NASA CONTRACTOR REPORT

NASA CR-1004

NASA CR-1004

GUIDANCE, FLIGHT MECHANICS AND TRAJECTORY OPTIMIZATION

Volume V - State Determination and/or Estimation

by G. E. Townsend, D. R. Grier, R. R. Palmer, R. J. Ruggiero, and A. S. Abbott

Prepared by
NORTH AMERICAN AVIATION, INC.
Downey, Calif.
for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • APRIL 1968

NASA CR-1004

GUIDANCE, FLIGHT MECHANICS AND TRAJECTORY OPTIMIZATION

Volume V - State Determination and/or Estimation

By G. E. Townsend, D. R. Grier, R. R. Palmer,
R. J. Ruggiero, and A. S. Abbott

Issued by Originator as Report No. SID 65-1200-5

Prepared under Contract No. NAS 8-11495 by
NORTH AMERICAN AVIATION, INC.
Downey, Calif.

for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

FOREWORD

This report was prepared under contract NAS 8-11495 and is one of a series intended to illustrate analytical methods used in the fields of Guidance, Flight Mechanics, and Trajectory Optimization. Derivations, mechanizations and recommended procedures are given. Below is a complete list of the reports in the series.

Volume I	Coordinate Systems and Time Measure
Volume II	Observation Theory and Sensors
Volume III	The Two Body Problem
Volume IV	The Calculus of Variations and Modern Applications
Volume V	State Determination and/or Estimation
Volume VI	The N-Body Problem and Special Perturbation Techniques
Volume VII	The Pontryagin Maximum Principle
Volume VIII	Boost Guidance Equations
Volume IX	General Perturbations Theory
Volume X	Dynamic Programming
Volume XI	Guidance Equations for Orbital Operations
Volume XII	Relative Motion, Guidance Equations for Terminal Rendezvous
Volume XIII	Numerical Optimization Methods
Volume XIV	Entry Guidance Equations
Volume XV	Application of Optimization Techniques
Volume XVI	Mission Constraints and Trajectory Interfaces
Volume XVII	Guidance System Performance Analysis

The work was conducted under the direction of C. D. Baker, J. W. Winch, and D. P. Chandler, Aero-Astro Dynamics Laboratory, George C. Marshall Space Flight Center. The North American program was conducted under the direction of H. A. McCarty and G. E. Townsend.

TABLE OF CONTENTS

LIST OF SYMBOLS

$A(t)$	matrix of coefficients relating $\dot{\vec{\delta}}$ and $\vec{\delta}$ (section 2.4)
A	linear transformation matrix (section 2.3)
a	semi-major axis for ellipse
B	matrix relating to the observables acquired at various epochs to a single epoch for the purpose of estimation (section 2.2) $B = \begin{bmatrix} H_1 \varphi(t_1, t_o) \\ \vdots \\ H_N \varphi(t_n, t_o) \end{bmatrix}$
	matrix relating errors in the observables to errors in the math model utilizes to generate the observables (section 2.5)
$B(\ )$	Bayes function
$b(\Theta)$	parameter bias
$\hat{C}, \hat{U}, \hat{S}, \hat{X}$	universal variables
cov (.)	covariance matrix
D	planet diameter
E	expected value (section 2.3)
E	Covariance matrix for the estimation errors (section 2.2)
E	eccentric anomaly for elliptic motion (section 2.4)
$\underline{e}$	error vector for observation process (section 2.3)
e	eccentricity for ellipse
$F_1(t, t_o)$ $F_2(t, t_o)$ $F_3(t, t_o)$	components of the position error for the case of circular motion as obtained by direct integration
$\underline{F}$	general vector function (section 2.3)
$\vec{F}(t)$	column vector of forcing functions relative to nominal trajectory (section 2.4)
$f(x)$	probability density function of X

f, g	functions of time utilized to relate $\vec{r}$ and $\vec{v}$ as functions of time in elliptic hyperbolic motion
H	matrix of partials of the observables with respect to the state
$\vec{h}$	angular momentum vector
[I]	identity matrix
i	orbital inclination
J	E^{-1}; utilized in constructing the recursive minimum variance estimator (section 2.2)
L( )	loss function (section 2.3)
L	latitude
$\underline{\ell}$	unit vector
M	mean anomaly
$\underline{dM}$	the vector of model uncertainties
m	sample size
$\underline{m}$	unit vector
$\hat{N}$	nodal unit vector
$\underline{N}_2$	vector from observer to occultation point on planet
n	mean motion
$\underline{n}$	unit vector
$\underline{o}$	unit vector
P( )	probability
p	semilatus rectum for ellipse
"pdf"	probability density function
Q	optimum weighting matrix for computing the state deviation vector as employed in the minimum variance approaches (section 2.2)
q_i	the i^{th} measured value of a navigation measurement
q_{i_o}	the nominal value of the i^{th} navigation measurement

$R(\delta, \theta)$	the risk resulting from an estimate of δ for θ
R	covariance matrix for the errors in the observables
$\underline{R_T}$	tracking station location
$\underline{R_V}$	vehicle position with respect to tracking station
$\vec{R}$	observer's position vector (section 2.1)
$\hat{R}, \hat{S}, \hat{W}$	radial, circumferential and normal (along h) unit vectors in rotating local coordinates
$\vec{r}$	radius vector
T	a statistic; i.e., a known function of a random sample (section 2.3)
T_s	sufficient statistic (section 2.3)
t	time measured from some convenient reference
$\vec{u}$	control vector
V	covariance matrix for estimation errors
$\underline{V}$	vehicle velocity vector
V	$\lvert\dot{\vec{r}}\rvert$
$\underline{V_p}$	velocity vector of a planet
x	$E-E_o$ elliptic motion hyperbolic motion; $F-F_o$ hyperbolic motion
$\vec{x}$	state vector $\left\{\vec{r}/V\right\}$
$\hat{x}, \hat{y}, \hat{z}$	unit vectors for a general inertial coordinate system
y	outcome of a statistical process
$\underline{\alpha}$	vector of errors in the observables ($\underline{\delta A}$) (section 2.2)
β	azimuth relative to north point on the horizon (section 2.4)
γ	$\sin^{-1}(\hat{r} \cdot \hat{v})$
δ	estimator for the parameters (section 2.3)
δq_i	deviation of the $i^{\underline{th}}$ measurement from nominal
$\delta\underline{R}$	position deviation vector

$\vec{\delta}$	state deviation (section 2.4)
$\delta\underline{x}$	vector of state deviations (sometimes referred to as δ) (section 2.2)
$\delta\underline{A}$	vector of observed minus computed residuals of the observables (section 2.2)
ΔA_{STAR}	uncertainty in star direction
ϵ	estimation error ($\delta - \theta$) (section 2.3)
$\underline{d\epsilon}$	the vector of instrument uncertainties
θ	parameter being estimated (section 2.3)
θ	true anomaly for elliptic motion (section 2.4)
$\Lambda(t, t_o)$	the matrix analogous to $\varphi(t, t_o)$ in the adjoint equations
μ	mean of a statistical process (section 2.3)
μ	GM; i.e., the gravitation constant for the central force field
$\vec{\rho}$	position of the observed body relative to the observer
$\hat{\rho}$	unit vector (section 2.5)
σ^2	variance of a statistical process
τ	orbital period for elliptic motion (section 2.4)
τ	time relative to some specified epoch in units such that the gravitation constant is 1 $[= \tau (t - t_o)]$
φ	$\cos^{-1} (r \cdot N)$
$\varphi(t, t_o)$	state transition matrix relating the errors in the state at t and t_o (section 2.4)
$\Phi_{I, II}$	composite state transition matrix from epoch II to epoch I (section 2.5)
Ψ	angle measured perpendicular to the plane of motion
Ω	right ascending node
ω	d/dt (θ) for circular motion (section 2.4)
ω	argument of periapse for elliptic or hyperbolic motion (section 2.4)

ω	spin rate for the earth (section 2.5)
$\underline{\omega}$	rotation vector of earth

SUPERSCRIPTS

$\rightarrow$	vector
$\wedge$	unit vector (sections 2.1, 2.4)
$\wedge$	universal variable (section 2.4)
$\wedge$	computed or estimated quantity (section 2.2)
$\sim$	observed (section 2.2)
$\sim\sim\sim$	expected value matrix (section 2.2)
$'$	a'priori estimate (section 2.2)
$\overline{\quad}$	average or mean (section 2.3)
$\cdot$	d/dt
(n)	denotes estimate based on all data processed through the epoch corresponding to the n^{th} piece of data
T	transpose
-1	inverse

SUBSCRIPTS

$\underline{\quad}$	vector
B	Bayes
i, n	variable index denoting sequence
LS	least squares
MV	minimum variance
o	initial
p	periapse

LIST OF FIGURES

LIST OF TABLES

1.0 STATEMENT OF THE PROBLEM

The basic problem to which this monograph will be addressed is the generalization of precise information regarding the trajectory of an observed vehicle and the coefficients of the math model employed in the prediction from a set of observations (different quantities, or the same type of measurement at different times) made of the vehicle. However, to completely define the nature of the problem to be discussed, it is necessary to outline the assumptions which are implicit in the analysis. First, the observations are assumed to have been contaminated with noise (errors) resulting from random errors in the instruments and recording devices utilized for the measurements. Thus, if α^* denotes the true value of any measured quantity and ϵ the corresponding error, then the observable (α) is

$$\alpha_i = \alpha_i^* + \epsilon_i$$

Next, the number of linearly independent measurements is assumed to be greater than the number of the parameters being estimated. This assumption precludes the situations where no solution exists or where the solution is unique and leaves the general problem of obtaining "best" estimates from the over-determined system of equations. Finally, it is assumed that in the absense of the contaminating noise, a unique solution would result which would relate the observables at the various measurement times.

The rigorous problem implicit in this set of assumptions is highly nonlinear. This nonlinearity is produced by the fact that most of the observables (for example, range, range-rate...) and the components of the instantaneous position and velocity vectors (the estimated parameters) are both very complex functions of time; as a matter of fact, no analytic solution exists which relates this set of parameters as a function of time. Thus, the means employed to generate the desired estimate (commonly referred to as the method of differential corrections) will be iterative in nature. The first step in the process will be to define a nominal trajectory by some set of parameter $(\vec{r}_o, \vec{v}_o)$ which represents the observations within some allowable error.

This definition avoids most of the mechanization problems by assuming that all of the nonlinear effects can be included in the reference trajectory with sufficient accuracy to allow subsequent computations to be precise. The second step is the approximation of the dynamics and the observation processes utilizing only the linear terms of the Taylor series as in Reference 1.1. That is:

$$\delta\vec{P}(t) = \varphi(t,t_o)\,\delta\vec{P}(t_o)$$

where $\vec{P}$ denotes the ordered vector of parameters being estimated; where $\varphi(t, t_o)$ is on n by n matrix of partial derivatives

$$\delta\vec{\alpha}(t) = H(t)\,\delta\vec{P}(t) + \vec{\epsilon}(t)$$

of the parameters at time t with respect to the same set of parameters at time t_o; where H(t) is the matrix of partial derivatives of the observables with respect to the parameters being estimated at the epoch of the observation; and where ϵ (t) is the vector of errors in the true observables.

Finally estimates of the parameters at some selected epoch (T) will be generated. These estimates will be selected such that some measure of "goodness" in the estimator is maximized when the available information regarding the statistics of the errors is provided.

The discussions of this monograph will be ordered to answer questions which arise regarding each of the steps in this process and will relate in detail the nature of the problem. To accomplish this objective, large amounts of the open literature have been reviewed. Though this material is generally referenced throughout the text to provide additional information on topics being discussed, some of the more pertinent references will be quoted in the following paragraphs to aid in establishing the nature of the discussions.

The initial investigations will be directed to the task of generating a reliable first approximation to the true trajectory. This step will be performed by utilizing the material presented in a previous monograph (Ref. 1.3) and classical work, principally of Laplace and Gauss. In this material, the true trajectory is approximated by a nearly equivalent conic section to obtain the position and velocity vectors which, if the force field were central, would yield the subset of the observables used to define the conic section. The solution is discussed in detail and precautionary steps which will assume more reliable solutions are presented. Thus, if no previous estimate of the trajectory or data from the vehicles guidance system at burnout are available, an accurate initial estimate can be generated.

The discussions will then turn to the development of the "optimum" estimates of the deviation vector $\delta\vec{P}(t)$. Particular attention will center on the development of simple measures of the degree of optimality in the estimator and the generation of the estimation equations and estimation error for these measures. These discussions will parallel much of the material presented in the open literature, though some of the steps are different to facilitate comprehension of the simplest physical process. The classical least squares, weighted least squares, and minimum variance estimators will be derived; then attention will turn to modern estimation in a recursive mode. The concepts of Kalman as presented in Reference 0.1 (subsequently adopted by Schmidt in Reference 0.2) and Battin (Reference 0.3) will be reviewed carefully since estimation in this mode is capable of correcting for some of the approximations made in developing the estimator itself. This latter observation is the result of the fact that the true trajectory (i.e., nonlinear) can be approximated by a series of discontinuous arcs, each of which obeys the linear model of the dynamics, to a better degree than a single arc satisfying the same linear model.

The filter concepts outlined in the previous paragraph are based on intuitive measures of "optimality." Further, they are tailored to problems where the statistics involved are Gaussian, where the dynamics can be adequately approximated by the linear model, and where the optimum estimator is a linear

function of the deviations in the observables. Thus, the problem of estimation is reintroduced in a more complete analysis to explain the exact nature of the material which it follows, and to demonstrate the mechanism whereby some of the simplifications just enumerated can be eliminated. In the process, the problem is demonstrated to be equivalent to that presented by Middleton in Reference 0.4. This material, while requiring a reasonable knowledge of statistical concepts, ties the general estimation problem into a verified analytic framework which is capable of demonstrating the effects of the availability of all data pertaining to the process.

Having thoroughly explored the general problem of estimation, attention turns to the development of material required to yield an estimate of the trajectory. To be specific, the matrices relating the dynamics at various times relative to the nominal trajectory (the State Transition Matrix denoted by $\varphi(t, t_o)$ and the matrix presenting the error data for the observables are derived. The first of these developments progresses from the basic formulation of the transition matrix (for example, Reference 0.5) to the generation of an analytic form for the case of conic motion (for example, References 0.6, 0.7, 0.8, 0.9). This development presents several alternate representations of the desired matrix and discusses the weaknesses in them. The second development is an extension of the material presented in a previous monograph (Reference 1.1) which shows the functions involved in the process and refers to error data available in the literature for construction of the weighting matrix.

The monograph concludes with a set of recommendations for the application of this material and one possible mechanization which will be selected to utilize the maximum amount of information in the data and to minimize the computational problems.

2.0 STATE OF THE ART

2.1 INITIAL ESTIMATES OF THE ORBIT

Since the computational rationale proposed for determining precise values of the instantaneous elements for the space trajectory is built upon the concept of differential corrections, care must be exercised to assure that the initial estimates of the nominal trajectory are sufficiently precise to allow all of the partial derivatives to be evaluated accurately and to assure that the estimation error lies within the neighborhood about the true trajectory which is small enough for the process to converge. The purpose of this discussion will be to develop several such techniques and to discuss the sources of error. To be specific, the methods of Laplace and Gauss as well as methods involving the use of range and range-rate data will be presented in detail. The utilization of position and velocity information obtained directly from an integrating accelerometer (Reference 1.1) will not be discussed at this time since this information can be be utilized only for those cases in which the trajectory is to be estimated from the epoch of injection to any other reference epoch (any other possibility requires updating burnout conditions to the epoch of problem initiation) and since the data thus provided need no further transformation (i.e., they can be utilized directly if transmitted to the ground or fed to the on-board computer).

2.1.1 Data Provided Include Range, Azimuth and Elevation at Two Epochs

For the case in which the ground-based radar utilized for tracking the satellite provides range, azimuth and elevation (or an equivalent set of data) as a function of time, the logic employed to obtain an estimate of the trajectory can be relatively simple. First, the station's position at the two epochs is computed (Reference 1.2) from

$$\vec{r}_T = \left[\frac{R_e}{\sqrt{\left(\frac{R_p}{R_e}\right)^2 \sin^2 L + \cos^2 L}} + H\right]\left[U_1 \hat{X} + U_2 \hat{Y}\right] + \left[\frac{R_p^2}{\sqrt{R_p^2 \sin^2 L + R_e \cos^2 L}} + H\right] U_3 \hat{Z} \qquad (1.1)$$

where R_e, R_p = equatorial and polar radii of the reference ellipsoid for the earth

H = altitude of the station relative to the reference ellipsoid

L = Geodetic latitude of the station

U_1, U_2, U_3 = 3 components of the unit vector from the station outward along the direction inclined by the angle L to the equatorial plane

$\hat{U}, \hat{E}, \hat{N}$ = the up, east, north unit vectors

$$\begin{Bmatrix} \hat{U} \\ \hat{E} \\ \hat{N} \end{Bmatrix} = \begin{bmatrix} \cos L & 0 & \sin L \\ 0 & 1 & 0 \\ -\sin L & 0 & \cos L \end{bmatrix} \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{Bmatrix} \hat{X} \\ \hat{Y} \\ \hat{Z} \end{Bmatrix}$$

α = right ascension of the station

= G.H.A. plus longitude

Then, the position of the satellite relative to the station is computed in the $\hat{U}, \hat{E}, \hat{N}$ directions and transformed by substituting for $\hat{U}, \hat{E}, \hat{N}$ in terms of $\hat{X}, \hat{Y}, \hat{Z}$.

$$\vec{\rho} = \rho \cos E_l \sin A_z \hat{E} + \rho \cos E_l \cos A_z \hat{N} + \rho \sin E_l \hat{U} \qquad (1.2)$$

At this point, the position vectors are defined as

$$\vec{r} = \vec{r}_T + \vec{\rho} \qquad (1.3)$$

and the velocity information is derived by employing Lambert's Theorem (Reference 1.3).

2.1.2 Data Provided Include Only Azimuth and Elevation (or Equivalent Data) at Three Epochs

For the case in which only angular data are available, a complete reformulation of the problem is necessary. However, as in most of the problems discussed in this monograph, there is no unique means of reducing the data. The discussions of subsequent paragraphs will present two such schemes: the methods of Laplace and Gauss.

These techniques were derived primarily for the case in which the orbit being determined was central relative to a body other than that utilized for the observations (the sun). Thus, when each technique has been prepared for the case for which it was originally intended, it will be extended to the case of primary interest - geocentric motion.

