)$ is given by

$$\underline{P}(m) \stackrel{\Delta}{=} E[\underline{x}(m) - \hat{\underline{x}}(m|m-1)] [\underline{x}(m) - \hat{\underline{x}}(m|m-1)]'$$

$$= E\{\underline{\phi}(m, m-1)[\underline{x}(m-1) - \hat{\underline{x}}(m-1|m-1)] + \underline{G}(m) \underline{w}(m)\}$$

$$\times \{\underline{\phi}(m, m-1)[\underline{x}(m-1) - \hat{\underline{x}}(m-1|m-1)] + \underline{G}(m) \underline{w}(m)\}'$$

$$= \underline{\phi}(m, m-1) \underline{V}(m-1) \underline{\phi}'(m, m-1) + \underline{G}(m) \underline{W}(m) \underline{G}'(m) \tag{27}$$

where $\underline{V}(m-1)$ is the error covariance matrix defined by

$$\underline{V}(j) \stackrel{\Delta}{=} E[\underline{x}(j) - \hat{\underline{x}}(j|j)] [\underline{x}(j) - \hat{\underline{x}}(j|j)]' \tag{28}$$

for $j = m-1$.

Using these results (Eqs. 25, 26, and 27), we can express $\ln p[\underline{x}(m) | \underline{r}_{1, m}]$, defined by Eq. 24, as

$$\ln p[\underline{x}(m)|\underline{r}_{1,m}] = -\frac{1}{2}[\underline{r}(m) - \underline{H}(m)\underline{x}(m)]'\underline{N}^{-1}(m)[\underline{r}(m) - \underline{H}(m)\underline{x}(m)]$$

$$- \frac{1}{2}[\underline{x}(m) - \hat{\underline{x}}(m|m-1)]'\underline{P}^{-1}(m)[\underline{x}(m) - \hat{\underline{x}}(m|m-1)] + c_o'' \qquad (29)$$

To obtain an expression for $\hat{\underline{x}}(m|m)$, we expand Eq. 29 keeping only terms which depend on $\underline{x}(m)$ and then we complete the square. The details are straightforward and result in the following recursive expressions for $\hat{\underline{x}}(m|m)$ and $\underline{V}(m)$:

$$\hat{\underline{x}}(m|m) = \underline{V}(m)[\underline{H}'(m)\underline{N}^{-1}(m)\underline{r}(m) + \underline{P}^{-1}(m)\hat{\underline{x}}(m|m-1)] \qquad (30)$$

and

$$\underline{V}(m) = [\underline{H}'(m)\underline{N}^{-1}(m)\underline{H}(m) + \underline{P}^{-1}(m)]^{-1} \qquad (31)$$

where $P(m)$ is defined by Eq. 27.

By simple matrix manipulations, Eqs. 30 and 31 can be placed in a form having some computational advantages which are mentioned below. Observe from Eq. 30 that

$$\hat{\underline{x}}(m|m) = \underline{V}(m)[\underline{H}'(m)\underline{N}^{-1}(m)\underline{r}(m) + \{\underline{H}'(m)\underline{N}^{-1}(m)\underline{H}(m) + \underline{P}^{-1}\}\hat{\underline{x}}(m|m-1)$$

$$- \underline{H}'(m)\underline{N}^{-1}(m)\underline{H}(m)\hat{\underline{x}}(m|m-1)] \qquad (32)$$

$$= \hat{\underline{x}}(m|m-1) + \underline{V}(m)\underline{H}'(m)N^{-1}(m)[\underline{r}(m) - \underline{H}(m)\hat{\underline{x}}(m|m-1)] \qquad (33)$$

Equation 33 is the desired expression for $\hat{\underline{x}}(m|m)$. By using the lemma of Appendix 1, we obtain the desired expression for $\underline{V}(m)$ from Eq. 31

$$\underline{V}(m) = \underline{P}(m) - \underline{P}(m)\underline{H}'(m)[\underline{H}(m)\underline{P}(m)\underline{H}'(m) + \underline{N}(m)]^{-1}\underline{H}(m)\underline{P}(m) \qquad (34)$$

where

$$\underline{P}(m) = \phi(m, m-1)\underline{V}(m-1)\phi'(m, m-1) + \underline{G}(m)\underline{W}(m)\underline{G}'(m) \qquad (35)$$

The most evident computational advantage in determining $\underline{V}(m)$ by Eq. 34 rather than Eq. 31 is that Eq. 34 requires the inversion of a single $p \times p$ matrix; whereas, Eq. 31 requires the inversion of three matrices, one of order $p \times p$ and two of order $m \times m$. A less evident advantage exists because the number of observables is very often less than the order of the state vector. In this instance, $p < m$ and Eq. 34 requires the inversion of a smaller matrix than Eq. 31. The advantage of the alternative expression for $\hat{x}(m|m)$, Eq. 33 rather than Eq. 30, is that the inverse of $\underline{P}(m)$ is not required.

Just as in the previous examples, $\underline{V}(m)$ does not depend on the observed data and can, therefore, be precomputed. Then because only simple matrix additions and multiplications are required, $\hat{\underline{x}}(m|m)$ can be rapidly computed as new data becomes available. Furthermore, the error performance can be investigated without carrying out a complete simulation.

Collecting all results, we have

Model:

$$\underline{x}(i) = \underline{\phi}(i, i-1)\,\underline{x}(i-1) + \underline{G}(i)\,\underline{w}(i)$$

$$\underline{r}(i) = \underline{H}(i)\,\underline{x}(i) + \underline{n}(i)$$

for $i = 1, 2, \ldots, m$

Estimation Equations:

(1) $$\underline{x}(m|m) = \hat{\underline{x}}(m|m-1) + \underline{V}(m)\,\underline{H}'(m)\,\underline{N}^{-1}(m)[\underline{r}(m) - \underline{H}(m)\,\hat{\underline{x}}(m|m-1)]$$

where

$$\hat{\underline{x}}(m|m-1) = \underline{\phi}(m, m-1)\,\hat{\underline{x}}(m-1|m-1)$$

(2) $$\underline{V}(m) = [\underline{H}'(m)\,\underline{N}^{-1}(m)\,\underline{H}(m) + \underline{P}^{-1}(m)]^{-1}$$

$$= \underline{P}(m) - \underline{P}(m)\,\underline{H}'(m)[\underline{H}(m)\,\underline{P}(m)\,\underline{H}'(m) + \underline{N}(m)]^{-1}\,\underline{H}(m)\,\underline{P}(m)$$

where

$$\underline{P}(m) = \underline{\phi}(m, m-1)\,\underline{V}(m-1)\,\underline{\phi}'(m, m-1) + \underline{G}(m)\,\underline{W}(m)\,\underline{G}'(m)$$

Equations 33 and 34 are difference equations for the optimum estimate and were first derived by Kalman and Bucy in 1961[2,3]. The associated initial conditions, $\hat{\underline{x}}(0|0)$ and $\underline{V}(0)$, are determined from the a priori knowledge of $\underline{x}_o$, the initial state. The known a priori mean is $\hat{\underline{x}}(0|0)$, $E[\underline{x}_o] = \hat{\underline{x}}(0)$: and the a priori covariance matrix is $\underline{V}(0)$, $E[\underline{x}_o - \hat{\underline{x}}_o][\underline{x}_o - \hat{\underline{x}}(0)]' = \underline{V}(0)$.

This brief derivation of the discrete Kalman-Bucy equations ignores several important issues which range from questions of the existence of inverse matrices to the uniqueness, stability, and asymptotic behavior of $\hat{\underline{x}}(m|m)$ and $\underline{V}(m)$. These issues are discussed in detail in the cited references.

3. NONLINEAR ESTIMATION THEORY

3.1 The Nonlinear Estimation Model

The nonlinear estimation model differs from the linear estimation model only to the extent of including a possibly nonlinear transformation of the state vector in the observations. The new model is described by the relations:

$$\underline{x}(i) = \underline{\phi}(i, i-1)\,\underline{x}(i-1) + \underline{G}(i)\,\underline{w}(i) \tag{36}$$

$$\underline{x}(0) = \underline{x}_o$$

and

$$\underline{r}(i) = \underline{h}[i : \underline{x}(i)] + \underline{n}(i) \tag{37}$$

for $i = 1, 2, \ldots, m$. We assume $\underline{h}[i : \underline{x}(i)]$ is suitably restricted so that, for instance, it has a Taylor series expansion

$$\underline{h}(i : \underline{x}) = \underline{h}(i : \underline{z}) + \sum_{k}^{m} (x_k - z_k) \frac{\partial}{\partial x_k} \underline{h}(i : \underline{x}) \Big|_{\underline{x} = \underline{z}} + \text{higher order terms}$$

$$= \underline{h}(i : \underline{z}) + \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{z}} (\underline{x} - \underline{z}) + \text{higher order terms} \tag{38}$$

where $\left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{z}}$ is a matrix of partial derivatives of $\underline{h}(i : \underline{x})$ evaluated at the point $\underline{x} = \underline{z}$; the k-row, ℓ-column element of the matrix, $\frac{\partial \underline{h}}{\partial \underline{x}}$, is $\frac{\partial}{\partial x_\ell} h_k(i : \underline{x})$. This matrix is commonly referred to as the Jacobian matrix.

3.2 The Nonlinear Estimation Problem

Just as in the linear case, we seek to estimate $\underline{x}(m)$, the state at the (moving) endpoint of the observation interval, based on all the accumulated observations, $\underline{r}_{1, m}$. The procedure we shall take is the following.

We assume first that the trajectory traced out by the sequence of points $\underline{x}(1)$, $\underline{x}(2), \ldots, \underline{x}(m)$ does not differ greatly from some trajectory described, for instance, by $\underline{z}(1)$, $\underline{z}(2), \ldots, \underline{z}(m)$. Then, for each i, we expand $\underline{h}[i : \underline{x}(i)]$ in a Taylor expansion about $\underline{z}(i)$ and retain only the first two terms. The observations are thereby linearized about the trajectory $\underline{z}(1), \underline{z}(2), \ldots, \underline{z}(m)$, and the results of the preceding section (i. e., the Kalman-Bucy equations) can be used to estimate the state at the end point of the observation interval, $\underline{x}(m)$.

The choice of the trajectory about which the linearization is performed depends on the particular application. We can describe the sequence $\underline{z}(1)$, $\underline{z}(2), \ldots, \underline{z}(m)$ as being prespecified or not depending upon whether or not it is known before data is received. The technique of linearizing about a prespecified trajectory and then applying the Kalman-Bucy equations to estimate the true trajectory has been used with success in a variety of space applications in which a vehicle was controlled so as to follow a prescribed path, such as a path to the moon.* The technique is also applicable to the high-altitude satellite navigation system discussed below when the user's vehicle is controlled so as to follow a prescribed course. Such a situation arises, for example, with the point-to-point travel of commercial aircraft through prescribed air corridors.