2.1.2.1 Laplace's Method

The discussions of Reference 1.3 showed that six arbitrary constants

were required to uniquely determine the motion of a body in a central force field. Thus, if the true force field is approximated by that produced by the dominant mass (or in the case of motion relative to the earth, by that produced by neglecting those terms arising from the nonspherical shape of the earth), a conic trajectory can be found utilizing three sets of observations composed of angular data (azimuth-elevation, right ascension-declination, etc.).

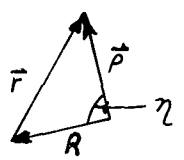
Consider the vector diagram and the corresponding equations below:

$$\vec{r} + \vec{R} = \rho\hat{\rho}$$

$$\dot{\vec{r}} + \dot{\vec{R}} = \dot{\rho}\hat{\rho} + \rho\dot{\hat{\rho}}$$

$$\ddot{\vec{r}} + \ddot{\vec{R}} = -\left(\frac{\vec{r}}{r^3} + \frac{\vec{R}}{R^3}\right) = \ddot{\rho}\hat{\rho} + 2\dot{\rho}\dot{\hat{\rho}} + \rho\ddot{\hat{\rho}}$$

$$= \frac{-\hat{\rho}}{r^3} + \left(\frac{1}{r^3} - \frac{1}{R^3}\right)\vec{R} = \ddot{\rho}\hat{\rho} + 2\dot{\rho}\dot{\hat{\rho}} + \rho\ddot{\hat{\rho}} \qquad (1.4)$$



where $\hat{\rho}$ and $\vec{R}$ at the three epochs are known or observable and where the units of time have been selected so that the gravitational constant is one [i.e., $\tau = \sqrt{\mu}\,(t-t_0)$]. Now if the three observations are acquired over a sufficiently small interval of time, the geometry of the problem can be approximated by expanding $\hat{\rho}$ in a Taylor series as follows:

$$\hat{\rho} = \hat{\rho}_o + \sum_{n=1}^{\infty} \frac{\tau^n}{n!} \left.\frac{d^n\hat{\rho}}{dt^n}\right|_{\tau_o} \qquad (1.5)$$

$$= \hat{\rho}_o + \tau\dot{\hat{\rho}}_o + \tfrac{1}{2}\tau^2\ddot{\hat{\rho}}_o + \ldots$$

This series will be terminated at the third term so that the three observations will completely define an initial estimate of $\vec{R}$ and $\dot{\vec{R}}$. Further, the second time point will be utilized for time reference so that the maximum value of τ is kept small (to assure the maximum accuracy in equation (1.5)).

Assume that three values of $\hat{\rho}$ are utilized in conjunction with the series expansion (1.5) to yield values for $\hat{\rho}_o$ and $\dot{\hat{\rho}}_o$. Now crossing $\hat{\rho}$ into $\dot{\hat{\rho}}$ and dotting into Equation (1.4) yields

$$\left(\frac{1}{r_o^3} - \frac{1}{R_o^3}\right)\left(\hat{\rho}_o \times \dot{\hat{\rho}}_o \cdot \vec{R}_o\right) = \rho_o\left[\hat{\rho}_o \times \dot{\hat{\rho}}_o \cdot \ddot{\hat{\rho}}_o\right] \qquad (1.6)$$

In like manner, cross $\hat{\rho}$ into $\ddot{\hat{\rho}}$ and dot Equation (1.4)

$$\left(\frac{1}{r_o^3} - \frac{1}{R_o^3}\right)\left(\hat{\rho}_o \times \ddot{\hat{\rho}}_o \cdot \vec{R}_o\right) = 2\dot{\rho}_o\left[\hat{\rho}_o \times \ddot{\hat{\rho}}_o \cdot \dot{\hat{\rho}}_o\right] = -2\dot{\rho}_o\left[\hat{\rho}_o \times \dot{\hat{\rho}}_o \cdot \ddot{\hat{\rho}}_o\right] \quad . \tag{1.7}$$

Now the procedure is to iterate Equation (1.6) and the law of cosines

$$\vec{r} = \vec{\rho} - \vec{R}$$

or

$$r^2 = \rho^2 + R^2 - 2R\rho\cos\eta \tag{1.8}$$

to solve for the correct value of $|\vec{r}|$ at τ_o. This value of r can then be utilized in Equation (1.7) to solve for $\dot{\rho}_o$; $\vec{r}_o$ can be found from

$$\dot{\vec{r}}_o + \dot{\vec{R}}_o = \dot{\rho}_o\hat{\rho}_o + \rho_o\dot{\hat{\rho}}_o$$

The nature of the simultaneous solution of Equations (1.6) and (1.8) is explored in some detail in Reference 1.4. This material develops an iteration procedure based on a single angular variable. The result of this procedure is an iteration process which can rely on graphical techniques for initial estimates of the parameter being estimated.

If the central body is the earth rather than the sun, differences in the formulation arise due to the fact that the acceleration of the observer is incorrect. For this case

$$\ddot{\vec{R}} = \frac{d^2}{dt^2}\left(-\vec{r}_\tau\right) = +\omega^2\vec{r}_\tau$$

and Equation (1.6), which was solved iteratively with the law of cosines, becomes

$$\left(\frac{1}{r_o^3} + \omega^2\right)\left(\hat{\rho}_o \times \dot{\hat{\rho}}_o \cdot \vec{r}_{\tau_o}\right) = \rho_o\left[\hat{\rho}_o \times \dot{\hat{\rho}}_o \cdot \ddot{\hat{\rho}}_o\right] \quad . \tag{1.6a}$$

Similarly, Equation (1.7), which was solved for $\dot{\rho}_o$, becomes

$$\left(\frac{1}{r_o^3} + \omega^2\right)\left(\hat{\rho}_o \times \ddot{\hat{\rho}}_o \cdot \vec{r}_{\tau_o}\right) = -2\dot{\rho}_o\left[\hat{\rho}_o \times \dot{\hat{\rho}}_o \cdot \ddot{\hat{\rho}}_o\right] \quad . \tag{1.7a}$$

The largest source of error in the process is the truncation of Equation (1.5) at the third term. This step means that the values of $\vec{r}_o$ and $\vec{V}_o$

which are obtained from the process will not represent the conic providing the three observations to the best degree. Thus, it is generally desirable to differentially correct these vectors before assuming that a solution is known. This process is readily accomplished utilizing the material presented in Section 2.4 of this monograph since

$$\Delta \vec{X} = \varphi(t, t_o)\, \Delta \vec{X}_o \tag{1.9}$$

and

$$\Delta \vec{y} = H \Delta \vec{X} \tag{1.10}$$

where $\Delta \vec{y}$ = a vector of observed minus computed residuals at the three epochs (τ_1, τ_2, τ_3)

H = the matrix relating errors in the observables to small errors in the position and velocity

$\Delta \vec{X}$ = a vector of position and velocity deviations ($d\vec{r}$, $d\vec{v}$)

Thus, the vector of errors at the epoch τ_2 ($\Delta \vec{X}_o$) can be estimated as

$$\Delta \vec{X}_o = \left[H \varphi\right]^{-1} \Delta \vec{y} \tag{1.11a}$$

and the previously computed values of $\vec{r}_o$ and $\vec{V}_o$ corrected

$$\vec{r}_o = \vec{r}_o + \Delta \vec{r}_o \tag{1.11b}$$

$$\vec{V}_o = \vec{V}_o + \Delta \vec{V}_o \tag{1.11c}$$

This second estimate of $\vec{r}_o$, $\vec{V}_o$ can now be utilized to generate a new error vector ($\Delta \vec{y}$) so that the process can continue until convergence is achieved.

2.1.2.2 Gauss's Method

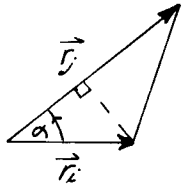
In Laplace's method, the approximations made to facilitate the solution were in the truncation of the Taylor series for $\hat{\rho}$. Gauss approached the problem by making an approximation in the dynamics rather than in the geometry. The method proceeds as follows: Since the motion is assumed planar, any of the three radii can be expressed as some linear function of the other two, i.e.,

$$\vec{r}_2 = c_1\vec{r}_1 + c_3\vec{r}_3 \qquad (\vec{r}_1 \neq k\vec{r}_3) \tag{1.12}$$

Thus

$$\vec{r}_2 \times \vec{r}_3 = c_1\vec{r}_1 \times \vec{r}_3$$

$$c_1 = \frac{\vec{r}_2 \times \vec{r}_3}{\vec{r}_1 \times \vec{r}_3}$$



But the area of the triangle of sides $\vec{r}_i$ and $\vec{r}_j$ is

$$A = [r_i, r_j] = 1/2 \text{ base x altitude}$$
$$= 1/2\ r_j\ (r_i \sin\alpha)$$
$$= 1/2\ |\vec{r}_i \times \vec{r}_j|$$

so that an alternate form of the solution for C_1 is

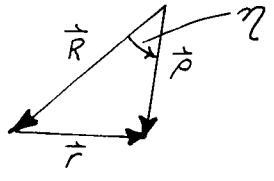
$$c_1 = \frac{[r_2, r_3]}{[r_1, r_3]} \tag{1.13}$$

Similarly

$$c_3 = \frac{\vec{r}_1 \times \vec{r}_2}{\vec{r}_1 \times \vec{r}_3} = \frac{[r_1, r_2]}{[r_1, r_3]} \tag{1.14}$$

These two scalar constants must now be related to the dynamics so that they can be determined. As a first step, consider the relation

$$\vec{r} = \vec{\rho} - \vec{R}$$



or its equivalent

$$\vec{\rho}_2 - \vec{R}_2 = c_1(\vec{\rho}_1 - \vec{R}_1) + c_3(\vec{\rho}_3 - \vec{R}_3) \tag{1.15}$$

or

$$c_1\vec{\rho}_1 - \vec{\rho}_2 + c_3\vec{\rho}_3 = c_1\vec{R}_1 - \vec{R}_2 + c_3\vec{R}_3 \quad .$$

This equation yields three linearly independent scalar relations in the unknowns $c_1, c_3, \vec{\rho}_i\ (i = 1,3)$.

At this point the dynamics of the problem are introduced in the form

$$\vec{r} = f\vec{r}_o + g\vec{v}_o \tag{1.16}$$

where $\vec{r}_0$, $\vec{V}_0$ will be taken to be the position and velocity of the observed body at the second of the three observation epochs. However, it is inconvenient to introduce the exact functional forms for f and g as presented in Reference 1.3 since this procedure would require knowledge of the parameters being estimated. Rather, series representations for these functions will be employed. reference 1.4 gives the form of these series through terms of the order T^2 (higher order terms would introduce parameters of the motion for which no estimate exists at this point in the solution) as

$$f_i = 1 - \tfrac{1}{2}\sigma T_j^3 \tag{1.17a}$$

$$g_i = T_j\left(1 - \tfrac{1}{6}\sigma T_j^3\right) \tag{1.17b}$$

$$\sigma = 1/r_2^3 \tag{1.17c}$$

where

$$T_1 = \sqrt{\mu}\,(t_3 - t_2)$$
$$T_2 = \sqrt{\mu}\,(t_3 - t_1)$$
$$T_3 = \sqrt{\mu}\,(t_2 - t_1)$$

Now the following cross products can be formed by employing Equations (1.16)

$$\vec{r}_2 \times \vec{r}_3 = g_3\vec{r}_2 \times \dot{\vec{r}}_2$$

$$\vec{r}_1 \times \vec{r}_2 = g_1\vec{r}_2 \times \dot{\vec{r}}_2 = -g_1\dot{\vec{r}}_2 \times \vec{r}_2$$

$$\vec{r}_1 \times \vec{r}_3 = \left(f_1\vec{r}_2 \times g_1\dot{\vec{r}}_2\right) \times \left(f_3\vec{r}_2 + g_3\dot{\vec{r}}_2\right)$$

$$= \left(f_1g_3 - f_3g_1\right)\vec{r}_2 \times \dot{\vec{r}}_2$$

and the scalar constants C_1 and C_3 evaluated.

$$C_1 = \frac{g_3}{f_1g_3 - f_3g_1} \tag{1.18a}$$

$$C_3 = \frac{-g_1}{f_1g_3 - f_3g_1} \tag{1.18b}$$

But, since series expressions have been employed in each of the f and g, the various factors in these expressions can be developed in series in powers of T. This step is accomplished as follows:

$$f_3 g_1 = \left(1 - \tfrac{1}{2}\sigma T_1^2\right)\left[-T_3\left(1 - \tfrac{1}{6}\sigma T_3^2\right)\right]$$

$$= -T_3 + \tfrac{1}{6}\sigma T_3^3 + \tfrac{1}{2}\sigma T_1^2 T^3$$

$$f_1 g_3 = T_1 - \tfrac{1}{6}\sigma T_1^3 - \tfrac{1}{2}\sigma T_1 T_3^2$$

where second-order terms in σ have been dropped. Thus,

$$f_1 g_3 - f_3 g_1 = (T_1 + T_3) - \tfrac{1}{6}\sigma\left(T_1^3 + 3T_1^2 T_3 + 3T_1 T_3^2 + T_3^3\right)$$

$$= (T_1 + T_3) - \tfrac{1}{6}\sigma (T_1 + T_3)^3$$

$$= T_2 - \tfrac{1}{6}\sigma T_2^3 = T_2\left(1 - \tfrac{1}{6}\sigma T_2^2\right)$$

and

$$C_1 = \frac{T_1\left(1 - \tfrac{1}{6}\sigma T^2\right)}{T_2\left(1 - \tfrac{1}{6}\sigma T_2^2\right)} \approx \frac{T_1}{T_2}\left[1 + \tfrac{1}{6}\sigma\left(T_2^2 - T_1^2\right)\right] \qquad (1.19a)$$

$$C_3 = \frac{T_1\left(1 - \tfrac{1}{6}\sigma T^2\right)}{T_2\left(1 - \tfrac{1}{6}\sigma T_2^2\right)} \approx \frac{T_1}{T_2}\left[1 + \tfrac{1}{6}\sigma\left(T_2^2 - T_3^2\right)\right] \qquad (1.19b)$$

Finally, dotting Equation (1.15) by $\hat{\rho}_1 \times \hat{\rho}_3$ yields

$$-\rho_2\left[\hat{\rho}_2 \cdot \hat{\rho}_1 \times \hat{\rho}_3\right] = C_1\left[\vec{R}_1 \cdot \hat{\rho}_1 \times \hat{\rho}_3\right] - \vec{R}_2 \cdot \hat{\rho}_1 \times \hat{\rho}_3 + C_3\left[\vec{R}_3 \cdot \hat{\rho}_1 \times \hat{\rho}_3\right] \qquad (1.20)$$

where the unknowns in Equations (1.19a), (1.19b), and (1.20) are ρ_2 and σ (i.e., $1/r^3$). As was the case with the approach of Laplace, a solution to the problem is possible by iterating Equation (1.20) with the law of cosines. The velocity is then determined from Equation (1.16). As before, the solution can be differentially corrected to compensate for assumptions made in the development of the first estimate of the solution.

In contrast to the method of Laplace, however, the method of Gauss is readily applicable to the task of determining geocentric orbits. Only a

change in notation is required.

$$\vec{R} = -\vec{r}_T \quad .$$

2.1.2.3 Modified Gauss and Laplace Methods

The formulations presented in the previous sections discuss the nature of the solution required to yield values of the position and velocity vectors at some epoch for the case where components of a unit vector from the observer toward the tracked body are given at three times. The procedures followed in these cases where either the dynamics or the geometry was approximated were detailed, and the nature or source of the errors was discussed.

There are variations of these two techniques, however, which have been developed and mechanized. Several of these are discussed in References 1.4 and 1.5. While some of these formulations have definite merit, they generally add nothing to the knowledge of the processes being employed and can thus be deleted in the presentation of different techniques of analysis.

2.1.3 Range and Range-Rate Data

For geocentric satellites where there is an opportunity to accurately measure the range and range-rate (Doppler) of the vehicle with respect to the observer, an alternate logic is required. One version of the required mechanization will be developed in the following paragraphs to facilitate comparison with the previous material.

Consider once again the equation

$$\vec{\rho} = \vec{r} + \vec{R}$$

or its scalar equivalent

$$\rho^2 = \vec{r}\cdot\vec{r} + 2\vec{r}\cdot\vec{R} + \vec{R}\cdot\vec{R} \tag{1.21}$$

and the first derivative with respect to time

$$\rho\dot{\rho} = \vec{r}\cdot\dot{\vec{r}} + \vec{r}\cdot\dot{\vec{R}} + \dot{\vec{r}}\cdot\vec{R} + R\cdot\dot{R} \tag{1.22}$$

But, the dynamics of this problem can be expressed as a function of time utilizing the f, g series (as was the case before) by expanding about the second of three observations as:

$$\vec{r} = f\vec{r}_0 + g\dot{\vec{r}}_0$$
$$\dot{\vec{r}} = \dot{f}\vec{r}_0 + \dot{g}\dot{\vec{r}}_0$$

Thus,

$$\vec{r}\cdot\dot{\vec{r}} = f\dot{f}r_0^2 + \vec{r}_0\cdot\dot{\vec{r}}_0(f\dot{g}+g\dot{f}) + g\dot{g}V_0^2$$

$$\vec{r}\cdot\dot{\vec{R}} = f\,\vec{r}_0\cdot\dot{\vec{R}} + g\,\dot{\vec{r}}_0\cdot\dot{\vec{R}}$$

$$\dot{\vec{r}}\cdot\vec{R} = \dot{f}\,\vec{r}_0\cdot\vec{R} + \dot{g}\,\dot{\vec{r}}_0\cdot\vec{R}$$

$$\vec{r}\cdot\vec{r} = f^2r_0^2 + 2fg\,\vec{r}_0\cdot\dot{\vec{r}}_0 + g^2V_0^2$$

$$\vec{r}\cdot\vec{R} = f\,\vec{r}_0\cdot\vec{R} + g\,\dot{\vec{r}}_0\cdot\vec{R}$$

Substitution of these approximations into the equations for range and range-rate yields

$$\rho^2 = f^2r_0^2 + 2fg\,\vec{r}_0\cdot\dot{\vec{r}}_0 + g^2V_0^2 + 2(f\vec{r}_0\cdot\vec{R} + g\,\dot{\vec{r}}_0\cdot\vec{R}) + \vec{R}\cdot\vec{R} \qquad (1.23)$$

and

$$\rho\dot{\rho} = f\dot{f}r_0^2 + \vec{r}_0\cdot\dot{\vec{r}}_0(f\dot{g}+g\dot{f}) + g\dot{g}V_0^2 + f\,\dot{\vec{r}}_0\cdot\dot{\vec{R}} + g\,\dot{\vec{r}}_0\cdot\dot{\vec{R}} + \dot{f}\,\vec{r}_0\cdot\vec{R} + \dot{g}\,\dot{\vec{r}}_0\cdot\vec{R} + \vec{R}\cdot\dot{\vec{R}} \qquad (1.24)$$

which can, in turn, be written as

$$\rho^2 - \vec{R}\cdot\vec{R} = ar_0^2 + b\,\vec{r}_0\cdot\dot{\vec{r}}_0 + cV_0^2 + \vec{r}_0\cdot\vec{P} + \dot{\vec{r}}_0\cdot\vec{Q} \qquad (1.25)$$

$$\rho\dot{\rho} - \vec{R}\cdot\dot{\vec{R}} = dr_0^2 + e\,\vec{r}_0\cdot\dot{\vec{r}}_0 + hV_0^2 + \vec{r}_0\cdot\vec{S} + \dot{\vec{r}}_0\cdot\vec{T} \qquad (1.26)$$

where

$$a = f^2 \qquad d = f\dot{f}$$
$$b = 2fg \qquad e = f\dot{g} + g\dot{f}$$
$$c = g^2 \qquad h = g\dot{g}$$
$$\vec{P} = 2f\vec{R} \qquad \vec{S} = f\dot{\vec{R}} + \dot{f}\vec{R}$$
$$\vec{Q} = 2g\vec{R} \qquad \vec{T} = g\dot{\vec{R}} + \dot{g}\vec{R}$$

Thus, since the only unknowns are the components of $\vec{r}_o$, $\vec{v}_o$ (the coefficients are approximately known functions of time), three sets of these equations (three epochs) will yield a solution.