Very often, on the other hand, no prespecified trajectory is available and $\underline{z}(i)$, for each i, must be generated as data is received. An example of this situation arises in the navigation context when the user does not follow a prescribed course as in a tactical or evasive maneuver. If the actual and estimated trajectories do not differ greatly, then it is natural to consider a

---

* The notion was introduced in this context by Smith et al. [7] and McLean et al. [8].

linearization about the estimated trajectory itself. That is, for each i, we can let

$$\underline{z}(i) = \hat{\underline{x}}(i \mid i) \qquad i = 1, 2, \ldots, m \tag{39a}$$

where $\hat{\underline{x}}(i \mid i)$ is the estimate of $\underline{x}(i)$ based on all the data available at time i, $\underline{r}_{1,i}$.

We shall see below that this choice for $\underline{z}(i)$, Eq. 39a, leads to nonlinear equations for the estimate and that these equations must be solved simultaneously. Since solving simultaneous, nonlinear equations is generally difficult, even with the aid of a computer, we shall propose an alternative choice for $\underline{z}(i)$ that obviates this particular problem. The alternative choice will work nearly as well as that of (39a) under most circumstances, certainly most circumstances where a linearization procedure will work at all. Let

$$\underline{z}(i) = \hat{\underline{x}}(i \mid i-1) \tag{39b}$$

where $\hat{\underline{x}}(i \mid i-1)$ is the extrapolated or predicted value of $\underline{x}(i)$ based on $\underline{r}_{1,i-1}$ which includes all of the data available prior to time i. The advantage of this choice is that while we are still led to nonlinear equations, they can be solved sequentially in a straightforward manner.

The sequence $\underline{z}(1)$, $\underline{z}(2), \ldots, \underline{z}(m)$ can also be viewed as the mean or the conditional mean (i. e., estimated mean of the stochastic sequence $\underline{x}(1), \underline{x}(2), \ldots, \underline{x}(m)$. From this viewpoint, a "prespecified" sequence $\underline{z}(i)$ corresponds to the case of a known mean from which it is assumed $\underline{x}(i)$, $i = 1, 2, \ldots, m$, does not differ greatly. A non-prespecified sequence $\underline{z}(i)$ then corresponds to the case of an unknown mean that must be estimated as data is accumulated.

Let us now illustrate these procedures by considering the nonlinear estimation problem. First we shall examine a special case paralleling the linear estimation problem of Section 2.2.1; that is, we shall remove all time dependence and employ only one observation. Then we shall reintroduce time dependence and treat the general nonlinear problem.

3.2.1 Special Case: Time-Independent Observations

Let

$$\underline{r} = \underline{h}(\underline{x}) + \underline{n} \tag{40}$$

where $\underline{x}$ and $\underline{n}$ are Gaussian random variables with known means, $\underline{x}_0$ and $\underline{0}$, and covariance matrices, $\underline{W}$ and $\underline{N}$, respectively. Our reason for choosing this special case is that it models the high-altitude satellite navigation system discussed in Section IV when only one observation is taken so that changes in the relative position of the user and satellites need not be considered. We assume that $\underline{x}$ does not deviate significantly from its mean, $\underline{x}_0$, and then expand $\underline{h}(\underline{x})$ about this point. To a close approximation, we then have

$$\underline{r} - \underline{h}(\underline{x}_o) = \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{x}_o} (\underline{x} - \underline{x}_o) + \underline{n} \tag{41}$$

We now want to apply the results of the linear estimation problem of Section 2.2.1. To this end, we note that $\underline{r} - \underline{h}(\underline{x}_o)$ may be taken as the observed signal, $\left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{x}_o}$ as the matrix $\underline{H} = \underline{H}(\underline{x}_o)$, and $\underline{x} - \underline{x}_o$ as a zero mean Gaussian random variable. It then follows from Eq. 13 that

$$\underline{x}^* = \underline{V}^*(\underline{x}_o) \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{x}_o} \underline{N}^{-1} [\underline{r} - \underline{h}(\underline{x}_o)] \tag{42}$$

where the asterisk on $\underline{x}^*$ indicates that the estimate is an approximation to the optimum estimate, $\hat{\underline{x}}$; we shall refer to $\underline{x}^*$ as being "quasi-optimum."

The matrix, $\underline{V}^*(\underline{x}_o)$, is given from Eq. 12 by

$$\underline{V}^*(\underline{x}_o) = \left[\underline{W}^{-1} + \left(\frac{\partial h}{\partial \underline{x}}\right)'_{\underline{x}_o} \underline{N}^{-1} \left(\frac{\partial h}{\partial \underline{x}}\right)_{\underline{x}_o}\right]^{-1} \tag{43}$$

An alternative expression for $\underline{V}^*(\underline{x}_o)$ can be obtained by using Eq. 14.

The error covariance matrix $\underline{V}$, defined by

$$\underline{V} = E[\underline{x} - \underline{x}^*][\underline{x} - \underline{x}^*]' \tag{44}$$

equals, or nearly equals, $\underline{V}^*(\underline{x}_o)$ when the error, $\underline{x} - \underline{x}^*$, is small.† More generally, the Cramer-Rao bound[11] can be used to demonstrate that $\underline{V}$ satisfies the inequality

$$\underline{V} \geq \underline{V}^*(\underline{x}_o) \tag{45}$$

whatever the size of the error. This inequality implies that the error ellipsoids associated with $\underline{V}$,

$$[\underline{x} - \underline{x}^*]' \underline{V} [\underline{x} - \underline{x}^*] = \text{constant}, \tag{46}$$

never lie within those associated with $\underline{V}^*(\underline{x}_o)$

$$[\underline{x} - \underline{x}^*]' \underline{V}^* (\underline{x}_o)[\underline{x} - \underline{x}^*] = \text{constant} \tag{47}$$

for the same constant. It follows that $\underline{V}^*(\underline{x}_o)$ can be used to study the best attainable performance of $\underline{x}^*$ as an estimate of $\underline{x}$. For this purpose, various norms of $\underline{V}^*(\underline{x}_o)$ can be used. Two possible norms, which were mentioned in

---

† We say $\underline{x} - \underline{x}^*$ is small if some related norm is small. A typical norm which can be used is $\mathrm{tr}[\underline{x} - \underline{x}^*][\underline{x} - \underline{x}^*]'$.

the linear case of Section 2.2.1, are the largest eigenvalue and trace of $\underline{V}^*(\underline{x}_o)$.

3.2.2 General Case: Time-Dependent Observations

For this case, we have the observed sequence given by Eq. 37 as

$$\underline{r}(i) = \underline{h}[i:\underline{x}(i)] + \underline{n}(i) \tag{37}$$

for $i = 1, 2, \ldots, m$. We now expand $\underline{h}[i:\underline{x}(i)]$ about $\underline{z}(i)$, for each i, and then use the small error assumption. At this point, $\underline{z}(i)$ can be either pre-specified or not. We shall consider the separate cases below. To a close approximation, the result is

$$\underline{\rho}(i) \triangleq \underline{r}(i) - \underline{h}[i:\underline{z}(i)] + \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{z}(i)} \underline{z}(i) = \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{z}(i)} \underline{x}(i) + \underline{n}(i) \tag{48}$$

for $i = 1, 2, \ldots, m$. The results of the general linear estimation problem can be applied when we identify $\underline{\rho}(i)$ as the observed signal and $(\partial \underline{h}/\partial \underline{x})_{\underline{z}(i)}$ as the matrix, $\underline{H}(i)$. From Eq. 33 we then obtain

$$\underline{x}^*(m|m) = \underline{x}^*(m|m-1) + \underline{V}^*(m) \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)'_{\underline{z}(m)} \underline{N}^{-1}(m)\left\{\underline{\rho}(m) - \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{z}(m)} \underline{x}^*(m|m-1)\right\} \tag{49}$$

$$\begin{aligned} &= \underline{x}^*(m|m-1) \\ &+ \underline{V}^*(m) \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)'_{\underline{z}(m)} \underline{N}^{-1}(m)\left\{\underline{r}(m) - \underline{h}[m:\underline{z}(m)] - \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{z}(m)} \right. \\ &\left. \times \ [\underline{x}^*(m|m-1) - \underline{z}(m)]\right\} \end{aligned} \tag{50}$$

where $\underline{V}^*(m)$ is a matrix specified from Eq. 31 by

$$\underline{V}^*(m) = \left[\underline{P}^{-1}(m) + \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)'_{\underline{z}(m)} \underline{N}^{-1}(m) \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{z}(m)}\right]^{-1} \tag{51}$$

and where

$$\underline{P}(m) = \underline{\phi}(m, m-1)\,\underline{V}^*(m-1)\,\underline{\phi}'(m, m-1) + \underline{G}(m)\,\underline{W}(m)\,\underline{G}'(m).$$

An alternative expression for $\underline{V}^*(m)$ can be given by using Eq. 34. A further simplification of Eq. 50 is possible if it is assumed that $\underline{z}(m)$ does not differ greatly from the extrapolated state, $\underline{x}^*(m \mid m-1)$:

$$\underline{z}(m) \approx \underline{x}^*(m \mid m-1) \tag{52}$$

Then

$$\underline{h}[m : \underline{x}^*(m|m-1)] \approx \underline{h}[m : \underline{z}(m)] + \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{z}(m)} [\underline{x}^*(m|m-1) - \underline{z}(m)]$$

and the expression for $\underline{x}^*(m|m)$ becomes

$$\underline{x}^*(m|m) = \underline{x}^*(m|m-1) + \underline{V}^*(m)\left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{z}(m)} \underline{N}^{-1}(m)\{\underline{r}(m) - \underline{h}[m:\underline{x}^*(m|m-1)]\} \tag{53}$$

Equations 51 through 53, along with the initial conditions,

$$\underline{x}^*(0|0) = E[\underline{x}(0)] \triangleq \hat{\underline{x}}_o \tag{54}$$

and

$$\underline{V}^*(0) = E[\underline{x}(0) - \hat{\underline{x}}_0][\underline{x}(0) - \hat{\underline{x}}_0]' \triangleq \underline{V}(0) \tag{55}$$

specify the quasi-optimum estimate of the state at the endpoint of the observation interval. It can be observed now that the trajectory followed by $\underline{z}(i)$, for each i, enters into the estimate through the Jacobian matrix, $(\partial \underline{h}/\partial \underline{x})_{\underline{z}(m)}$, which occurs both in Eqs. 51 and 53.

In the case of prespecified trajectories, $\underline{z}(i)$, for each i, is known in advance so that $(\partial \underline{h}/\partial \underline{x})_{\underline{z}(m)}$ and, therefore, $\underline{V}^*(m)$ can be precomputed. We then observe that

(1) $\underline{x}^*(m|m)$ can be determined rapidly and simply as new data becomes available. This is because the "gain" matrix, $\underline{V}^*(m)(\partial\underline{h}/\partial\underline{x})_{\underline{z}(m)}\underline{N}^{-1}(m)$, can be precalculated so that the only "on-line" matrix operations required in Eq. 53 are additions and multiplications and not inversions.