This solution process is simplified for some cases of geocentric satellite motion, where the three observations are acquired over a relatively small time period, since several approximations can be made to simplify the problem. First, it is noted that since the time interval on either side of the second observation is small, the position vector for the observer ($\vec{R}$ (t)) can be approximated by $\vec{R}$ (t_o). Secondly, it is noted that the scalar product $\vec{r} \cdot \dot{\vec{R}}$ is small at each of the three epochs for most of the problems of interest, since the two vectors are nearly normal. Finally, because the time interval is small and because $\vec{r} \cdot \dot{\vec{R}}$ is small, the scalar product can be approximated by the term $g\vec{r}_o \cdot \dot{\vec{R}}$ (one of the terms of $\vec{r} \cdot \dot{\vec{R}}$). For this case, a reasonable approximation of the components of $\vec{r}_o$ and $\vec{v}_o$ can be obtained by rewriting the solution in the form

$$\begin{bmatrix} \rho_1^2 - R_1^2 \\ \rho_2^2 - R_2^2 \\ \rho_3^2 - R_3^2 \\ \rho_1\dot{\rho}_1 - \vec{R}_1\cdot\dot{\vec{R}}_1 \\ \rho_2\dot{\rho}_2 - \vec{R}_2\cdot\dot{\vec{R}}_2 \\ \rho_3\dot{\rho}_3 - \vec{R}_3\cdot\dot{\vec{R}}_3 \end{bmatrix} = \begin{bmatrix} f_1^2 & 2f_1g_1 & g_1^2 & 2f_1 & 2g_1 & 0 \\ f_2^2 & 2f_2g_2 & g_2^2 & 2f_2 & 2g_2 & 0 \\ f_3^2 & 2f_3g_3 & g_3^2 & 2f_3 & 2g_3 & 0 \\ f_1\dot{f}_1 & f_1\dot{g}_1 + g_1\dot{f}_1 & g_1\dot{g}_1 & \dot{f}_1 & \dot{g}_1 & g_1 \\ f_2\dot{f}_2 & f_2\dot{g}_2 + g_2\dot{f}_2 & g_2\dot{g}_2 & \dot{f}_2 & \dot{g}_2 & g_2 \\ f_3\dot{f}_3 & f_3\dot{g}_3 + g_3\dot{f}_3 & g_3\dot{g}_3 & \dot{f}_3 & \dot{g}_3 & g_3 \end{bmatrix} \begin{bmatrix} r_o^2 \\ \vec{r}_o\cdot\dot{\vec{r}}_o \\ v_o^2 \\ \vec{r}_o\cdot\vec{R}_o \\ \dot{\vec{r}}_o\cdot\vec{R}_o \\ \dot{\vec{r}}_o\cdot\dot{\vec{R}}_o \end{bmatrix}$$

or

$$\vec{Y} = [M]\ \vec{X} \tag{1.27}$$

Thus, the parameters denoted by $\vec{X}$ can be evaluated by inverting the matrix M as

$$\vec{X} = [M]^{-1} \vec{Y} \tag{1.28}$$

The components of $\vec{r}_o$, $\vec{v}_o$ are then obtained in a straight forward manner.

As was the case with the methods of Gauss and Laplace, a differential corrections process is required to adjust for errors in the representation of $\vec{r}$ and $\vec{v}$ as functions of the time from the reference epoch and in the assumptions made to obtain the first approximation of $\vec{r}_o$, $\vec{v}_o$.

The case in which range or range-rate data alone are acquired at six epochs can, of course, also yield values of $\vec{r}_o$, $\vec{v}_o$. These applications are special cases, however, of the material presented in the preceding paragraphs, and will therefore not receive special attention.

2.1.4 Precautionary Numerical Operations

Regardless of the approach taken in developing the initial estimate of the position and velocity vectors, the quality of the solution will be dependent on the quality of the data utilized and the time interval between the observations (the sensitivity of the solution to errors in the observables will increase as the time interval decreases). Thus, it is essential to assure that as many of the errors as possible have been adjusted. In particular,

1. the affects of refraction
2. the affects of signal propagation time etc.
3. inclusion in the observations and instrument biases
4. recording errors
5. etc.

must be determined and compensated for. However, these steps in themselves will not assure a good estimate since the sensors utilized for the observations are not perfect.

Therefore, normal practice utilizes a preliminary smoothing of the data acquired over three intervals of time to produce three estimates of the true observables in the sense of least squares (or weighted least squares). This smoothing can be accomplished by fitting the data to a line (if the intervals are small) but is normally accomplished by employing a parabola. The process is mechanized as follows; (assuming that the data in one interval obey the equation)

$$y_i = a + b t_i + c t_i^2 \tag{1.29}$$

where y_i = ith observed value of one component of the observation vector in the interval $A < t < B$

a, b, c = coefficients of parabola utilized for the purposes of smoothing the data

A matrix equation is now prepared

$$\vec{Y} = \begin{bmatrix} 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_n & t_n^2 \end{bmatrix} \begin{Bmatrix} a \\ b \\ c \end{Bmatrix} \equiv T \begin{Bmatrix} a \\ b \\ c \end{Bmatrix} \quad) \tag{1.30}$$

and the least-squares estimate of the coefficients are generated (Section 2.2)

$$\begin{Bmatrix} a \\ b \\ c \end{Bmatrix} = \left[T^T T \right]^{-1} T^T \vec{Y}$$

When this process has been performed for each component of observation vector in the interval $A \leq t \leq B$, a smoothed estimate of the vector is prepared at an epoch in the interval (normally, $t = (B - A)/2$). At this point, the process is repeated for the other two intervals of time. As a final output then, there are three sets of smoothed estimates of the observables which, when utilized, will produce generally superior values of $\vec{r}_o$, $\vec{v}_o$.

2.2 ORBIT IMPROVEMENT

2.2.1 Introduction

The basic process for determining the position and velocity deviations from an estimated trajectory involves the measurement of any position, velocity or time dependent set of parameters and the construction of the linear system of equations relating the observables to the parameters being estimated at the time of measurement.

$$\delta \underline{A}(t) = H(t)\, \delta \underline{X}(t) \tag{2.1}$$

[In this notation, $\delta \underline{A}$ (t) is the m-vector of observed minus computed values of the observables; H (t) is an m by n (n is normally 6) matrix of partial derivatives of the observations with respect to the state which was developed in Reference 1.1; and $\delta \underline{X}$ (t) is the n-vector of state deviations. These equations normally do not completely determine the state since the observables collected at any one epoch generally number less than the number of components in the vector $\delta \underline{X}$ (t). Thus, data collected at different epochs are referenced to some standard epoch through the use of the state transition matrix (Section 2.4) as

$$\delta \underline{X}(t) = \varphi(t, t_o)\, \delta \underline{X}(t_o)$$

So

$$\delta \underline{A}(t) = H(t)\, \varphi(t, t_o)\, \delta \underline{X}(t_o) \tag{2.2}$$

and the task becomes one of generating an estimate of $\delta \underline{X}$ (t_0).

The estimation of $\delta \underline{X}$ (t_0) for the special case of an evenly determined set of data can, of course, be performed by simply inverting the set of equations. For example, if

$$\delta A = \begin{Bmatrix} \delta a_1 \\ \delta a_2 \\ \vdots \\ \delta a_6 \end{Bmatrix} = \begin{bmatrix} H(t_1)\varphi(t_1, t_o) \\ H(t_2)\varphi(t_2, t_o) \\ \vdots \\ H(t_6)\varphi(t_6, t_o) \end{bmatrix} \begin{bmatrix} \delta x_1 \\ \delta x_2 \\ \vdots \\ \delta x_6 \end{bmatrix} \tag{2.3}$$

$$\equiv B\, \delta \underline{X}(t_o)$$

where δa_i denote scalar quantities and where the t_i are not necessarily unique. Then,

$$\delta \underline{X}(t_o) = B^{-1}\, \delta \underline{A}$$

or

$$\delta \underline{X}(t) = \varphi(t, t_o)\, B^{-1} \delta \underline{A} \tag{2.4}$$

However, since errors exist in the instrument utilized to perform the measurements, and in the mathematical model utilized to compute both the H (t) matrix and the computed values of the observables (used directly to define the observed minus computed residuals), the true and computed values of the state deviation vector $\underline{\delta X}\ (t_o)$ will differ. Consequently, it will be necessary to distinguish between the three types of deviations employed in this analysis (actual, measured, computed). This distinction will be accomplished by adopting the notation

$$\delta\underline{\tilde{A}} = \delta\underline{A} + \underline{\alpha}$$

where $(\sim)$ denotes measured and where $\underline{\alpha}$ is the vector of errors in the observed data; and the notation

$$\delta\underline{\hat{X}}_n \equiv \delta\underline{\hat{X}}(t_n) = \delta\underline{X}(t_n) + \underline{\epsilon}_n$$

where $(\hat{})$ denotes computed and where $\underline{\epsilon}$ is the vector of errors in the computed state deviation vector.

Thus for an evenly determined set of data

$$\delta\underline{\hat{X}}_n = \varphi(t_n, t_o)B^{-1}\delta\underline{\tilde{A}} \tag{2.5}$$

and the error in computing $\underline{\delta X}_n$ is

$$\underline{\epsilon} = \Phi(t_n, t_o)B^{-1}\underline{\alpha}$$

This equation can be used to compute the covariance matrix of the estimation errors E from the covariance matrix of the measurement errors R. Adopting the notation

$$E \equiv \widetilde{\underline{\epsilon}\ \underline{\epsilon}^T}$$

to mean the expected value of the matrix $\underline{\epsilon}\ \underline{\epsilon}^T$, the notation

$$R \equiv \widetilde{\underline{\alpha}\underline{\alpha}^T}$$

to denote the errors in the observables and using the material of Appendix A allows the matrix of estimation errors to be written

$$E_m = \Phi(t_n, t_o)(B^T R^{-1} B)^{-1}\Phi^T(t_n, t_o) \tag{2.6}$$

since

$$(B^{-1})^T = (B^T)^{-1}$$

2.2.2 Data Filtering Techniques

In the introduction to this material, equations were derived for computing the position and velocity perturbations when precisely six navigation sightings were made. Further, equations were derived that related errors in the observations to errors in the computed position and velocity deviations for this case. On evaluation of these error equations, it is found that errors

in the observables of relatively small magnitudes can produce errors in the computed position perturbations that are completely unacceptable. The question then arises as to how additional sighting might be used to obtain a better estimate. Several methods of accomplishing this objective will be considered.

However, before considering this material it will be noted that any or all of the various estimation processes can be employed. The choice, should however, depend on the amount of information known about the errors in the observables. Thus, attempts will be made to demonstrate the accuracy (estimation error) of each approach and to explain the differences in precision obtained in terms of the assumptions made in deriving the estimator. In all cases, however, the assumption of second order statistical distributions is implicit (the discussions repeatedly employ Appendix A). Thus, information pertaining to higher moments is not employed and the "goodness" of the estimator should be suspect for non-Gaussian errors.

2.2.2.1 Least Squares Estimation

The method of least squares is perhaps the oldest and most easily understood of the general techniques for smoothing over-determined sets of data. For this reason, it will be considered first. The logic behind this filter is that the squares of the deviations in the observed minus computed observables from the estimated straight line (in m-dimensional space) defined by the equations

$$\delta \underline{A} = \mathcal{B}\, \delta \underline{X} \tag{2.7}$$

should be as small as possible. Alternatively, the moment of the deviations above the estimated line will equal the moment of the deviations below the line. This statement of the problem is equivalent to computing the line such that a comparison function equal to the summation of the squares of the differences in the observed and computed values of the values of δA is as small as possible, i.e.,

$$\frac{\partial L}{\partial \underline{X}} = \frac{\partial}{\partial X} \sum e_i^2 = 0 \qquad e_i \text{ are not defined.}$$

But, the sum of the squares of the measurement residuals L, can be written as

$$L = \overset{1xm\ mx1}{\underline{e}^T \underline{e}} = e_1^2 + e_2^2 + e_3^2 + \ldots + e_m^2 \tag{2.8}$$

or, by direct substitution, as

$$L = (\widetilde{\delta A} - B\,\delta\hat{X}_p)^T(\widetilde{\delta A} - B\,\delta\hat{X}_o)$$

Thus, the variation in this scalar comparison function can be related to a variation in the estimate as follows

$$\Delta L = -\Delta\,\delta\hat{X}_p^T B^T(\widetilde{\delta A} - B\,\delta\hat{X}_o) - (\widetilde{\delta A} - B\,\delta\hat{X}_o)^T B\,\Delta\,\delta\hat{X}_p \qquad (2.9)$$

$$\Delta L = -\left[(B^T\widetilde{\delta A} - B^TB\,\delta\hat{X}_o)^T\Delta\,\delta\hat{X}_o\right]^T - \left[(B^T\widetilde{\delta A} - B^TB\,\delta\hat{X}_p)^T\Delta\,\delta\hat{X}_o\right]$$

L will have an extreme value if $\Delta L = 0$ for any value of $\Delta\delta\underline{X}_o$. This will be the case if

$$(B^T\widetilde{\delta A} - B^TB\,\delta\hat{X}_o) = 0 \quad , \qquad (2.10)$$

or solving for $\delta\hat{X}_o$,

$$\delta\hat{X}_o = (B^TB)^{-1}B^T\widetilde{\delta A} \quad . \qquad (2.11)$$

The errors in the computed estimate of the state vector can now be related to errors in the measurement of $\widetilde{\delta A}$

$$\delta\hat{X}_o = \delta X_p + \epsilon \quad , \quad \widetilde{\delta A} = \delta A + \alpha$$

and

$$\delta A = B\,\delta X_o \quad .$$

Using Equation (2.11),

$$\epsilon = \delta\hat{X}_o - \delta X_o$$

$$= (B^TB)^{-1}B^T(\delta A + \alpha) - \delta X_o$$

$$= (B^TB)^{-1}B^T(B\,\delta X_o + \alpha) - \delta X_o$$

Then

$$\epsilon = (B^T B)^{-1} B^T \underline{\alpha} \quad . \tag{2.12}$$

Equation (2.12) can then be used to relate the covariance matrix of the estimation error to the covariance matrix of the measurement errors as follows:

$$E = \widetilde{\underline{\epsilon}\,\underline{\epsilon}^T} = \left[(B^T B)^{-1} B^T\right] \widetilde{\underline{\alpha}\,\underline{\alpha}^T} \left[(B^T B)^{-1} B^T\right]^T$$

Therefore,

$$\boxed{E = (B^T B)^{-1} B^T R B (B^T B)^{-1}} \tag{2.13a}$$

$$\boxed{\hat{\delta\underline{\chi}}_o = (B^T B)^{-1} B^T \widetilde{\delta\underline{A}}} \tag{2.13b}$$

2.2.2.2 Weighted Least Squares Estimation

The least squares estimate neglected information regarding the distribution of the measurement errors. Thus, if this information is known, a better estimate of the m-dimensional line utilized to fit the data can be obtained. This estimate is generated by modifying the comparison function in such a manner that moments for the errors which correspond to the higher quality observations are weighted higher. That is, the comparison function, L, of the previous analysis becomes

$$L = \frac{e_1^2}{\widetilde{\alpha_1 \alpha_1}} + \frac{e_2^2}{\widetilde{\alpha_2 \alpha_2}} + \frac{e_3^2}{\widetilde{\alpha_3 \alpha_3}} + \cdots + \frac{e_n^2}{\widetilde{\alpha_n \alpha_n}} \tag{2.14}$$

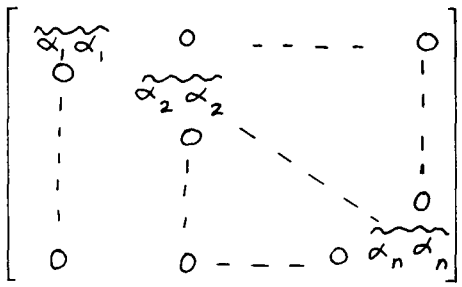
or

$$L = \underline{e}^T U^{-1} \underline{e} \tag{2.15}$$

where

$$\underline{e} = \underline{\delta\tilde{A}} - B\underline{\delta\hat{X}_o} \tag{2.16}$$

$$U = \tag{2.17}$$



In this equation the square of the measurement residual is weighted by a factor that is inversely porportional to the expected mean square value of the measurement error. Therefore, if the expected mean square measurement residual for a particular measurement is large, the contribution of the term to the comparison function will be small.