(2) The error performance can be investigated using $\underline{V}^*(m)$ without making any observations or involved computer simulations.

On the other hand, $\underline{V}^*(m)$ cannot be precomputed when there is no pre-specified trajectory; that is, when $\underline{z}(i)$ must be generated as data is received. If $\underline{z}(i)$ is defined by (39a):

$$\underline{z}(i) = \underline{x}^*(i|i) \qquad i = 1, 2, \ldots, m \qquad \text{(Eq. 39a, repeated)}$$

then (51) and (53) are coupled and must be solved <u>simultaneously</u> Alternatively, if $\underline{z}(i)$ is defined by (39b),

$$\underline{z}(i) = \underline{x}^*(i|i-1) \qquad i = 1, 2, \ldots, m \qquad \text{(Eq 39b, repeated)}$$

then (51) and (53) are coupled but they can be solved <u>sequentially</u>. The error performance cannot be investigated in either of these cases without an involved computer simulation. Note also that if $\underline{x}^*(m|m)$ does differ from $\underline{x}^*(m|m-1)$ enough so that $(\partial\underline{h}/\partial\underline{x})_{\underline{x}^*(m|m)}$ is appreciably different from $(\partial\underline{h}/\partial\underline{x})_{\underline{x}^*(m|m-1)}$, then solving Eqs. (51) and (53) sequentially is not sufficient. The linearization of $\underline{h}$ about the final estimate must agree with the linearization used in the solution of these equations.

4. APPLICATION OF NONLINEAR ESTIMATION THEORY TO THE PROBLEM OF NAVIGATING WITH HIGH-ALTITUDE SATELLITES

We now want to apply the results derived in the preceding sections to simplified versions of the high-altitude satellite navigation system described in the introduction. We shall study the problem in two setps. First, we examine the simplest case where only one observation is taken so that user and satellite position changes need not be considered. Then, second, we shall introduce multiple observations and include simple user motion.

4.1 Special Case I: Fixed-Position and Single Observation

The geometry associated with the navigation system for the case of three fixed satellites, one fixed control station, and one fixed user is shown in Fig. 1. We shall use an earth-centered rectangular coordinate system. The position vectors associated with the satellites, control station, and user are defined in Table I.

	Position
Satellites j = 1, 2, 3	$\underline{p}_j$
Control Station	$\underline{p}_c$
User	$\underline{p}_u$

Table I

Let h be the user's height measured from the earth's center:

$$h = |\underline{p}_u| = [\underline{p}'_u \underline{p}_u]^{1/2} \tag{56}$$

h is nominally equal to the radius of the earth.

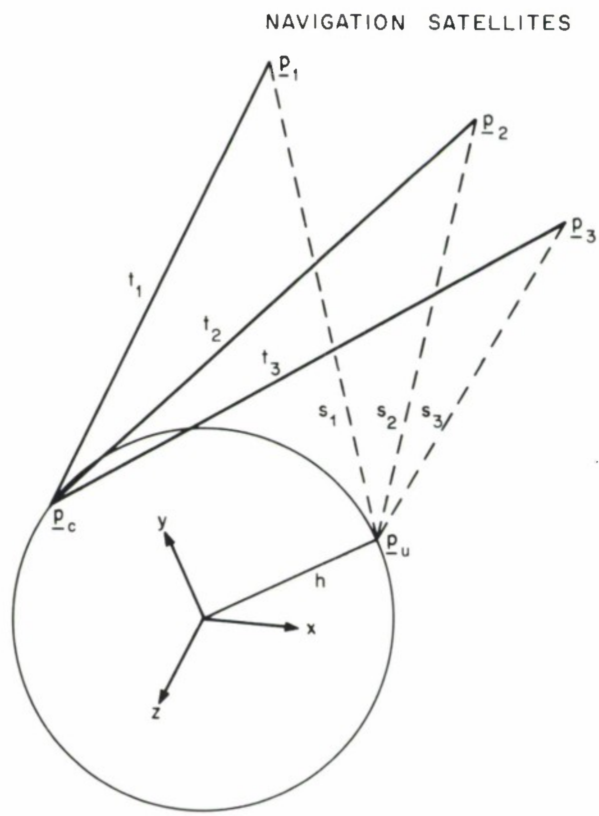

Figure 1. Geometry of the high-altitude satellite navigation system

The user-to-satellite and control-station-to-satellite distances indicated in Fig. 1 are given by

$$s_j = |\underline{p}_j - \underline{p}_u| = \{(\underline{p}_j - \underline{p}_u)'(\underline{p}_j - \underline{p}_u)\}^{1/2} \tag{57}$$

and

$$\underline{t}_j = |\underline{p}_j - \underline{p}_c| = \{(\underline{p}_j - \underline{p}_c)'(\underline{p}_j - \underline{p}_u)\}^{1/2} \tag{58}$$

for j = 1, 2, 3, respectively.

Let τ_j be the time difference between the clock of the j^{th} satellite and a master clock. The vector

$$\underline{\tau} = \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} \tag{59}$$

is then a time-error vector associated with the three satellite clocks. Similarly, let τ_o be the time error associated with the user's clock (again relative to the master clock). We assume that τ_i, for i = 0, 1, 2, and 3, is measured in distance units [see footnote (*) on next page].