The weighted least squares (WLS) estimate of the state deviation vector can now be generated as in the case of the simple least squares problem. First, the comparison function is expressed in terms of the parameters of the problem

$$L = (\underline{\delta\tilde{A}} - B\underline{\delta\hat{X}_o})^T U^{-1}(\underline{\delta\tilde{A}} - B\underline{\delta\hat{X}_o}) \; . \tag{2.18}$$

Then the first variation of the comparison function, L, with respect to the estimate of the state vector is formed

$$\Delta L = -\Delta\underline{\delta\hat{X}_o}^T B^T U^{-1}(\underline{\delta\tilde{A}} - B\underline{\delta\hat{X}_o}) - (\underline{\delta\tilde{A}} - B\underline{\delta\hat{X}_o})^T U^{-1} B\Delta\underline{\delta\hat{X}_o} \; .$$

Again, this equation can be written as

$$\Delta L = -\left[(B^T U^{-1}\underline{\delta\tilde{A}} - B^T U^{-1} B\underline{\delta\hat{X}_o})^T \Delta\underline{\delta\hat{X}_o}\right]^T - \left[(B^T U^{-1}\underline{\delta\tilde{A}} - B^T U^{-1} B\underline{\delta\hat{X}_o})^T \Delta\underline{\delta\hat{X}_o}\right] \; .$$

since

$$(U^{-1})^T = (U^T)^{-1} = U^{-1} \; .$$

But, since both of the terms of this equation are scalars, the transpose of the bracketed term is equal to itself. This fact indicates that the loss function will have a stationary value when the estimate is chosen so that

$$B^T U^{-1} \widetilde{\delta A} - B^T U^{-1} B \, \hat{\delta X}_o = 0$$

Or, solving this equation for $\hat{\delta X}_o$, the WLS estimate becomes

$$\hat{\delta X}_o = (B^T U^{-1} B)^{-1} B^T U^{-1} \widetilde{\delta A} \tag{2.19}$$

The method of deriving the equations that are required to relate the measurement errors to the errors in the computed estimate is identical to that used for the least squares case. The estimation error is defined as

$$\epsilon = \hat{\delta X}_o - \delta X_o \; ;$$

therefore, from Equation 2.19

$$\epsilon = (B^T U^{-1} B)^{-1} B^T U^{-1} (B \, \delta X_o + \alpha) - \delta X_o \; , \tag{2.20}$$

However, since

$$\widetilde{\delta A} = (\delta A + \alpha) = (B \, \delta X_o + \alpha)$$

Equation (2.20) reduces to

$$\epsilon = (B^T U^{-1} B)^{-1} B^T U^{-1} \alpha$$

The covariance matrix of the estimation error is thus

$$E = \widetilde{\epsilon}\,\widetilde{\epsilon}^T = \left[(B^T U^{-1} B)^{-1} B^T U^{-1} \right] \widetilde{\alpha}\,\widetilde{\alpha}^T \left[(B^T U^{-1} B)^{-1} B^T U^{-1} \right]^T$$

If $\widetilde{\alpha}\,\widetilde{\alpha}^T$ is defined to be R, then these equations may be summarized as

$$E = \left[(B^T U^{-1} B)^{-1} B^T U^{-1} \right] R \left[(B^T U^{-1} B)^{-1} B^T U^{-1} \right]^T \tag{2.21}$$

$$\hat{\delta X} = (B^T U^{-1} B)^{-1} B^T U^{-1} \widetilde{\delta A} \tag{2.22}$$

Now, for the special case in which the measurement errors are uncorrelated; i.e.,

$$R = U$$

Then

$$E = \left[(B^T U^{-1} B)^{-1} B^T U^{-1}\right] U \left[(B^T U^{-1} B)^{-1} B^T U^{-1}\right]^T .$$

But, since U is a diagonal matrix,

$$U^T = U$$
$$(U^{-1})^T = U^{-1}$$

and the covariance matrix of the estimation error becomes

$$E = (B^T U^{-1} B)^{-1} B^T (U^{-1} U) U^{-1} B \left[(B^T U^{-1} B)^{-1}\right]^T$$

which reduces to

$$E = (B^T U^{-1} B)^{-1} \qquad (2.21a)$$

It is of further interst to note if the variances of the measurement errors are equal, then U can be written as

$$U = \sigma^2 I \quad \text{and} \quad U^{-1} = \frac{1}{\sigma^2} I$$

where σ^2 is the variance of each of the measurements. The estimate for $\underline{\delta X}$ then becomes

$$\hat{\underline{\delta X}}_0 = \left(\frac{1}{\sigma^2} B^T I B\right)^{-1} \frac{1}{\sigma^2} B^T I \, \widetilde{\underline{\delta A}}$$
$$\hat{\underline{\delta X}}_0 = (B^T B)^{-1} B^T \, \widetilde{\underline{\delta A}}$$

This equation indicates that for the case in which the variances of the measurement errors are equal the weighted least squares estimate will reduce to the least squares estimate.

2.2.2.3 Minimum Variance Estimation

In developing the estimation equations for both the least squares and the weighted least squares filters, a loss function was utilized which was simply related to the moments of the errors. The estimation equation was then formulated to minimize this loss function. In neither case was the statistical information pertaining to correlations in the components of the error vector ($\underline{\alpha}$) utilized. Thus, at this point a different approach to the problem will be formulated. To be specific, that estimate (defined as optimum) of the state deviation vector which is a linear function of the measurements, which minimizes each element of the estimation error covariance matrix and which corresponds to the constraint that the estimation error is not influenced by the quantity being estimated will be developed. That is, the form of the estimate is to be

$$\delta\hat{\underline{X}}_o = Q\,\delta\widetilde{\underline{A}} \; . \tag{2.23}$$

But,

$$\delta\hat{\underline{X}}_o = \delta\underline{X}_o + \underline{\epsilon}_o \; , \qquad \delta\widetilde{\underline{A}} = \delta\underline{A} + \underline{\alpha} \; , \text{ and } \quad \delta\underline{A} = B\,\delta\underline{X}_o$$

Thus, by direct substitution,

$$\delta\underline{X}_o + \underline{\epsilon}_o = Q\left[\delta\underline{A} + \underline{\alpha}\right] = Q\left[B\,\delta\underline{X}_o + \underline{\alpha}\right] \tag{2.24}$$

so that the error in the estimate is

$$\underline{\epsilon}_o = \left[QB - I\right]\delta\underline{X}_o + Q\,\underline{\alpha} \tag{2.25}$$

This equation indicates that if the error in the estimate is to be independent of the quantity being estimated, $\delta\underline{X}_o$, it is necessary that

$$QB - I = 0 \tag{2.26}$$

The estimation error then becomes

$$\underline{\epsilon}_o = Q\,\underline{\alpha}$$

and the estimation error covariance matrix becomes

$$\underline{\underline{\epsilon}}_o\,\underline{\underline{\epsilon}}_o^T = Q\,\underline{\underline{\alpha}}\,\underline{\underline{\alpha}}^T Q^T$$

or defining

$$\underline{\epsilon}_o\,\underline{\epsilon}_o^T \equiv E \quad \text{and} \quad \underline{\alpha}\,\underline{\alpha}^T \equiv R \tag{2.27}$$

$$E = QRQ^T$$

The problem now is to determine Q such that it will minimize E subject to the constraint that

$$QB - I = 0$$

This step can be accomplished by adjoining the constraint equation to E using matrix Lagrange multipliers as follows. Since

$$QB - I = 0 \quad , \quad B^T Q^T - I = 0$$

$$QB\lambda - \lambda = \lambda^T B^T Q^T - \lambda^T = 0 \tag{2.28}$$

where λ is an arbitrary n by n matrix multiplier that is not a function of Q. Thus

$$F = E + (QB - I)\lambda + \left[(QB - I)\lambda\right]^T$$

$$= QRQ^T + QB\lambda + \lambda^T B^T Q^T - \lambda - \lambda^T. \tag{2.29}$$

Now, recalling that $R^T \equiv R$, the first variation of $F\ (= \Delta F)$ becomes

$$\delta F = \left\{\delta Q\left[RQ^T + B\lambda\right]\right\} + \left\{\delta Q\left[RQ^T + B\lambda\right]\right\}^T$$

But the δQ are arbitrary. Therefore, each element of E will have an extreme value if

$$RQ^T + B\lambda = 0 \tag{2.30}$$

or

$$Q^T = -R^{-1}B\lambda \ . \tag{2.31}$$

Thus, premultiplying by B^T yields

$$B^T Q^T = -B^T R^{-1} B\lambda .$$

But,

$$B^T Q^T = I \ ;$$

thus,

$$\lambda = -(B^T R^{-1} B)^{-1} \tag{2.32}$$

so that

$$Q^T = R^{-1}B(B^T R^{-1}B)^{-1} , \tag{2.33a}$$

and

$$Q = (B^T R^{-1} B)^{-1} B^T R^{-1} \tag{2.33b}$$

since both R and $(B^T R^{-1} B)$ are symmetric. Thus, the minimum variance estimation equation is

$$\hat{\delta \underline{X}}_o = (B^T R^{-1} B)^{-1} B^T R^{-1} \widetilde{\delta \underline{A}} \tag{2.34}$$

It is interesting to note that when the errors in the measurements are uncorrelated (i.e., R = U), the minimum variance and weighted least squares estimates are identical.

Equations (2.27), (2.33a), and (2.33b) can be used to express the errors in the estimate as follows:

$$E = QRQ^T = (B^TR^{-1}B)^{-1}B^TR^{-1}B(B^TR^{-1}B)^{-1}$$

$$= (B^TR^{-1}B)^{-1}(BR^{-1}B)(B^TR^{-1}B)^{-1}$$

therefore, in summary

$$E = (B^TR^{-1}B)^{-1} \tag{2.35}$$

$$\delta\hat{\underline{X}}_o = (B^TR^{-1}B)^{-1}B^TR^{-1}\delta\underline{\tilde{A}} \tag{2.36}$$

In the development of these equations, Q was constrained to make E invariant with respect to the characteristics of $\delta\underline{X}_o$. Thus, these equations should be used to estimate $\delta\underline{X}_o$ when the statistical characteristics of $\delta\underline{X}_o$ are unknown. However, when the statistical characteristics of $\delta\underline{X}_o$ are known, that is, if the covariance matrix of $\delta\underline{X}_o$ is known, this constraint should be removed. When this is done, the resulting estimation equation can be determined as follows. From Equation (2.25),

$$\widetilde{\underline{\epsilon}_o\underline{\epsilon}_o^T} = \left[QB - I\right]\widetilde{\delta\underline{X}_o\delta\underline{X}_o^T}\left[QB - I\right]^T + Q\widetilde{\underline{\alpha}\underline{\alpha}^T}Q^T$$
$$+\left[QB - I\right]\widetilde{\delta\underline{X}_o\underline{\alpha}^T}Q^T + Q\widetilde{\underline{\alpha}\delta\underline{X}_o^T}\left[QB - I\right]^T$$

Now, defining

$$\widetilde{\delta\underline{X}_o\delta\underline{X}_o^T} \equiv V = V^T, \quad \widetilde{\delta\underline{X}_o\underline{\alpha}^T} = 0, \text{ and } \widetilde{\underline{\alpha}\delta\underline{X}_o^T} = 0$$

and rewriting the estimation error yields

$$E = Q\left[BVB^T + R\right]Q^T - QBV - VB^TQ^T + V \tag{2.37}$$

The variation in E that is produced by a variation in Q is thus

$$\Delta E = \left\{\delta Q\left[(BVB^T + R)Q^T - BV\right]\right\} + \left\{\delta Q\left[(BVB^T + R)Q^T - BV\right]\right\}^T$$

Again, if ΔE is to be zero for the case where δQ is arbitrary

$$(BVB^T + R)Q^T - BV = 0 \quad ; \tag{2.38}$$

that is, E will have an extreme value if

$$Q^T = (BVB^T + R)^{-1} BV \quad . \tag{2.39a}$$

Thus

$$Q = VB^T(BVB^T + R)^{-1} \tag{2.39b}$$

since $R = R^T$ and $V = V^T$.

The optimum estimate is now found by substitution into Equation (2.23)

$$\hat{\delta \underline{X}}_o = VB^T(BVB^T + R)^{-1} \widetilde{\delta \underline{A}} \tag{2.40}$$

The error in the estimate is determined using Equations (2.37) and (2.39b) as

$$E = VB^TQ^T - QBV - VB^TQ^T + V \quad .$$

Thus summarizing,

$$E = V - VB^T(BVB^T + R)^{-1} BV \tag{2.41a}$$

$$\hat{\delta \underline{X}}_o = VB^T(BVB^T + R)^{-1} \widetilde{\delta \underline{A}} \tag{2.41b}$$

Equations (2.41) can be put into a form that makes comparison with Equations (2.35)and (2.36) much simpler since from Equation (2.39)

$$Q = VB^T(BVB^T + R)^{-1}$$

$$= (B^TR^{-1}B + V^{-1})^{-1}(B^TR^{-1}B + V^{-1})VB^T(BVB^T + R)^{-1}$$

$$= (B^T R^{-1} B + V^{-1})^{-1} (B^T R^{-1} B V B^T + B^T R^{-1} R)(B V B^T + R)^{-1}$$

$$Q = (B^T R^{-1} B + V^{-1})^{-1} (B^T R^{-1})(B V B^T + R)(B V B^T + R)^{-1}.$$

Thus

$$Q = (B^T R^{-1} B + V^{-1})^{-1} B^T R^{-1}. \tag{2.42}$$

The estimation equation can then be written as

$$\delta \hat{\underline{X}}_o = (B^T R^{-1} B + V^{-1})^{-1} B^T R^{-1} \delta \underline{\tilde{A}} \tag{2.43}$$

Using Equation (2.41a) and (2.42), the error in the estimate can then be written in the form

$$E = V - Q B V$$

$$= V - (B^T R^{-1} B + V^{-1})^{-1} B^T R^{-1} B V$$

$$= (B^T R^{-1} B + V^{-1})^{-1} (B^T R^{-1} B + V^{-1}) V - (B^T R^{-1} B + V^{-1})^{-1} B^T R^{-1} B V$$

$$E = (B^T R^{-1} B + V^{-1})^{-1} \left[B^T R^{-1} B V + I - B^T R^{-1} B V \right]$$

thus

$$E = (B^T R^{-1} B + V^{-1})^{-1} \tag{2.44a}$$

$$\delta \hat{\underline{X}}_o = (B^T R^{-1} B + V^{-1})^{-1} B^T R^{-1} \delta \underline{\tilde{A}} \tag{2.44b}$$

While Equations (2.41) and (2.44) differ considerably in form, they are equivalent and will yield the same results. Notice that Equations (2.44) differ from the corresponding Equations (2.35 and (2.36) only through the presence of the additive term V^{-1}; i.e., if $V^{-1} = 0$, the equations become identical.

Before leaving this discussion, it is worthy of note to demonstrate that the process employed in this technique to derive an optimum estimate of $\delta \underline{X}$ (i.e., the minimization of a matrix) is equivalent to one which a scalar loss function is constructed. One such loss function could be the summation of the eigen-values of the covariance matrix (see Appendix C).

Consider the scalar

$$\mathcal{L} = A^T E A$$

where A is an arbitrary vector whose dimensionality is n and where E is the n by n matrix of estimation errors. Since A is arbitrary, it can be independent of the parameters of interest so that

$$\Delta \mathcal{L} = A^T \Delta E A$$

Thus, if $\Delta \mathcal{L} = 0$, ΔE must also be zero provided E is free of constraints. Thus, a sufficient condition for any scalar measure of a matrix to be minimum is for $\Delta E = 0$.

It is also worthy to note that no explicit assumption has been made regarding the distribution of the errors. True, only second order statistics are utilized so that the estimate will not be optimum in a larger sense (use all of the information available) unless the errors are Gaussian. However, this estimate can be generated. A minimum variance unbiased measure of the performance degradation will be discussed in Section 2.3.

Finally, it is noted that the minimum variance estimate generated in this manner is unbiased since the conditional expectation of the estimate, $\delta \hat{\underline{X}}_o$, is $\delta \underline{X}_o$.

2.2.2.4 Iterative Form of the Minimum Variance Estimator

The equation that was derived in the previous section for computing the mv estimate was

$$\delta \hat{\underline{X}}_o = (B^T R^{-1} B)^{-1} B^T R^{-1} \widetilde{\delta \underline{A}} \tag{2.45}$$

This equation is useful when all of the measurements are to be processed at one time. Quite often, however, it is desirable to process the data that is currently available to formulate an initial estimate, and then to compute new estimates as additional data becomes available. This desirability arises from several distinct factors. First, the numerical operations themselves would be considerably simplified if only the most recent observation was being processed. (The amount of data can become staggeringly large.) Second, the trajectory is, in fact, nonlinear so that errors of assumed linearity in the transition matrix and in the observation problem combine to make translation to the fixed reference epoch very inaccurate as the time from this epoch becomes large (this fault can be avoided if the reference trajectory is re-defined by adding the reference position and velocity to the computed deviations and restarting the estimation process). For these reasons, an iterative (or repetitive) form of the minimum variance estimator will be developed.