The unknown parameter (or state) vector describing the navigation system is given in terms of the defined quantities as:

$$\underline{x} = \begin{bmatrix} \underline{p}_u \\ \underline{p}_c \\ \underline{p}_1 \\ \underline{p}_2 \\ \underline{p}_3 \\ \underline{\tau} \\ \tau_o \end{bmatrix} \qquad \text{(19 components)} \tag{60}$$

The user attempts to determine his position by observing the control station monitoring signals, the satellite timing signals (relative to his clock), and his height. We shall simplify the problem at this point by assuming that the control station can make very accurate--in fact, perfect--measurements of the satellite positions and satellite clock errors. In this instance, the control station positions, $\underline{p}_c$, and satellite clock errors, $\underline{\tau}$, are eliminated from the problem and the user's observable may be taken as

$$\underline{r} = \underline{h}(\underline{x}) + \underline{n} \tag{61}$$

where

$$\underline{h}(\underline{x}) = \begin{bmatrix} h_1(\underline{x}) \\ h_2(\underline{x}) \\ h_3(\underline{x}) \\ h_4(\underline{x}) \end{bmatrix} = \begin{bmatrix} s_1 + \tau_o \\ s_2 + \tau_o \\ s_3 + \tau_o \\ h \end{bmatrix} = \begin{bmatrix} |\underline{p}_1 - \underline{p}_u| + \tau_o \\ |\underline{p}_2 - \underline{p}_u| + \tau_o \\ |\underline{p}_3 - \underline{p}_u| + \tau_o \\ |\underline{p}_u| \end{bmatrix}$$

and $\underline{n}$ is the observation noise with an assumed covariance

$$E[\underline{n}\,\underline{n}'] = \begin{bmatrix} \sigma_r^2 & & & 0 \\ & \sigma_r^2 & & \\ & & \sigma_r^2 & \\ 0 & & & \sigma_h^2 \end{bmatrix}$$

Here σ_r^2 is the ranging-error variance that results, for example, from receiver noise, propagation effects, and short-term clock instabilities; σ_h^2 is the height error variance.

The vector $\underline{x}$, defined by:

---

* (from previous page) Just as radar echo delays are converted to distances by using the speed of light.

$$\underline{x} = \begin{bmatrix} \underline{p}_u \\ \underline{p}_1 \\ \underline{p}_2 \\ \underline{p}_3 \\ \tau_o \end{bmatrix} \qquad \text{(13 components)}$$

is assumed to have a mean $E(\underline{x}) = \underline{x}_0$ which accounts for the nominal positions of the satellite and user and also the nominal value of the user's clock error. The a priori error covariance matrix is defined by

$$\underline{W} = E(\underline{x} - \underline{x}_0)(\underline{x} - \underline{x}_0)' \qquad \text{(dimension } 13 \times 13\text{)}$$

The diagonal components of $\underline{W}$ are the a priori variances associated with the satellite and user positions and with the user's clock.

Define the Jacobian matrix by

$$\frac{\partial \underline{h}}{\partial \underline{x}} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_{13}} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_2}{\partial x_{13}} \\ \vdots & \vdots & & \\ \frac{\partial h_4}{\partial x_1} & \frac{\partial h_4}{\partial x_2} & \cdots & \frac{\partial h_4}{\partial x_{13}} \end{bmatrix} \tag{62}$$

Then, from (43), $\underline{V}^*(\underline{x}_o)$, the error covariance matrix associated with estimating $\underline{x}$ when $E[\underline{x}] = \underline{x}_o$, is given by

$$\underline{V}^*(\underline{x}_0) = \left[\underline{W}^{-1} + \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)'_{\underline{x}_0} \underline{N}^{-1} \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{x}_0}\right]^{-1} = \underline{W} - \underline{W}\left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)'_{\underline{x}_0}$$

$$\times \left[\left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{x}_0} \underline{W} \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)'_{\underline{x}_0} + \underline{N}\right]^{-1} \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{x}_0} \underline{W} \tag{63}$$

Equation (63) is being used in a computer study of the navigation problem for the special case described. Various user positions, $\underline{x}_0$, and a priori statistics, $\underline{W}$, are being examined. The results will appear in a companion report.

4.2 Special Case II: Simple Motion and Multiple Observations

In this example, we shall make the following assumptions:

(i) The control station can make very accurate measurements of the satellite positions and clock errors; this removes $\underline{p}_c$ and $\underline{\tau}$ from the problem.

(ii) There are three fixed satellites.

(iii) The user undergoes motion along some nominal prescribed trajectory; that is, his mean trajectory is known. His "state,"

$$\underline{s}_u(i) = \begin{bmatrix} \underline{p}_u(i) \\ \text{- - - - -} \\ \underline{v}_u(i) \end{bmatrix} \qquad i = 1, 2, \ldots, m \tag{64}$$

consisting of his position and velocity, satisfies

$$\underline{s}_u(i) = \underline{\Phi}_u(i, i-1)\, \underline{s}_u(i-1) + \underline{G}_u(i)\, \underline{w}_u(i) \qquad i = 1, 2, \ldots, m \tag{65}$$

where $\underline{\Phi}_u(i, i-1)$, $\underline{G}_u(i)$, and $\underline{w}_u(i)$ are the same as defined by Eq. (1) of Section 2.1 except that here we allow $\underline{w}_u(i)$ to have a nonzero mean. The nonzero mean of $\underline{w}_u(i)$ is such that $\underline{s}_u(i)$, $i = 1, 2, \ldots, m$, follows the prescribed trajectory on the average; that is, we are assuming that the (known) mean of the stochastic sequence $\underline{s}_u(i)$, $i = 1, 2, \ldots, m$, is the prescribed trajectory about which the linearization is carried out.