The matrices contained in Equation (2.3.5) can be partitioned into sub-matrices as follows

$$B = \begin{bmatrix} \underline{B}_1 \\ \underline{B}_2 \\ \underline{B}_3 \\ \vdots \\ \underline{B}_m \end{bmatrix} , \qquad \widetilde{\delta \underline{A}} = \begin{bmatrix} \widetilde{\delta \underline{A}}_1 \\ \widetilde{\delta \underline{A}}_2 \\ \widetilde{\delta \underline{A}}_3 \\ \vdots \\ \widetilde{\delta \underline{A}}_m \end{bmatrix}$$

$$B^T = \left[B_1^T \mid B_2^T \mid B_3^T \mid \cdots \mid B_m^T \right] ,$$

$$R = \left[\begin{array}{c|c|c|c|c} R_1 & 0 & 0 & \cdots & 0 \\ \hline 0 & R_2 & 0 & \cdots & 0 \\ \hline 0 & 0 & R_3 & \cdots & 0 \\ \hline \vdots & \vdots & \vdots & & \\ \hline 0 & 0 & 0 & & R_m \end{array}\right] .$$

where it is noted that the subscripts in these equations now refer to sets of quantities. Thus, δA_1 , refers to the first set of measurements, and $\delta \underline{A}_n$ means the nth set of measurements. This notation contrasts with previous usage (Equations (2.2) and (2.3)) where the subscripts referred to individual measurements. It is, of course, understood that the matrices B and R are partitioned so that the sub-matrices are conformable. This division guarantees that the required multiplications can be performed. A further assumption has been made concerning the R matrix in that the different sets of measurements are assumed to be uncorrelated. Correlation between the individual measurements of any measurement set is, of course, permitted. Under these assumptions, the inverse of R can be written as

$$R^{-1} = \begin{bmatrix} R_1^{-1} & 0 & 0 & \cdots & 0 \\ 0 & R_2^{-1} & 0 & \cdots & 0 \\ 0 & 0 & R_3^{-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \\ 0 & 0 & 0 & & R_n^{-1} \end{bmatrix}$$

as can be seen by considering the relation

$$RR^{-1} = I$$

Now, the matrix product $B^T R^{-1}$ is

$$B^T R^{-1} = \left[B_1^T R_1^{-1} \mid B_2^T R_2^{-1} \mid B_3^T R_3^{-1} \mid \cdots \mid B_n^T R_n^{-1} \right] ,$$

and the product $B^T R^{-1} B$ will then be

$$B^T R^{-1} B = B_1^T R_1^{-1} B_1 + B_2^T R_2^{-1} B_2 + \cdots + B_n^T R_n^{-1} B_n \quad .$$

or

$$B^T R^{-1} B = \sum_{i=1}^{n} B_i^T R_i^{-1} B_i \quad . \tag{2.46}$$

The matrix product $B^T R^{-1} B$ becomes a function of the number of submatrices contained in the product. Therefore, let this product be defined as follows:

$$J_n \equiv B^T R^{-1} B = \sum_{i=1}^{n} B_i^T R_i^{-1} B_i \quad . \tag{2.47}$$

The matrix J_n can be written as

$$J_n = \sum_{i=1}^{n-1} B_i^T R_i^{-1} B_i + B_n^T R_n^{-1} B_n \tag{2.48}$$

and in iterative form as

$$J_n = J_{n-1} + B_n^T R_n^{-1} B_n \quad . \tag{2.49}$$

Similarly, the matrix $B^T R^{-1} \delta\widetilde{A}$ can be written as

$$B^T R^{-1} \delta\widetilde{A} = B_1^T R_1^{-1} \delta\widetilde{A}_1 + B_2^T R_2^{-1} \delta\widetilde{A}_2 + \ldots + B_n^T R_n^{-1} \delta\widetilde{A}_n \tag{2.50}$$

Now defining

$$\xi_n \equiv B^T R^{-1} \delta\widetilde{A} = \sum_{i=1}^{n} B_i^T R_i^{-1} \delta\widetilde{A}_i \tag{2.51}$$

the iterative form of ξ_n will be,

$$\xi_n = \xi_{n-1} + B_n^T R_n^{-1} \delta\widetilde{A}_n \quad . \tag{2.52}$$

Equations (2.45), (2.49), and (2.51) can be used to express the estimate of the state vector that is obtained by processing all of the data up to and including the nth set as follows:

$$\delta\hat{X}_o^{(n)} = J_n^{-1} \xi_n \quad , \tag{2.53}$$

and the estimate obtained from processing all data up to and including the (n-1) set is

$$\delta\hat{X}_o^{(n-1)} = J_{n-1}^{-1} \xi_{n-1} \tag{2.54}$$

The superscripts have been added to $\delta\hat{X}_o$ to indicate the quantity of measurement data used in making the estimate. But, substitution of Equations (2.49) and (2.52) into Equation (2.53) yields the estimate of $\delta\hat{X}_o^n$ as

$$\delta\hat{X}_o^{(n)} = (J_{n-1} + B_n^T R_n^{-1} B_n)^{-1} (\xi_{n-1} + B_n^T R_n^{-1} \delta\widetilde{A}_n) \tag{2.55}$$

from which

$$(J_{n-1} + B_n^T R_n^{-1} B_n)\,\delta\hat{\underline{X}}_o^{(n)} = \underline{\xi}_{n-1} + B_n^T R_n^{-1}\,\delta\widetilde{\underline{A}}_n \tag{2.56}$$

But, from Equation (2.54),

$$\underline{\xi}_{n-1} = J_{n-1}\,\delta\hat{\underline{X}}_o^{(n-1)}. \tag{2.57}$$

Thus, Equation (2.56) can be written as

$$\begin{aligned}(J_{n-1} + B_n^T R_n^{-1} B_n)\,\delta\hat{\underline{X}}_o^{(n)} &= J_{n-1}\,\delta\hat{\underline{X}}_o^{(n-1)} + B_n^T R_n^{-1}\,\delta\widetilde{\underline{A}}_n \\ &= J_{n-1}\,\delta\hat{\underline{X}}_o^{(n-1)} + B_n^T R_n^{-1} B_n\,\delta\hat{\underline{X}}_o^{(n-1)} \\ &\quad - B_n^T R_n^{-1} B_n\,\delta\hat{\underline{X}}_o^{(n-1)} + B_n^T R_n^{-1}\,\delta\widetilde{\underline{A}}_n\end{aligned} \tag{2.53}$$

which reduces to

$$\begin{aligned}(J_{n-1} + B_n^T R_n^{-1} B_n)\,\delta\hat{\underline{X}}_o^{(n)} &= (J_{n-1} + B_n^T R_n^{-1} B_n)\,\delta\hat{\underline{X}}_o^{(n-1)} \\ &\quad + B_n^T R_n^{-1}(\delta\widetilde{\underline{A}}_n - B_n\,\delta\hat{\underline{X}}_o^{(n-1)}).\end{aligned} \tag{2.59}$$

Finally, multiplying both sides of Equation (2.59) by the inverse of $(J_{n-1} + B_n^T R_n^{-1} B_n)$ yields

$$\delta\hat{\underline{X}}_o^{(n)} = \delta\hat{\underline{X}}_o^{(n-1)} + (J_{n-1} + B_n^T R_n^{-1} B_n)^{-1} B_n^T R_n^{-1}(\delta\widetilde{\underline{A}}_n - B_n\,\delta\hat{\underline{X}}_o^{(n-1)}), \tag{2.60}$$

where from Equation (2.49)

$$J_n = J_{n-1} + B_n^T R_n^{-1} B_n \tag{2.61}$$

Equations (2.60) and (2.61) are the iterative equations that are required to compute the minimum variance estimate.

In a similar manner, the recursive form of the covariance matrix for the estimation errors can be developed. Consider the non-recursive form

$$E = (B^T R^{-1} B)^{-1}$$

Comparison of this equation with Equation (2.47) shows that the covariance matrix of the error in the nth estimate is related to J_n as follows: $E_n = J_n^{-1}$ Thus, this substitution into Equations (2.60) and (2.61) yields:

$$\delta\hat{\underline{X}}_o^{(n)} = \delta\hat{\underline{X}}_o^{(n-1)} + (E_{n-1}^{-1} + B_n^T R_n^{-1} B_n)^{-1} B_n^T R_n^{-1} (\delta\tilde{A}_n - B_n \delta\hat{\underline{X}}_o^{(n-1)}),$$

$$E_n = (E_{n-1}^{-1} + B_n^T R_n^{-1} B_n)^{-1}$$

where: $\underline{\delta\hat{\underline{X}}_o^{(0)} \equiv \text{arbitrary}}$ AND $\underline{E_o^{-1} \equiv 0}$

At this point it is interesting to note that the conditional expectation of $\delta\hat{\underline{X}}_o^{(n)}$ is biased by the memory of all previous estimates. Further, if the time intervals between data points are approximately equal, and if the errors in the observables are comparable, the bias will increase as the number of data points increases. This fact may appear to be the result of an error, since the form of this estimate is a direct consequence of Equation (2.44). However, it is noted in the way of an explanation that the solution is biased only in a local sense. The result of the complete reduction problem will still be unbiased, since the initial conditions for the problem were unbiased; i.e., the matrix E_o^{-1} was the null matrix rather than some initial estimate of this quantity.

Now noting that

$$\delta\hat{\underline{X}}_o^{(n)} = \varphi(t_o, t_n)\, \delta\hat{\underline{X}}_n^{(n)}$$

$$\delta\hat{\underline{X}}_o^{(n-1)} = \varphi(t_o, t_{n-1})\, \delta\hat{\underline{X}}_{n-1}^{(n-1)}$$

$$\varphi(t_n, t_{n-1}) = \varphi^{-1}(t_o, t_n)\, \varphi(t_o, t_{n-1})$$

$$B_n = H_n\, \varphi(t_n, t_o)$$

allows the first of these equations (2.60) to be rewritten as

$$\delta\hat{\underline{X}}_n^{(n)} = \varphi(t_n, t_{n-1})\, \delta\hat{\underline{X}}_{n-1}^{(n-1)} + \varphi(t_n, t_o)\, J_n^{-1} B_n^T R_n^{-1} (\delta\tilde{A}_n - H_n\, \varphi(t_n, t_{n-1})\, \delta\hat{\underline{X}}_{n-1}^{(n-1)})$$

But

$$\varphi(t_n,t_o)\,J_n^{-1}B_n^T = \left[\varphi(t_n,t_o)\,J_n^{-1}\varphi^T(t_n,t_o)\right]H_n^T$$

$$\equiv J_n^{-1}(t_n)\,H_n^T$$

$$= \left[\varphi(t_n,t_{n-1})\,J_{n-1}(t_{n-1})\,\varphi(t_n,t_{n-1}) + H_n^T R_n^{-1} H_n\right]^{-1} H_n^T$$

and the estimator reduces to

$$\delta\hat{\underline{X}}_n^{(n)} = \varphi(t_n,t_{n-1})\,\delta\hat{\underline{X}}_{n-1}^{(n-1)} + \left[\varphi(t_n,t_{n-1})E_{n-1}^{-1}(t_{n-1})\,\varphi^T(t_n,t_{n-1})\right.$$

$$\left. + H_n^T R_n^{-1} H_n\right]^{-1} H_n^T R_n^{-1}\left(\delta\tilde{\underline{A}}_n - H_n\,\varphi(t_n,t_{n-1})\,\delta\hat{\underline{X}}_{n-1}^{(n-1)}\right)$$

$$= \delta\hat{X}_{n-1}^{(n)} + Q_n\left[\delta\tilde{A}_n - H_n\,\delta\hat{X}_{n-1}^{(n)}\right]$$

where $\delta\hat{X}_{n-1}^{(n)}$ = estimate of $\delta\underline{X}$ at the epoch t_n based on all information processed through t_{n-1} (n-1 sets)

$$Q_n \equiv \left[J_{n-1}^{(n)} + H_n^T R_n^{-1} H_n\right]^{-1} H_n^T R_n^{-1}$$

$J_{n-1}^{(n)}$ = estimate of J at the epoch t_n based on information processed through t_{n-1}

$$= \varphi(t_n,t_{n-1})\,J_{n-1}^{(n-1)}\,\varphi^T(t_n,t_{n-1})$$

The second equation for estimation error then becomes

$$J_n^{(o)} = H_n^T R_n^{-1} H_n + J_{n-1}^{(o)}$$

$$\varphi(t_0,t_n)J_n^{(n)}\varphi^T(t_0,t_n) = \varphi^T(t_n,t_0)H_n^T R_n^{-1} H_n \varphi(t_n,t_0) + \varphi(t_0,t_{n-1})J_{n-1}^{(0)}\varphi^T(t_0,t_{n-1})$$

or

$$J_n^{(n)} = H_n^T R_n^{-1} H_n + \varphi(t_n,t_{n-1})J_{n-1}^{(n-1)}\varphi^T(t_n,t_{n-1}) \equiv H_n R_n^{-1} H_n + J_{n-1}^{(n)}$$

Thus, in summary

$$J_{n-1}^{(n)} = \varphi(t_n,t_{n-1})\, J_{n-1}^{(n-1)}\varphi^T(t_n,t_{n-1})$$

$$Q_n = \left[J_{n-1}^{(n)} + H_n^T R_n^{-1} H_n\right]^{-1} H_n^T R_n^{-1}$$

$$\delta\hat{\underline{X}}_n^{(n)} = \varphi(t_n,t_{n-1})\,\delta\hat{\underline{X}}_{n-1}^{(n-1)} + Q_n\left[\delta\widetilde{\underline{A}}_n - H_n\varphi(t_n,t_{n-1})\,\delta\hat{\underline{X}}_{n-1}\right]$$

$$J_n^{(n)} = H_n^T R_n^{-1} H_n + J_{n-1}^{(n)}$$

$$J_0^{(0)} \equiv 0 \quad ; \quad \delta\underline{X}_0 \equiv \text{arbitrary}$$

This set of equations allows an estimate of the state to be generated for the epoch t_n from a priori estimate of the state at this epoch $[\varphi(t_n,t_{n-1})\delta\underline{X}_{n-1}^{(n-1)}]$ and the observed minus computed residuals available at t_n. Further, since the initial conditions for $J_0^{(0)}$ and $\delta\underline{X}_0$ are specified, the process can be initiated at any time. However, the question exists as to how information which might be available at t_0 for $J_0^{(0)}$ could be utilized. It might be argued that such a process is simply a continuation of some previous analysis and that the initial conditions could be substituted directly. However, this argument is not satisfying; and, therefore, a more rigorous proof will be constructed.

2.2.2.5 Schmidt-Kalman Filter Via Minimum Variance

The development of the Schmidt-Kalman (Reference 0.1) estimation equations is very similar to the development of the minimum variance equations, but differs in the respect that an initial estimate of $\underline{\delta X}_0$, $\underline{\delta X}_0'$ is assumed to be available for the purpose of biasing the estimator in a total sense toward a'priori estimate. The optimum estimate is thus assumed to be a linear function of both the a'priori estimate $\underline{\delta X}_0'$ and the measurements $\underline{\delta A}$ and is formulated to minimize the elements of the estimation error covariance matrix subject to the constraint that the estimation error is not a function of $\underline{\delta X}_0$.

The derivation of the filter equations will then require the following definitions:

$\underline{\delta X}_0'$ = a'priori estimate of $\underline{\delta X}_0$

$\underline{e}_0'$ = error in a'priori estimate of $\underline{\delta X}_0$

E' = $\underset{\sim}{\underline{e}_0' \; \underline{e}_0'^T}$ = covariance matrix of a'priori estimation errors

And, as was stated, the form of the optimum estimate is assumed to be

$$\underline{\hat{\delta X}}_0 = P\underline{\hat{\delta X}}_0' + Q\,\underline{\tilde{\delta A}} \tag{2.64}$$

where P and Q are chosen to fullfill the conditions discussed previously. But,

$$\underline{\hat{\delta X}}_0 \equiv \underline{\delta X}_0 + \underline{e}_0 \;,\quad \underline{\hat{\delta X}}_0' \equiv \underline{\delta X}_0 + \underline{e}_0' \;,\quad \underline{\tilde{\delta A}} \equiv \underline{\delta A} + \underline{\alpha} \;,$$
$$\text{and} \quad \underline{\delta A} \equiv B\,\underline{\delta X}_0 \;,\quad \text{so that} \quad \underline{\tilde{\delta A}} = B\,\underline{\delta X}_0 + \underline{\alpha}$$

Thus, substitution of these definitions into equation (2.64) yields

$$\underline{\delta X}_0 + \underline{e}_0 = P\underline{\delta X}_0 + P\underline{e}_0' + QB\,\underline{\delta X}_0 + Q\underline{\alpha} \;,$$

and the error in the estimate is obtained as

$$\underline{e}_0 = (P + QB - I)\,\underline{\delta X}_0 + P\underline{e}_0' + Q\underline{\alpha} \;. \tag{2.65}$$

It can be seen from equation (2.65) that if the error in the estimate is to be independent of $\underline{\delta X}$, then

$$(P + QB - I) = 0 \tag{2.66}$$

This constraint allows P to be determined as a function of Q

$$P = I - QB \tag{2.67}$$

and allows the form of the estimation equation (2.64) to be written as

$$\underline{\delta \hat{X}}_o = \underline{\delta \hat{X}}'_o + Q(\underline{\delta \tilde{A}} - B\,\underline{\delta \hat{X}}'_o) \tag{2.68}$$

Therefore, it remains to select Q so as to minimize the elements of the covariance matrix of $\underline{e}_o$. This task can be accomplished using equations (2.65) and (2.67), as follows:

$$\begin{aligned} \underline{e}_o &= (I - QB)\,\underline{e}'_o + Q\,\underline{\alpha} \\ \underline{e}_o \underline{e}_o^T &= (I-QB)\,\underline{e}'_o \underline{e}'^T_o (I-QB)^T + Q\,\underline{\alpha}\underline{\alpha}^T Q^T \\ &\quad + (I-QB)\,\underline{e}'_o \underline{\alpha}^T Q^T + Q\,\underline{\alpha}\,\underline{e}'^T_o (I-QB)^T . \end{aligned} \tag{2.69}$$

Now, defining

$\underline{\alpha}\underline{\alpha}^T = R$

$\underline{e}'_o \underline{\alpha}^T \equiv U$ then $\underline{\alpha}\,\underline{e}'^T_o = (\underline{e}'_o \underline{\alpha}^T)^T = (\underline{e}'_o \underline{\alpha}^T)^T = U^T$

allows the covariance of the estimation error to be written as

$$\begin{aligned} E &= Q\left[BE'B^T + R - BU - U^T B^T\right] Q^T \\ &\quad - Q\left[BE' - U^T\right] - \left[E'B^T - U\right] Q^T + E' \end{aligned} \tag{2.70}$$

The first variation of E with respect to Q is now required as the first step in obtaining the relationship for the optimum linear estimate.

$$\begin{aligned} \delta E &= \left\{\delta Q\left[\left(BE'B^T + R - BU - U^T B^T\right) Q^T - \left(BE' - U^T\right)\right]\right\} \\ &\quad + \left\{\delta Q\left[BE'B^T + R - BU - U^T B^T\right) Q^T - \left(BE' - U^T\right)\right]\right\}^T \end{aligned}$$

But, δQ is arbitrary; thus, the elements of E will have an extreme value if

$$\left(BE'B^T + R - BU - U^T B^T\right) Q^T - \left(BE' - U^T\right) = 0$$

or, solving for Q

$$Q = \left[E'B^T - U\right]\left[BE'B^T + R - BU - U^T B^T\right]^{-1} \tag{2.71}$$

In most applications, the measurement errors and the errors in the a'priori estimate of the state are uncorrelated

$$U = 0 \quad \text{and} \quad U^T = 0$$

Thus, Q becomes

$$Q = E' B^T \left[B E' B^T + R \right]^{-1} \tag{2.72}$$

The covariance matrix of the estimation error for this case can be determined from equations (2.70) and (2.72) as

$$\begin{aligned} E &= E' + Q\left[BE'B^T + R\right]Q^T - QBE' - E'B^TQ^T \\ &= E' + E'B^T\left[BE'B^T + R\right]^{-1}\left[BE'B^T + R\right]Q^T - QBE' - E'B^TQ^T \\ E &= E' - QBE' \quad . \end{aligned} \tag{2.73}$$