(iv) The user's clock is stable but has a fixed offset, τ_o, from true time (as indicated on a master clock).

(v) Multiple observations are taken.

The state vector $\underline{x}(i)$, i = 1, 2, . . . , m, associated with the estimation problem is defined by

$$\underline{x}(i) = \begin{bmatrix} \underline{p}_1(i) \\ \underline{p}_2(i) \\ \underline{p}_3(i) \\ \underline{s}_u(i) \\ \tau_o \end{bmatrix} \qquad \text{(16 components)} \tag{66}$$

where, because of the above assumptions,

$$\underline{x}(i) = \begin{bmatrix} \underline{I} & \underline{0} & \underline{0} & \underline{0} & 0 \\ \underline{0} & \underline{I} & \underline{0} & \underline{0} & 0 \\ \underline{0} & \underline{0} & \underline{I} & \underline{0} & 0 \\ \underline{0} & \underline{0} & \underline{0} & \underline{\phi}_u(i, i-1) & 0 \\ \underline{0} & \underline{0} & \underline{0} & \underline{0} & 1 \end{bmatrix} \underline{x}(i-1) + \begin{bmatrix} \underline{0} \\ \underline{0} \\ \underline{0} \\ \underline{G}_u(i) \\ \underline{0} \end{bmatrix} \underline{w}_u(i)$$

$$\triangleq \underline{\phi}(i, i-1)\,\underline{x}(i-1) + \underline{G}(i)\,\underline{w}_u(i) \qquad \text{for } i = 1, 2, \ldots, m \tag{67}$$

The observed sequence satisfies

$$\underline{r}(i) = \underline{h}[\underline{x}(i)] + \underline{n}(i) \qquad i = 1, 2, \ldots, m \tag{68}$$

where

$$\underline{h}[\underline{x}(i)] = \begin{bmatrix} |\underline{p}_1 - \underline{p}_u(i)| + \tau_o \\ |\underline{p}_2 - \underline{p}_u(i)| + \tau_o \\ |\underline{p}_3 - \underline{p}_u(i)| + \tau_o \\ |\underline{p}_u(i)| \end{bmatrix} \tag{69}$$

and $\underline{n}(i)$ is the observation noise with an assumed covariance

$$E[\underline{n}(i)\,\underline{n}'(j)] = \begin{bmatrix} \sigma_r^2 & & & 0 \\ & \sigma_r^2 & & \\ & & \sigma_r^2 & \\ 0 & & & \sigma_h^2 \end{bmatrix} \delta_{ij} \qquad (70)$$

We have assumed for this simple model that the additive observation noise is an uncorrelated sequence. In practice, correlated sequences would also be encountered and these would be treated by expanding the dimension of $\underline{x}(i)$.

Define the Jacobian matrix by

$$\frac{\partial}{\partial \underline{x}} \underline{h}[\underline{x}(i)] = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_{13}} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_2}{\partial x_{13}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial h_4}{\partial x_1} & \frac{\partial h_4}{\partial x_2} & & \frac{\partial h_4}{\partial x_{13}} \end{bmatrix}_{\underline{x}(i)} \qquad (71)$$

Then, from (51), the error covariance matrix associated with estimating $\underline{x}(m)$ when $E[\underline{x}(m)] = \underline{z}(m)$, is given by

$$\underline{V}^*(m) = \left[\underline{P}^{-1}(m) + \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)'_{\underline{z}(m)} \underline{N}^{-1}(m) \left(\frac{\partial \underline{h}}{\partial \underline{x}}\right)_{\underline{z}(m)} \right]^{-1} \qquad (72)$$

where

$$\underline{P}(m) = \underline{\Phi}(m, m-1)\, \underline{V}^*(m-1)\, \underline{\Phi}'(m, m-1) + \underline{G}(m)\, \underline{W}(m)\, \underline{G}'(m)$$

and $\underline{W}(m)$ is the a priori covariance matrix defined by

$$\underline{W}(m) = E[\underline{w}(m) - E\,\underline{w}(m)]\,[\underline{w}(m) - E\,\underline{w}(m)]'$$

Note that $\underline{V}^*(m)$ is determined from the difference equation (72) by starting with the initial condition $\underline{V}^*(0) = \underline{V}_0$, the a priori error covariance matrix, and serially calculating $\underline{V}^*(1)$, $\underline{V}^*(2), \ldots, \underline{V}^*(m)$.

Equation (72) is being used in a computer study of the navigation problem for the special case described. Various simple user trajectories and a priori statistics, $\underline{w}(m)$, are being examined. The results will appear in a companion report.

Preliminary Conclusions

We have indicated with two simple examples how the nonlinear estimation procedure of section 3 can be used for the problem of navigating with high-altitude satellites. Several questions that must be addressed before the technique can be applied successfully to the more general navigation problem are:

(1) Models must be developed for describing the motion of satellites placed in a near-synchronous orbit and for characterizing the errors made in determining their coordinates;

(2) Models must be developed for statistically describing the effects of timing errors due to atmospheric effects, such as refraction. These effects generally depend on the relative user-to-satellite position. Consequently, the associated covariance matrices will depend on the state $(\underline{x})$;

(3) An investigation of the sensitivity of the navigation model to changes in assumptions or parameters (such as the a priori error covariance matrix) should be made.