Therefore,

$$E = E' - E'B^T\left[BE'B^T + R\right]^{-1}BE' \tag{2.74}$$

where

$$\underline{\hat{\delta X}}_o = \underline{\hat{\delta X}}'_o + Q\left[\underline{\tilde{\delta A}} - B\,\underline{\hat{\delta X}}'_o\right]$$

$$Q = E'B^T\left[BE'B^T + R\right]^{-1}$$

These filter equations, (2.74), can be written in a different form, as follows:

$$\begin{aligned} Q &= E'B^T\left[BE'B^T + R\right]^{-1} \\ &= \left(B^TR^{-1}B + E'^{-1}\right)^{-1}\left(B^TR^{-1}B + E'^{-1}\right)E'B^T\left(BE'B^T + R\right)^{-1} \\ &= \left(B^TR^{-1}B + E'^{-1}\right)^{-1}\left(B^TR^{-1}BE'B^T + B^TR^{-1}R\right)\left(BE'B^T + R\right)^{-1} \\ &= \left(B^TR^{-1}B + E'^{-1}\right)^{-1}B^TR^{-1}\left(BE'B^T + R\right)\left(BE'B^T + R\right)^{-1} \end{aligned}$$

or

$$Q = \left(B^TR^{-1}B + E'^{-1}\right)^{-1}B^TR^{-1} \,, \tag{2.75}$$

Thus, the equation for E can be rewritten using equations (2.73) and (2.75) as

$$E = E' - QBE'$$

$$= \left[(B^T R^{-1} B + E'^{-1})^{-1}(B^T R^{-1} B + E'^{-1})\right] E' - \left[(B^T R^{-1} B + E'^{-1}) B^T R^{-1}\right] BE'$$

$$= (B^T R^{-1} B + E'^{-1})^{-1} \left[B^T R^{-1} BE' + I - B^T R^{-1} BE'\right]$$

or

$$E = (B^T R^{-1} B + E'^{-1})^{-1} \tag{2.76}$$

Thus, using equations (2.75) and (2.76), the set of equations analogous to (2.74) can be written as

$$\boxed{\begin{aligned} E &= (B^T R^{-1} B + E'^{-1})^{-1} \\ \underline{\delta \hat{X}_o} &= \underline{\delta \hat{X}'_o} + Q\left[\underline{\delta \tilde{A}} - B\,\underline{\delta \hat{X}'_o}\right] \\ Q &= (B^T R^{-1} B + E'^{-1})^{-1} B^T R^{-1} \end{aligned}} \tag{2.77}$$

Note that (2.77) requires the inversion of a matrix of the dimensionality of the state vector as opposed to the dimensionality of the observation vector, as was the case with equation (2.74). Also note that the estimation error for this case does not involve subtraction. The first of these differences is a definite disadvantage due to the fact that there is an increased chance for numerical error due to loss of significance when inverting. However, the second difference is an advantage since it avoids the problems associated with assuring positive eigen-values which might result as E approaches the null matrix in the other formulation.

Also, note that this form of the estimation equation is exactly the same as that obtained by transforming the minimum variance estimator to the recursive mode. Improvement in the estimate can, however, be expected, since provision has been made to begin the process with values of $J_o^{(o)}$ and $\underline{\delta X}_o$, other than those employed in the minimum variance case.

Finally, note that the estimators (2.74) and (2.77) are both biased. However, in contract to the recursive minimum variance estimates, this bias exists on the overall and local senses. This fact graphically displays the effect of a'priori information in the data reduction problem since the solution is weighted in the direction of the available data.

When the statistical properties of $\underline{\delta X}_o$ are known; i.e.,

$$\widetilde{\underline{\delta X}_o\,\underline{\delta X}_o^T} = V \qquad \widetilde{\underline{\delta X}_o\,\underline{e}_o'^T} = W \qquad \widetilde{\underline{e}_o'\,\underline{\delta X}_o^T} = W^T$$

$$\widetilde{\underline{\delta X}_o\,\underline{\alpha}^T} = S \qquad \widetilde{\underline{\alpha}\,\underline{\delta X}_o^T} = S^T \text{ and } \widetilde{\underline{e}'\,\underline{\alpha}^T} = U \qquad \widetilde{\underline{\alpha}\,\underline{e}_o'^T} = U^T, \tag{2.78}$$

this additional information can be used to improve the estimate of $\underline{\delta X_0}$. The equations required to utilize this information can be obtained by simply removing the constraint equation (2.66). When this step is performed, the covariance matrix of the estimation error will be found from equation (2.65) to be

$$\begin{aligned} E = {} & (P+QB-I)V(P+QB-I)^T + PE'P^T + QRQ^T \\ & + \left\{(P+QB-I)WP^T + (P+QB-I)SQ^T + PNQ^T\right\} \\ & + \left\{(P+QB-I)WP^T + PNQ^T + (P+QB-I)SQ^T\right\}^T \end{aligned} \tag{2.79}$$

Equation (2.79) can be written in the form

$$E = PAB^T + PC^TQ^T - PF - F^TP^T + QGQ^T + QCP^T - QD - D^TQ^T + V \tag{2.80}$$

where

$$\begin{aligned} A &= V + E' + W + W^T \\ G &= BVB^T + R + BS + (BS)^T \\ C &= BV + BW + S^T + U^T \\ D &= BV + S^T \\ F &= V + W^T \end{aligned} \tag{2.81}$$

The variation of E with respect to both P and Q can be written, using equation (2.80), as follows:

$$\begin{aligned} \delta E = {} & \left\{\delta P\left[AP^T + C^TQ^T - F\right]\right\} + \left\{\delta P\left[AP^T + C^TQ^T - F\right]\right\}^T \\ & + \left\{\delta Q\left[GQ^T + CP^T - D\right]\right\} + \left\{\delta Q\left[GQ^T + CP^T - D\right]\right\}^T \end{aligned} \tag{2.82}$$

But, since δP and δQ are arbitrary, equation (2.82) indicates that δE will be zero if

$$AP^T + C^TQ^T = F \tag{2.83}$$

$$CP^T + GQ^T = D \tag{2.84}$$

Equation (2.83) can be multiplied by A^{-1} to yield

$$P^T + A^{-1}C^TQ^T = A^{-1}F \tag{2.85}$$

and this result can be rewritten as

$$CP^T + CA^{-1}C^TQ^T = CA^{-1}F \tag{2.86}$$

Now, subtracting equation (2.86) from (2.84) yields

$$\left[G - CA^{-1}C^T\right]Q^T = D - CA^{-1}F$$

Therefore,

$$Q^T = \left[G - CA^{-1}C^T\right]^{-1}\left[D - CA^{-1}F\right] \tag{2.87}$$

or

$$Q = \left[D - CA^{-1}F\right]^T\left[G - CA^{-1}C^T\right]^{-1} \tag{2.88}$$

since $\left[G - CA^{-1}C^T\right]^T = \left[G^T - C(A^T)^{-1}C^T\right] = \left[G - CA^{-1}C^T\right]$.

Equation (2.83) now implies that

$$P^T = A^{-1}\left[F - C^TQ^T\right] .$$

Thus,

$$P = \left[F^T - QC\right]A^{-1} ; \qquad since\ A = A^T . \tag{2.89}$$

In summary, the equations required to formulate the estimate are

$$\underline{\delta\hat{X}}_o = P\,\underline{\delta\hat{X}}_o' + Q\,\underline{\widetilde{\delta A}}$$

where

$$\begin{aligned} P &= \left[F^T - QC\right]A^{-1} \\ Q &= \left[D - CA^{-1}F\right]^T\left[G - CA^{-1}C^T\right]^{-1} \\ A &= V + E' + W + W^T \\ G &= BVB^T + R + BS + (BS)^T \\ C &= BV + BW + S^T + U^T \\ D &= BV + S^T \\ F &= V + W^T \end{aligned} \tag{2.90}$$

$$V = \underset{\sim}{\underline{\delta X_o\,\delta X_o^T}} \qquad R = \underset{\sim}{\underline{\alpha\alpha^T}} \qquad W = \underset{\sim}{\underline{\delta X_o\,e_o'^T}} \qquad S = \underset{\sim}{\underline{\delta X_o\,\alpha^T}} \qquad U = \underset{\sim}{\underline{e_o'\alpha^T}}$$

A number of alternate forms for P and Q exist and should be investigated to determine the form most suitable for a particular application. One such variation can be obtained from equation (2.84) by post-multiplying by $C^T G^{-1}$,

$$C^T G^{-1} C P^T + C^T Q^T = C^T G^{-1} D \tag{2.91}$$

and subtracting from (2.83). This process yields

$$\left[A - C^T G^{-1} C\right] P^T = F - C^T G^{-1} D$$

from which

$$P^T = \left[A - C^T G^{-1} C\right]^{-1} \left[F - C^T G^{-1} D\right] , \tag{2.92}$$

or

$$P = \left[F - C^T G^{-1} D\right]^T \left[A - C^T G^{-1} C\right]^{-1} , \tag{2.93}$$

since $A = A^T$ and $G = G^T$.

The corresponding expression for Q can be obtained using equations (2.84) as follows

$$Q^T = G^{-1}\left[D - C P^T\right]$$

Thus,

$$Q = \left[D^T - D C^T\right] G^{-1}$$

This derivation of the Kalman estimator employed a minimum variance concept to arrive at the optimum estimate. If the statistics are Russian, this procedure will yield the optimum estimate in a larger sense (see Section 2.3) since the higher moments are zero.

However, if the statistics are non-Gaussian, the resultant estimator (a biased minimum variance estimator) will not be optimum (again in the larger sense) since it neglects all knowledge of any higher order moments in the distributions of the errors. Thus future discussions will provide a reformulation of this problem from the standpoint of a much more general concept of loss and optimum estimate selection. This discussion (presented in Section 2.3) will develop the specific case for Gaussian errors and will show that the resultant estimation is, in general, superior to any other which can be formulated.

2.3 STATISTICAL ESTIMATION THEORY

2.3.1 Introduction

The discussions presented in section 2.2 lead to the simple development of a series of computational algorithms which defined an estimate of the state deviation vector in terms of a series of observations and the initial conditions. However, implicit in this material were the assumptions that

1) the dynamical model was linear

2) the observation model was linear

3) the optimum estimate of the state deviation was a linear function of the observed minus computed values of the observables

4) only second-order statistics were necessary.

Further, the "loss" functions employed to develop optimum estimates of the state deviation vector, while similar, were intuitive, thus giving rise to questions regarding the uniqueness of the estimates generated. For these reasons, it is now desirable to re-examine the estimation problem to demonstrate the manner in which these assumptions can be relaxed and to show that all of these estimators are special cases of a more general family of estimators. Specific attention will be focused on:

1) the criteria to be utilized in determining the optimality of the estimate

2) the statistical properties of the variables, and

3) the form of the function relating the observables and the quantities being estimated.

In general, the particular problems of interest are representative of a class of problems which is the subject of the general theory of parameters estimation as set forth in statistical decision theory. Therefore, the fundamental concepts of the theory of parameter estimation form a basis for an adequately unified approach to fulfill the present requirements. It should be noted that the simple derivation of filtering methods employs some of the basic concepts of the theory of parameter estimation explicitly, while others are almost always implicitly involved. However, when these concepts are not consistently employed on an explicit basis, their applicability and usefulness are not fully realized or exploited. In the subsequent sections on estimation, the basic concepts of the general theory of parameter estimation are presented for the primary purpose of formulating a more unified approach to determining filtering methods than the simple approaches outlined previously. The discussions do not present an exhaustive treatment of the subject, nor is one intended; rather, primary emphasis is placed upon the basic concepts which

have general applicability and particular usefulness to the present problems. Nonetheless, adequately complete discussions of the concepts are presented so that extensions can be formulated and applied to those problems which require them.

It should be emphasized at the outset that the problem of state estimation in space navigation and guidance is completely equivalent to the problem of transmission and reception of information in a noisy communication channel. All of the methods utilized for solving the latter problem are totally and directly applicable to the former one. Further, since extensive application of the general theory of parameter estimation has been made to the general problem of communications in the presence of noise, leading to a general theory of statistical communications, the same approach to state estimation will yield a general theory of statistical navigation and guidance. Should questions arise during the discussions, it is likely that answers can be found in References such as 0.4.

In the discussions, it is assumed that the reader is generally familiar with the fundamentals of probability theory. To be specific, knowledge is assumed of: (1) continuous joint probability distribution and density functions; (2) marginal and conditional probability density functions; and (3) conditional expectations. Though an extensive knowledge of statistics is not required. This level of familiarity must be assumed since to do otherwise would require the development of all of the statistical concepts to be employed. Thus, should the terminology be unfamiliar, the reader is referred to any of a number of excellent references (References 0.4, 3.1, 3.2, etc.).

The discussions begin (section 2.3.2) with basic notation and definitions, which are of particular concern to the present subject, and which are not necessarily emphasized in the fundamentals of probability theory. This material is followed (section 2.3.3) with a description of the basic problem to be considered in mathematical form and a discussion of some physical interpretations.

In section 2.3.4, the rudiments of parameter estimation are discussed, and a basic description of the problem is given which emphasizes the underlying concepts. These discussions define estimators, estimation error and basic properties of estimators. Properties of a "good" estimator are then discussed to form a basic concept of estimation. Next, more general criteria for estimation are defined in terms of loss functions and associated risk. Properties of estimators which are based on estimator risk follow these discussions. Finally, sufficiency of an estimator is discussed, a sufficient statistic is defined, and a test for sufficient statistics is given.

The determination of particular estimators is discussed in section 2.3.5 with primary attention given to useful methods for determining estimators. The determination of minimum variance unbiased estimators by means of sufficient statistics and complete probability density functions is discussed, and the method of least-squares estimation is shown to yield minimum variance estimators under the condition of statisticsl independence of the sample. Attention then turns to the development of a lower bound for estimator variance and the determination of Bayes estimators. The general solution (in terms of the

Bayes estimator) for the case of squared error loss is shown to be the conditional expectation. (More general loss functions are also considered.) Following this material, Bayes risk is discussed and a comparison of a Bayes estimator and a minimum variance estimator is given to illustrate the difference in the results. Finally, minimum risk estimation is discussed to emphasize the use of sufficient statistics in minimizing risk for the general class of convex loss functions. (Maximum likelihood estimation is discussed and shown to be equivalent to Bayes estimation with a simple loss function.)

These discussions conclude (section 2.3.6) with the application of Bayes estimation to the case of many degrees of freedom. The first result is the formulation of the technique to develop the optimum estimator for the general vector case and for general statistical distributions. The second result is the extension of this material to the general linear case, where the statistics are Gaussian. This extension develops a proof that (under these assumptions) the Kalman estimator is the optimum estimator. This fact could not be developed from the material presented in section 2.2.

2.3.2 Basic Definitions

It is the purpose of the following sections to develop a basic understanding of the general problem of parameter estimation. This effort must necessarily begin with the following basic definitions of the essential elements of the problem.

2.3.2.1 Random Processes

A random process is defined herein as any phenomenon for which repeated observations, under a given set of conditions, do not yield identical results. In general, random processes are characterized by variations in outcomes for repeated equivalent trails. These variations in outcomes or observations are considered as the"randomness" of the process, which is equivalent to uncertainty in the outcome of the process. As a contrary example, consider a process whose behavior is completely described by a known system of differential equations. Theoretically, it is possible to completely determine the behavior of such a process if a sufficient finite set of initial conditions are known. Thus, it would be possible to completely specify the future behavior of such a process if an adequate set of observations are made at some time. Such a process is said to possess deterministic regularity. However, until such time that all physical laws are explicitly established for the microscopic and infinitesimal domains, the concept of random physical processes must be admitted, accepted, and dealt with.

Alternatively, a random process could be defined as one which does not possess deterministic regularity, and subsequent outcomes cannot be predicted with certainty from a set of observations of the process. However, a random process can possess definite properties of behavior which make possible a description on a statistical basis. Such random processes are said to possess statistical regularity. In such cases, even though particular outcomes of the process cannot be specified, it is possible to specify the relative frequency or probability of occurrence of outcomes for the process. That is, let y denote the outcome of a random process, then the process is

described by specifying the probability that y will lie in some arbitrary interval. The common notation is as follows:

$$P(a < y \leq b) = \int_a^b f(y)\,dy \tag{3.1}$$

In this notation $P(a < y \leq b)$ is referred to as the probability distribution function which is the probability that y lies in the interval $a < y \leq b$, and $f(y)$ is the probability density function of y.

Thus, random processes are explicitly described by specifying their probability density functions. In Section 2.3.3 a detailed mathematical description is given for random processes of particular interest in the present discussions.

For the sake of notational convenience, "pdf" will be used to denote "probability density function" in the text, and "f(y)" will denote the pdf of y in equations. However, it should be noted that if x and y are two different random processes the pdf of x, f(x), is not equal to the pdf of y, f(y), even for y = x.

2.3.2.2 Parameters

If a random process possesses statistical regularity, then it can be described by specifying all of its statistical properties, which is equivalent to specifying the pdf of the process. The "parameters" of a random process are defined as the smallest set of elements which specify the statistical properties of the process or its pdf. In general, all statistical averages, or moments, of a random process are required to specify it. However, for many processes of interest, all moments are not required, and a smaller set of parameters suffices to specify the process.

For example, in the case of a Gaussian distributed random process, it is only necessary to specify the mean and variance of the process, since the pdf of the process, f(y), is specified by these two parameters; i.e., if y is the outcome of a Gaussian distributed random process, then

$$f(y) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right] \tag{3.2}$$

where μ and σ^2 (the mean and variance of y, respectively) represent the two parameters which specify the pdf of a Gaussian random process.

In general, the set of n parameters which specifies the statistical properties of the pdf of a random process will be denoted by $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$. The pdf of a random process will be shown as a function of the parameter set θ in terms of the conditional pdf, given θ; i.e., $f(y/\theta)$. For the Gaussian pdf $\theta = (\theta_1, \theta_2) = (\mu, \sigma^2)$ and $f(y/\theta) = f(y/\mu, \sigma^2)$.

2.3.2.3 Random Samples

A random sample is defined as a collection of observations or outcomes of a random process. Specifically, a random sample of size m is a set which contains as elements the results of m observations of a random process. A random sample will be denoted by $Y = (y_1, y_2, \ldots, y_m)$. It is important to note that a random sample has a joint pdf which essentially specifies the probability of the simultaneous occurrences of the m observations. That is,

$$P\left[a_1 < y_1 \leq b_1 ;\ a_2 < y_2 \leq b_2 ;\ \cdots ;\ a_m < y_m \leq b_m\right] = \int_{a_1}^{b_1}\int_{a_2}^{b_2}\cdots\int_{a_m}^{b_m} f(y_1, y_2, \cdots, y_m)\, dy_m \cdots dy_2\, dy_1 \tag{3.3}$$

For convenience, the following notation will be used to denote the probability of simultaneous occurrences of a random sample of m observations.

$$P\left[D(Y)\right] = \int_{D(Y)} f(Y)\, dY \tag{3.4}$$

Where D(Y) is any m dimensional domain of interest, P[D(Y)] is the probability that the random sample will lie in D(Y), f(Y) is the joint pdf of the random sample Y, dY is an m dimensional infinitesimal, and it is understood that the integral must be performed over the domain D(Y).

In general, the m observations of a random sample can be statistically independent or dependent. The random sample is defined to be independent if

$$f(Y/\theta) = f(y_1/\theta)\cdot f(y_2/\theta)\cdots f(y_m/\theta) = \prod_{i=1}^{m} f(y_i/\theta) \tag{3.5}$$

where $f(y_i/\theta)$ is the pdf, given the parameters θ, for a single observation of the random process. Otherwise the random sample is statistically dependent.

As a particular example of $f(y/\theta)$, consider m independent observations of a Gaussian random process. For this case,

$$f(Y/\theta) = f(Y/\mu, \sigma^2)$$

$$f(Y/\theta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)\exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{m}(y_i-\mu)^2\right] \tag{3.6}$$

2.3.2.4 Random Variables

A random variable is defined as a real-valued function of F(y) which exists and is defined for each outcome of a random process. Of course, the outcomes for many random processes are actually random variables; i.e., F(y) = y. Such random processes are considered to be quantitative or numerical processes, e.g., random voltages, pressures, errors, etc. On the other hand, random processes exist which are non-numerical, such as the tossing of a coin where the outcome is either a heads or tails. However, it is possible

to define a random variable for this random process by assigning numbers to the outcomes or by defining the random variable to the number of heads in m tosses of a coin, etc.

The importance of the concept of a random variable lies in the fact that many of the arithmetic, algebraic and analytical operations which are defined for real-valued functions are meaningful for random variables, whereas they are not for the outcomes of all random processes. Thus, additions, subtractions, multiplications, transformations, etc., are applicable to random variables.

In general, if y possesses statistical regularity, then the random variable F(y) possesses statistical regularity. Thus, F(y) is generally specified by a pdf which is derived from that of y. The derivation of f [F(y)] from f(y) is, in general, not a simple transformation, except in the simple case of F(y) = ay for which f [F(y)] = a · f(y), where a is a constant. However, the expected values of random variables are usually required, and the pdf of F(y) is not required since the expected value of any random variable F(y) can be determined from pdf of y as follows:

$$E[F(y)] = \int_{-\infty}^{\infty} F(y)\, f(y)\, dy$$

where E denotes "expected value." The conditional expectation of F(y), given θ, is defined as

$$E[F(y)/\theta] = \int_{-\infty}^{\infty} F(y)\, f(y/\theta)\, dy \quad .$$

2.3.2.5 Statistic

A statistic is a known function of a random sample of observations of a random process whose outcomes are random variables. It follows that a statistic is a random variable; however, a random variable is not necessarily a statistic. For example, in the case of a Gaussian random variable y with unknown mean value μ, the function $y - \mu$ is a random variable; however, it is not a statistic since μ is not known. On the other hand, y - c, where c is a known constant, is a statistic.

The important difference is that a statistic is defined as a known function of a random variable sample set Y which does not contain any unknown elements. Thus, a particular sample Y specifies the statistic. Of course, statistics are not unique since many known functions of a random variable sample can be defined. In general, the set of statistics defined for the random variable sample set Y will be denoted by
$T(Y) = [T_1(Y), T_2(Y), T_3(Y) \ldots .]$.

A statistic represents a transformation of the sample set Y from the space of random variable observations to the space of statistics. The transformation is not unique for a given statistic. For example, a statistic can be the simple linear sum of random variable observations in which case many

different sample sets transform into the same point in the statistic space (i.e., various sample sets can yield the same numerical value of the statistic).

The conditional expectation of T(Y), given Θ, is determined from the conditional joint pdf of Y, i.e.,

$$E\left[T(Y)/\Theta\right] = \int_{Y} T(Y) f(Y/\Theta)\, dY$$

where the integral is m dimensional and must be performed over all y in Y.

2.3.2.6 Complete Probability Density Functions

In the consideration of parameter estimation, a question often arises concerning the uniqueness of the expectation of a random variable. This question can be resolved in terms of the property of completeness of pdf's. A random process, y, is specified by its pdf, $f(y/\Theta)$, which is a function of Θ. For any interval of Θ, $A < \Theta < B$, $f(y/\Theta)$ defines a family of pdf's. The following definition defines a complete family of pdf's.

DEFINITION: let $f(y/\Theta) > 0$ for any interval of $a < y < b$ and zero otherwise and let F(y) be a random variable defined on the interval $a < y < b$ and independent of Θ.

Now, if

$$E\left[F(y)\right] = 0 \qquad (A < \Theta < B)$$

if, and only if, $F(y) \equiv 0$ for $a < y < b$, then $f(y/\Theta)$ is defined to be a complete family of pdf's. Conversely, if there exists some F(y) which is not identically zero in $a < y < b$ for which $E\left[F(y)\right] = 0$ for all $A < \Theta < B$, then $f(y/\Theta)$ is not a complete family of pdf's.

It should be noted that a pdf can be complete with respect to certain parameters. For example, consider the Gaussian pdf, $f(y/\mu, \sigma^2)$, to wit,

$$f(y/\mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right]$$

If the mean value μ is known, then any random variable F(y) which is defined as an odd function about μ has zero expected value and, hence, $f(y/\mu, \sigma^2)$ is not complete with respect to σ^2 for known μ. On the other hand, if μ is unknown, then $f(y/\mu, \sigma^2)$ is complete with respect to both μ and σ^2 since the only F(y) which has zero expected value for all μ and σ^2 is given by $F(y) \equiv 0$. Of course, $f(y/\mu, \sigma^2)$ is complete for known σ^2 if μ is unknown.

The use of completeness can be demonstrated by considering $f(y/\mu, \sigma^2)$ for $\sigma^2 = 1$. Let $F(y) = y^2$; then $E\left[F(y)\right] = E(y^2) = \mu^2 + 1$. The property of completeness determines that there is no other $F(y) \neq y^2$ such that $E[F(y)] = \mu^2 + 1$. For if $E[F(y)] = E(y^2)$, then $E[F(y) - y^2] = 0$. But since $f(y/\mu, 1)$ is complete, $F(y) - y^2$ must be identically zero for all y; i.e., $F(y) \equiv y^2$. Thus, since $f(y/\mu, 1)$ is complete, the $E[F(y)]$ is uniquely determined.

2.3.3 A Mathematical Description of Random Processes

In order to promote an understanding of parameter estimation with respect to the particular problems of interest, a mathematical description is discussed in this section which describes, in general form, the set of random processes to be considered.

2.3.3.1 A General Form

For present purposes random processes will be defined in general form as follows

$$\underline{y} = \underline{F}(\underline{\theta}) + \underline{e} \tag{3.7}$$

where

$\underline{y}$ = m x 1 Observation vector

$\underline{\theta}$ = n x 1 Parameter vector

$\underline{F}$ = m x 1 Transformation vector

$\underline{e}$ = m x 1 Noise vector

Equation (3.7) defines a random process as a general non-linear function of the parameter vector $\underline{\theta}$. The general linear case is defined by $\underline{F}(\underline{\theta}) = A\underline{\theta}$; i.e.,

$$\underline{y} = A\underline{\theta} + \underline{e} \tag{3.8}$$

where

A = m x n Transformation matrix

The foregoing terms have the following properties.

$\underline{e}$ Is always random

$\underline{\theta}$ Is either random or non-random

$\underline{F}$ Is always non-random

A Is always non-random

$\underline{y}$ Is always random

Thus, for present purposes, a random process is defined as the linear sum of two processes $\underline{F}(\theta)$ and $\underline{e}$. The process $\underline{e}$ is always a random process. However, since $\underline{F}$ is always non-random, the process $\underline{F}(\underline{\theta})$ is randon if, and only if, $\underline{\theta}$ is random. It is extremely important to note that for non-random $\underline{\theta}$, the process $\underline{F}(\underline{\theta})$ possesses deterministic regularity. Furthermore, the process $\underline{F}(\underline{\theta})$ possesses conditional deterministic regularity, given θ; i.e., for a particular $\underline{\theta}$, $\underline{F}(\underline{\theta})$ is deterministically regular. Of course, $\underline{y}$ is always random

since $\underline{e}$ is always random.

A more detailed discussion of the general form of random processes is given in subsequent sections. However, at present it is advisable to discuss some physical interpretations of the random process as defined herein.

2.3.3.2 Some Physical Interpretations

In the general form of a random process, the vector $\underline{\Theta}$ represents the essential unknowns which are to be determined. For example, in the linear case, $\underline{\Theta}$ represents the "state" vector which is commonly used to denote the position and velocity deviations of a spacecraft from a reference trajectory. However, in the present definition $\underline{\Theta}$ is quite general and can represent other parameters such as the coefficients of a polynomial fit to data (as in the case of least-squares curve fitting) or the bias arising from a random process, etc.

The vector $\underline{F}(\underline{\Theta})$ represents some observable or measurable physical phenomena which is dependent on $\underline{\Theta}$. In general, $\underline{F}$ represents a set of known non-random functions of $\underline{\Theta}$. In the case where $\underline{\Theta}$ represents spacecraft position and velocity deviations from a reference trajectory, $\underline{F}(\underline{\Theta})$ usually represents deviations of space measurements from the reference. In this case, $\underline{F}(\underline{\Theta})$ usually becomes $A\underline{\Theta}$ where A is the matrix of first partial derivatives of space measurements with respect to spacecraft position and velocity. It should be noted that, in general, $\underline{F}$ is not necessarily constant; e.g., $\underline{F}$ can vary as a non-random function of time. For example, $\underline{\Theta}$ can denote trajectory injection deviations which are propogated into subsequent values of trajectory deviations as a function of time. In such case, $\underline{F}(\underline{\Theta})$ is a function of a time-dependent transition matrix.

The vector $\underline{e}$ represents errors in measurements of $\underline{F}(\underline{\Theta})$ or, equivalently, the uncertainty in observations. If the errors in measurements were not present, then,theoretically, the parameters $\underline{\Theta}$ could be determined directly from the inverse of $\underline{y} = \underline{F}(\underline{\Theta})$; i.e., for $\underline{e} \equiv 0$, $\underline{\Theta} = \underline{F}^{-1}(\underline{y})$ where $\underline{F}^{-1}$ exists and y is an adequate set. However, in the general case, $\underline{e}$ is present and $\underline{\Theta}$ cannot be determined from $\underline{F}^{-1}$ without the risk of large errors in the results, e.g., in the linear case where $\underline{y} = A\Theta + \underline{e}$, $A^{-1}y = \Theta + A^{-1}\underline{e}$ and the error in taking $\Theta = A^{-1}y$ is $A^{-1}\underline{e}$ which could be large depending on the nature of A^{-1} and $\underline{e}$.

The vector $\underline{y}$ is a random variable vector which represents the random process in general. A particular value of $\underline{y}$ represents a particular observation of the random process. The totality of observations comprises a random sample set Y which contains all observations of the random process. The vector $\underline{y}$ and the set Y can be interpreted in two equivalent forms. In general, the totality of observations can be taken in a time sequence of simultaneous observations of the elements of $\underline{y}$. In this case, Y contains the set of particular observations of $\underline{y}$, i.e., $Y = (\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_m)$. However, the vector $\underline{y}$ can represent the totality of observations with the elements of Y as subvectors of $\underline{y}$, in which case Y and $\underline{y}$ are the same set. The particular interpretation of $\underline{y}$ is selected on the basis of convenience for a particular problem.

2.3.3.3 An Ensemble of Non-Stationary Random Processes

In the general form of a random process, as defined in equation (3.7), the term $\underline{F}(\underline{\theta})$ is dependent on $\underline{\theta}$ and can vary as a function of time. Thus, the general form of random processes actually defines an ensemble of non-stationary random processes. Each particular value of $\underline{\theta}$ defines a particular member of the ensemble of random processes defined. Since $\underline{F}(\underline{\theta})$ can vary with time for any $\underline{\theta}$, the ensemble members are non-stationary; however, if $\underline{F}(\underline{\theta})$ is independent of time, i.e., constant for all particular $\underline{\theta}$, then each member of the ensemble is stationary, assuming that the random process $\underline{e}$ is stationary.

Two important aspects of the ensemble of random processes should be strongly emphasized: First, it is important to note that when a set of observations are made of a random process, a particular member of an ensemble is observed and determining $\underline{\theta}$ is equivalent to determining which member of the ensemble has been observed. That is, during a sequence of observations, $\underline{\theta}$ is a constant which, of course, is unknown. The object of making observations is to determine $\underline{\theta}$, if possible. It should be noted that for each ensemble member, the deterministic regularity resulting from $\underline{F}(\underline{\theta})$ is reflected as statistical regularity in the observations $\underline{y} = \underline{F}(\underline{\theta}) + \underline{e}$.

Second, it is also important to note that although $\underline{\theta}$ is an unknown to be determined, certain information is often available concerning the behavior of $\underline{\theta}$. Indeed, $\underline{\theta}$ is often a random variable with known pdf. In such cases, it is possible to determine the probability of occurrence of particular members of the ensemble. In general, members of the ensemble occur with various probabilities and certain members can occur with zero probability. If $\underline{\theta}$ is a random variable, then the ensemble of random processes is specified by the joint pdf of $\underline{\theta}$. Generally, $\underline{\theta}$ exists over some parameter space, Ω, and the pdf of $\underline{\theta}$ determines the probability of occurrence of a particular $\underline{\theta}$ in the parameter space Ω.

2.3.4 Rudiments of Parameter Estimation

2.3.4.1 A Basic Description of Parameter Estimation

The problem of parameter estimation can be described most succinctly in the following manner:

1) A random process exists which is a function of or characterized by a set of parameters.

2) The parameters are not explicitly known nor can they be directly observed.

3) Knowledge of the parameters is required to perform some particular task.

4) Observations of the random process can be made which yield a set of sample data.

5) The set of sample data provides the only means for determining the required parameters.

The foregoing is equivalent to the method of inductive inference, which is referred to as statistical inference when random processes are involved. That is, from a particular set of observations, conclusions are drawn which concern some general aspects of the process under observation. The method is depicted in Figure 2.3.1. It is a basic theorem of formal logic that the method of inductive inference is intrinsically uncertain. It is not possible to make generalizations with certainty on the basis of a set of particular observations. The situation is apparent when random processes are involved. Nonetheless, useful inferences can be made if the procedures involved are judiciously formulated. This is the general concern of the theory of statistical inference and the particular concern of statistical decision theory.

In the general theory of statistical inference, the basic problem which is considered, is that of making a decision under existing conditions of uncertainty. Two types of uncertainty are recognized: randomness and the lack of knowledge concerning the state of nature. It is required to make a decision concerning the state of nature, and it is desirable to make the "best decision possible" under the circumstances. In effect, the decision is an estimation of the state of nature; thus, if the state of nature is determined by a set of parameters, the decision problem of the parameter estimation.

The foregoing situation is completely equivalent to the problem of determining the set of parameters $\underline{\theta}$ of the random process $\underline{y} = \underline{F}(\underline{\theta}) + \underline{e}$. The uncertainty due to randomness is equivalent to the errors in measurements or observations, denoted by $\underline{e}$, and the uncertainty due to the unknown state of nature is equivalent to that of the unknown parameters $\underline{\theta}$. Of course, "best decision possible" is equivalent to "best parameter estimation possible." Thus, the problem of parameter estimation becomes that of determining estimations of $\underline{\theta}$ from a set of observation data which are "best" with respect to some applicable criteria. Some comments concerning the general philosophy are in order.

At the outset (before observations are taken) there exists uncertainty about the parameters $\underline{\theta}$. However, some knowledge of $\underline{\theta}$ can exist a'priori since the pdf of $\underline{\theta}$ is frequently known; i.e., the probability of occurrence of a particular $\underline{\theta}$ can be known a'priori. Thus, in general, complete uncertainty concerning $\underline{\theta}$ does not exist at the outset. Nonetheless, in each particular case, it is required to know $\underline{\theta}$ with less uncertainty on greater certainty (as contrasted with absolute certainty) than exists a'priori. The statistical regularity of the process, which is characterized by the parameter $\underline{\theta}$, is present in the observations of the process; hence, the observations must contain intrinsic information concerning the parameters $\underline{\theta}$. Thus, the object of parameter estimation becomes that of extracting information concerning $\underline{\theta}$ from the observations thereby decreasing the a'posteriori (after observations) uncertainty of $\underline{\theta}$ or, equivalently, increasing the a'posteriori knowledge of $\underline{\theta}$. It is interesting to note that the problem is identically equivalent to the problem of information transmission and reception in the presence of noise, which is considered in the general theory of statistical communication. Indeed, the problem is one of extracting information concerning $\underline{\theta}$ from the observations $\underline{y}$ in the presence of measurement or observation error $\underline{\theta}$.