Appendix

We wish to establish the lemma:

If

$$\underline{S}^{-1} = \underline{T}^{-1} + \underline{M}'\,\underline{U}^{-1}\,\underline{M} \tag{A1}$$

where $\underline{S}^{-1}$, $\underline{T}^{-1}$, and $\underline{U}^{-1}$ are symmetric positive-definite matrices; then

$$\underline{S} = \underline{T} - \underline{T}\,\underline{M}'[\underline{M}\,\underline{T}\,\underline{M}' + \underline{U}]^{-1}\,\underline{M}\,\underline{T} \tag{A2}$$

Proof:

The existence of the inverse matrices required for Eq. A2 follows from the positive-definiteness of $\underline{S}^{-1}$, $\underline{T}^{-1}$, and $\underline{U}^{-1}$. We establish the Lemma by direct substitution.

$$\underline{S}\,\underline{S}^{-1} = [\underline{T} - \underline{T}\,\underline{M}'(\underline{M}\,\underline{T}\,\underline{M}' + \underline{U})^{-1}\,\underline{M}\,\underline{T}]\,[\underline{T}^{-1} + \underline{M}'\,\underline{U}^{-1}\,\underline{M}]$$

$$= \underline{I} - \underline{T}\,\underline{M}'[(\underline{M}\,\underline{T}\,\underline{M}' + \underline{U})^{-1}\,\underline{M} - \underline{U}^{-1}\,\underline{M} + (\underline{M}\,\underline{T}\,\underline{M}' + \underline{U})^{-1}\,\underline{M}\,\underline{T}\,\underline{M}'\,\underline{U}^{-1}\,\underline{M}]$$

$$= \underline{I} - \underline{T}\,\underline{M}'[(\underline{M}\,\underline{T}\,\underline{M}' + \underline{U})^{-1}\,\underline{U} - \underline{I} + (\underline{M}\,\underline{T}\,\underline{M}'\ [(\underline{U})^{-1}\,\underline{M}\,\underline{T}\,\underline{M}']\ \underline{U}^{-1}\,\underline{M}$$

$$= \underline{I} - \underline{T}\,\underline{M}'[(\underline{M}\,\underline{T}\,\underline{M}' + \underline{U})^{-1}(\underline{M}\,\underline{T}\,\underline{M}' + \underline{U}) - \underline{I}]\ \underline{U}^{-1}\,\underline{M}$$

$$= \underline{I}$$

References

1. T. J. Goblick, Jr., "Navigation with High-Altitude Satellites: A Study of Ranging Errors," Lincoln Laboratory Technical Note 1966-46 (26 August 1966).

2. R. Kalman and R. Bucy, "New Results in Linear Filtering and Prediction Theory," ASME J. of Basic Engineering 83, 95-108 (March 1961).

3. R. Kalman, "New Methods and Results in Linear Prediction and Filtering Theory," RIAS Technical Report 61-1 (1961).

4. F. C. Schweppe, "An Introduction to Estimation Theory for Dynamical Systems," Lincoln Laboratory Group Report 22G-15 (August 1963).

5. H. Rauch, F. Tung, and C. Striebel, "Maximum Likelihood Estimates of Linear Dynamic Systems," AIAA Journal 3, 1445-1450 (August 1965).

6. Y. Ho and R. Lee, "A Bayesian Approach to Problem in Stochastic Estimation and Control," IEEE Trans. on Automatic Control, 333-339 (October 1964).

7. G. L. Smith, S. F. Schmidt, and L. A. McGee, "Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle," NASA Technical Report R-135 (1962).

8. J. McLean, S. F. Schmidt, and L. A. McGee, "Optimal Filtering and Linear Prediction Applied to a Mid-Course Navigation System for the Circumlunar Mission," NASA TN D-1208 (1962).

9. R. Ohap and A. Stubberud, "A Technique for Estimating the State of a Nonlinear System," IEEE Trans. on Automatic Control, 150-155 (April 1965).

10. H. Cox, "Recursive Nonlinear Filtering," Proc. of the National Electronics Conference 21, 770-775 (1965).

11. H. Cramer, Mathematical Methods of Statistics (Princeton University Press, Princeton, New Jersey, 1946).

ACKNOWLEDGMENT

This study is based in part on unpublished studies of T. J. Goblick, Jr., B. Reiffen, F. C. Schweppe, and R. Teoste. The navigation model used in Section 4.1 is originally due to F. C. Schweppe. I am grateful to T. J. Goblick, Jr., for many helpful discussions and his comments on early versions of the report.

DOCUMENT CONTROL DATA – R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Lincoln Laboratory, M.I.T.	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP None

3. REPORT TITLE

Navigation with High-Altitude Satellites: A Study of Errors in Position Determination

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Note

5. AUTHOR(S) *(Last name, first name, initial)*

Snyder, Donald L.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
6 February 1967	42	13

8a. CONTRACT OR GRANT NO. AF 19(628)-5167 *b.* PROJECT NO. 649L *c.* *d.*	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Note 1967-11 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)* ESD-TR-67-60

10. AVAILABILITY/LIMITATION NOTICES

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
None	Air Force Systems Command, USAF

13. ABSTRACT

In this report we investigate the accuracy with which the position of a receiver can be determined by use of a high-altitude satellite navigation system. The navigation system is modelled as a problem of nonlinear estimation in the presence of random disturbances. Equations are derived for describing positioning errors by using linearization and the Kalman-Bucy filtering equations.

14. KEY WORDS

satellite navigation
position determination