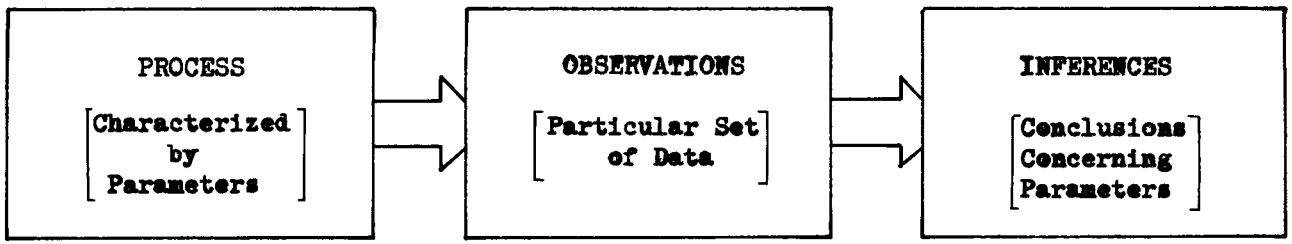


Figure 2.3.1: Inductive Inference

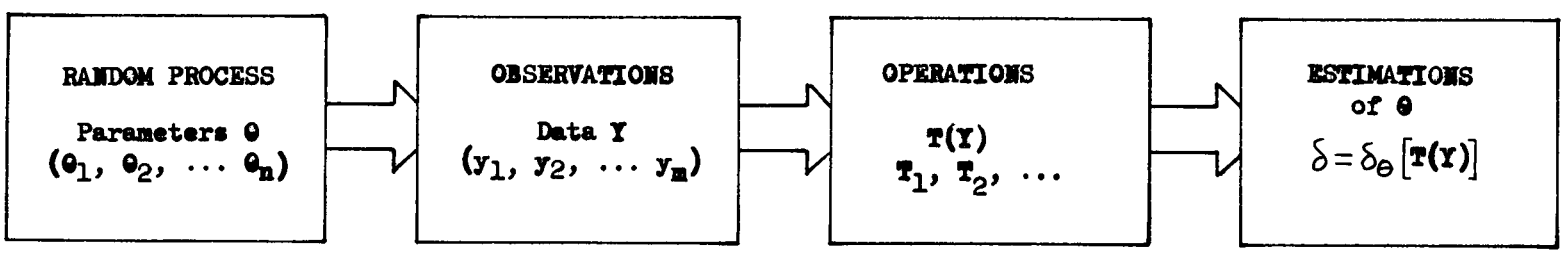


Figure 2.3.2: Basic Procesure of Parameter Estimation

In general, it is not possible to determine the parameters $\underline{\theta}$ with absolute certainty (unless an infinite number of observations are made, which is generally not practical); thus, some residual or a'posteriori uncertainty must be expected. The concept of "best parameter estimation possible" is equivalent to the minimization of a'posteriori uncertainty or the extraction of maximum information from the observations.

In general, parameter estimations are derived from operations which are performed on the set of observations or random sample data. These operations are defined as estimators and are denoted herein by δ. The general procedure of parameter estimation is depicted in Figure 3.2. In the following section, estimators δ are discussed in further detail.

2.3.4.2 Estimator

In general, an estimator δ is a known function of statistics (see Section 2.3.2.5) which provides an estimation of a parameter θ. Of course, an estimator is a random variable; in fact, estimators are a subclass of the class of statistics. (That is, an estimator is a statistic but a statistic is not necessarily an estimator.) In many cases the difference can be trivial; however, in the consideration of parameter estimation, it is extremely important to consider all possible known functions of random samples, i.e., statistics. The ambiguity can be resolved by considering a statistic as an admissible estimator if its conditional expectation, given θ, contains θ explicitly. That is, if $E(T/\theta) = \theta + b(\theta)$, then the statistic T is an admissible estimator δ for θ, where $b(\theta)$ is some function of θ, which is referred to as estimator "bias." It follows that

$$E(\delta/\theta) = \theta + b(\theta) \ ,$$

Thus, in general, admissible estimators can contain a "bias" term $b(\theta)$ which is a function of θ. However, all admissible estimators are not necessarily desirable since estimator bias, $b(\theta)$, whereas admissible, is not necessarily desirable in particular cases.

It should be noted that the term "estimator" denotes a function, whereas "estimation" or "estimate" generally denotes a particular value of an estimator as a function of a particular random sample.

There are two major objectives in the problem of parameter estimation. The first of these objectives is to determine parameter estimators δ with minimum uncertainty; and the second is to appraise the estimator's uncertainty in terms of its magnitude and behavior as a function of significant factors; e.g., sample size, random process characteristics, etc. Estimator uncertainty is usually measured by estimator error, which is discussed below.

2.3.4.3 Estimator Error

In general, parameter estimations are inductive inferences and as such are always subject to uncertainty. The primary objective of parameter estimation analyses is to determine estimators with minimum acceptable uncertainty. In order to approach this problem on a mathematical basis, it is necessary to express estimation uncertainty in explicit mathematical form.

Thus, with each estimator there is associated an estimator error, denoted by ϵ, which is defined as the linear difference of a parameters and its estimation, to wit,

$$\epsilon = \delta - \theta \tag{3.9}$$

The problem of parameter estimation can be considered as equivalent to an analysis of estimator error. Thus, at the outset, several general properties of should be noted. First, estimator error is a function of the sample data set and the parameter θ; i.e.,

$$\epsilon(Y,\theta) = \delta\left[T(Y)\right] - \theta \tag{3.10}$$

Second, since ϵ is a function of random variables, it is a random variable and must be analyzed on a statistical basis, i.e., ϵ is specified by determining its pdf, $f(\epsilon)$. Third, since ϵ is a function of Y and θ, $f(\epsilon)$ is a function of $f(Y)$ and $f(\theta)$, to wit,

$$f(\epsilon) = G\left[f(Y), f(\theta)\right] \tag{3.11}$$

Fourth, the particular form of $f(\epsilon)$ is dependent on the particular estimator $\delta\left[T(Y)\right]$.

In the analysis of parameter estimators, primary emphasis is placed upon the statistical behavior of estimator error. This behavior is dependent upon the characteristics of the random process (especially as a function of the parameters being estimated) and the particular form of the estimator as a function of the sample data. The ultimate objective in parameter estimation is to determine parameter estimators which yield acceptable behavior of the estimator error. Of course, suitable statistical properties of estimator error must be used as the object of analysis. The most significant properties are the first moment, first central moment, and second moment, which are usually referred to as the mean value, the variance and the mean squared value of estimator error, respectively. These quantities are defined below:

(1) Mean value of $\epsilon = \bar{\epsilon}$

$$\begin{aligned}\bar{\epsilon} &= E(\epsilon) \\ &= \int_{\epsilon} \epsilon\, f(\epsilon)\, d\epsilon \\ &= E(\delta) - E(\theta) \\ &= \bar{\delta} - \bar{\theta}\end{aligned}$$

(2) Variance of ϵ = $V(\epsilon)$

$$V(\epsilon) = E(\epsilon - \bar{\epsilon})^2$$

$$= \int_{\epsilon} (\epsilon - \bar{\epsilon})^2 f(\epsilon)\, d$$

(3) Mean-Squared Value of ϵ = $\overline{\epsilon^2}$

$$\overline{\epsilon^2} = E(\epsilon^2)$$

$$= \int_{\epsilon} \epsilon^2 f(\epsilon)\, d\epsilon$$

The mean value of $\bar{\epsilon}$ is commonly referred to as estimator bias and if $\bar{\epsilon} = 0$, then ϵ is considered unbiased. The most significant aspect of estimation error is the effect of $\bar{\epsilon}$ on $V(\epsilon)$ and $\overline{\epsilon^2}$. This effect is stated below.

If for all θ, the conditional expectation of ϵ, given θ, is zero, then the mean-squared value of ϵ is equal to $V(\epsilon)$ and the estimator variance, $V(\delta)$. This can be shown in the following manner [E_θ denotes expectation with respect to θ and $\bar{\delta}_\theta$ denotes the conditional expectation $E(\delta/\theta)$]:

$$\bar{\epsilon} = E_\theta E(\epsilon/\theta)$$

$$E(\epsilon/\theta) = \bar{\delta}_\theta - \theta$$

Thus, if $E(\epsilon/\theta) = 0$ for all θ, then $\bar{\epsilon} = 0$ and $\overline{\epsilon^2} = V(\epsilon)$. That $\overline{\epsilon^2} = V(\delta)$ for $E(\epsilon/\theta) = 0$ is shown below:

$$\overline{\epsilon^2} = E_\theta E(\epsilon^2/\theta)$$

$$E(\epsilon^2/\theta) = E\left[(\delta - \theta)^2/\theta\right]$$

$$= E\left[(\delta - \bar{\delta}_\theta + \bar{\delta}_\theta - \theta)^2/\theta\right]$$

$$= E\left\{\left[(\delta - \bar{\delta}_\theta)^2 + 2(\delta - \bar{\delta}_\theta)(\bar{\delta}_\theta - \theta) + (\bar{\delta}_\theta - \theta)^2\right]/\theta\right\}$$

$$= V(\delta/\theta) + E\left\{\left[\theta^2 - 2\theta\bar{\delta}_\theta + \bar{\delta}_\theta^2 + 2(\delta\bar{\delta}_\theta - \theta\delta + \theta\bar{\delta}_\theta - \bar{\delta}_\theta^2)\right]/\theta\right\}$$

$$= V(\delta/\theta) + \theta^2 + \bar{\delta}_\theta^2 - 2\theta\bar{\delta}_\theta + 2\bar{\delta}_\theta^2 - 2\bar{\delta}_\theta^2$$

$$= V(\delta/\theta) + \theta^2 - 2\theta\bar{\delta}_\theta + \bar{\delta}_\theta^2$$

$$= V(\delta/\theta) + (\bar{\delta}_\theta - \theta)^2$$

$$E(\epsilon^2/\theta) = V(\delta/\theta) + E^2(\epsilon/\theta) \qquad (3.12)$$

The second term is the square of $E(\epsilon/\theta)$ and is always positive, thus,

$$\overline{\epsilon^2} \geq V(\delta) \tag{3.13}$$

with equality if $E(\epsilon/\theta) = 0$ for all θ. Therefore, the following is true:

If $E(\epsilon/\theta) = 0$ for all θ, then $\overline{\epsilon^2} = V(\epsilon) = V(\delta)$.

It should be noted that $E(\epsilon/\theta) = 0$ for all θ is a sufficient condition for $\overline{\epsilon^2} = V(\epsilon) = V(\delta)$. In general, the condition is also necessary if the pdf of θ is non-zero from all θ. On the other hand, if pdf of θ is zero for some interval of θ, then $E(\epsilon/\theta)$ can be non-zero for the same interval of θ, and $\overline{\epsilon^2} = V(\epsilon) = V(\delta)$. This is true because if $f(\theta) = 0$ for some interval of θ, then those θ in this interval occur with zero probability and contribute nothing to the expectation with respect to θ. It should also be noted that if $\bar{\epsilon} = 0$, then $\overline{\epsilon^2} = V(\epsilon)$; however, it does not follow that $\overline{\epsilon^2} = V(\delta)$, since for $\bar{\delta} = \theta$ it is not required that $E(\epsilon/\theta) = 0$.

2.3.4.4 Basic Properties of Good Estimators

The primary motivation for the concept of a "good" estimator is derived from a consideration of the basic properties of estimator error discussed previously. Consider three different estimators δ_1, δ_2 and δ_3 for a parameter θ. Let the estimator errors for the three estimators be denoted by ϵ_1, ϵ_2 and ϵ_3, respectively, and assume that they have the following properties for all θ.

$$E(\epsilon_1/\theta) = 0$$

$$E(\epsilon_2/\theta) = 0$$

$$E(\epsilon_3/\theta) = \overline{\epsilon^3}$$

$$V(\epsilon_1) < V(\epsilon_2) < V(\epsilon_3)$$

The pdf's of δ_1, δ_2 and δ_3 are depicted below in Figure 2.3.3.

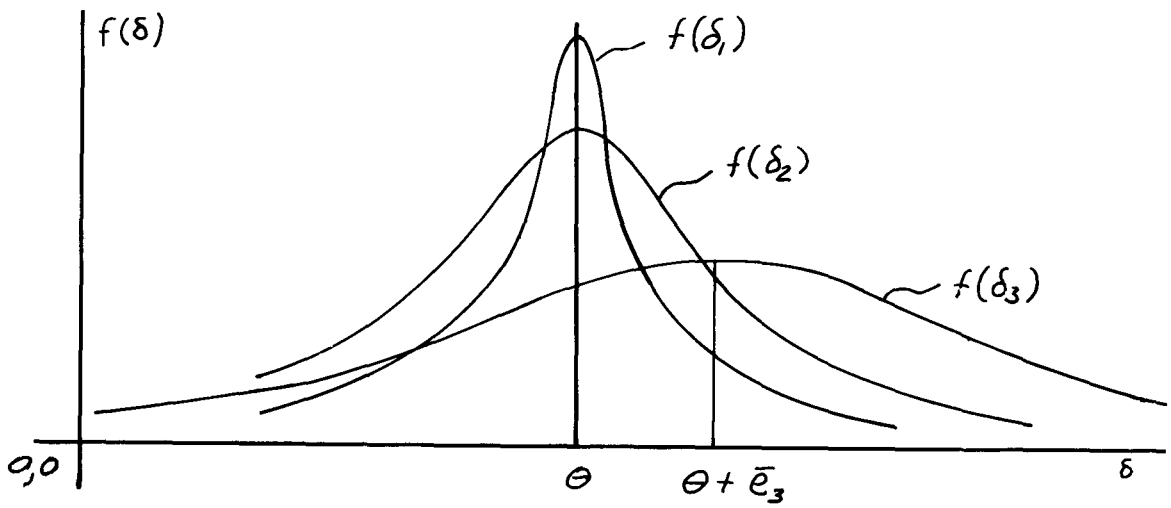


Figure 2.3.3: Illustration of Estimator pdf's

Inspection of the pdf's of δ_1, δ_2 and δ_3 leads to the conclusion that δ_1 is the "best" estimator of the three since it has the minimum variance of the three and is unbiased. This conclusion is further substantiated by considering the mean-squared values of ϵ_1, ϵ_2 and ϵ_3. From equation (4.3) and the specified properties for ϵ_1, ϵ_2 and ϵ_3 it is seen that

$$\overline{\epsilon_1^2} < \overline{\epsilon_2^2} < \overline{\epsilon_3^2}$$

Thus, it would again be concluded that δ_1 is the best estimator of δ_1, δ_2 and δ_3, since its estimation error has the minimum mean-squared value of the three considered.

In the foregoing situation, there is no problem selecting the best estimator of the three considered. However, in the general case some difficulty can arise. Consider a fourth estimator δ_4 whose variance is less than that of δ_1 but which has a non-zero mean value; i.e., $E(\epsilon_4/\theta) = e_4 \neq 0$. Of the four estimators δ_4 is the one with minimum variance; however, it is not necessarily the one with minimum mean-squared error. That is, although $V(\epsilon_4) < V(\epsilon_1)$, it is not necessarily true that $\overline{\epsilon_4^2} < \overline{\epsilon_1^2}$ since $\epsilon_4 \neq 0$.

Thus, in the set of all possible estimators for θ the one which has minimum variance does not necessarily have minimum mean-squared-error. However, if the estimator with minimum variance also has zero mean value for all θ then the minimum variance estimator is also the one with minimum mean-squared-

error in the subset. But, the minimum variance estimator in this subset is not necessarily the minimum mean-squared-error estimator of the total set of estimators.

It is apparent that, in general, some question exists concerning the selection of a "best" estimator. The final answer is usually dependent upon the particular problem being considered. Nonetheless, from the foregoing considerations some properties of "good" estimators can be formulated which actually classify or define sets of estimators as follows.

2.3.4.4.1 Unbiased Estimators

An estimator is defined as an "unbiased" estimator of Θ if

$$E(\delta/\theta) = \Theta \quad \text{for all } \Theta .$$

It follows that an unbiased estimator has zero mean-value of estimator error, i.e.,

$$E(\epsilon/\theta) = 0 \quad \text{for all } \Theta .$$

2.3.4.4.2 Minimum-Variance Estimators

An estimator is defined as a "minimum-variance" estimator of Θ if it has the smallest variance of all estimators of θ, for all Θ.

2.3.4.4.3 Minimum-Variance Unbiased Estimators

An estimator is defined as a "minimum-variance unbiased" estimator if it has the smallest variance of all unbiased estimators of θ, for all Θ.

2.3.4.4.4 Minimum Mean-Squared-Error Estimators

An estimator is defined as a "minimum mean-squared-error" estimator of Θ if it has the smallest mean-squared estimator error of all estimators of θ, for all Θ.

2.3.4.5 Loss Functions and Risk

In the previous section, basic criteria for good estimators were based upon estimator bias and variance, and mean-squared value of estimator error. Although these criteria are generally acceptable, there always exists the problem of selection between a minimum-variance and a minimum mean-squared error estimator when they are different. Furthermore, these criteria lack generality in terms of total performance or behavior of an estimator. A more general form of estimator performance criteria is formulated from the following considerations.

In the general situation of parameter estimation, there exists more penality or "loss" which is associated with an incorrect estimation. If this was not true, then nothing essential could be gained by the efforts of determining and using good estimators. The loss can be measured in terms of a non-negative function of estimator error which is a monotonically increasing function of

estimator error magnitude. Any function of estimator error with these two fundamental properties is referred to as a loss function which will be denoted by $L(\epsilon)$. Thus, a class of loss function is defined by

$$L(\epsilon) = 0 \quad \text{for} \quad \epsilon = 0$$

$$L(\epsilon_2) = L(\epsilon_1) \quad \text{for} \quad \epsilon_2 = \epsilon_1$$

$$L(\epsilon_2) > L(\epsilon_1) \quad \text{for} \quad |\epsilon_2| > |\epsilon_1|$$

$$L(\epsilon) = L(-\epsilon)$$

It is not necessary to restrict loss functions to the class with these properties, although the class has general applicability. A class of loss functions can be defined more generally as non-negative, monotonically non-decreasing with a unique coincident minimum with estimator error magnitude. Such loss functions allow a cost of measurement to be included and are minimum, rather than zero, for $\epsilon = 0$. Also, constant loss is allowed for an interval of estimation error ϵ.

Of course, negative loss would be considered as a gain; however, this is not allowed for non-zero estimator error and for zero error the loss is simply zero or some minimum. Obviously, there are no unique loss functions of estimator error; thus, a particular loss function should be selected on the basis of a particular problem. It should be noted that "squared-error" is an acceptable loss function, i.e., $\epsilon^2 = L(\epsilon)$. Others can be as easily defined.

The loss function $L(\epsilon)$ measures the loss incurred in making an error ϵ in an estimation of θ where it is assumed that some appropriate $L(\epsilon)$ can be defined for each problem of interest. Of course, $L(\epsilon)$ is a random variable dependent on estimator error $\epsilon = \delta - \theta$. Thus, $L(\epsilon) = L(\delta, \theta)$ is a function of δ and θ and measures the loss incurred in estimating θ with δ. The average value of $L(\delta, \theta)$ is the loss to be expected in using the estimator δ for θ. Expected loss is generally referred to as the "risk" taken, i.e., the risk is expected to be lost on the average. The risk in estimating θ by δ is a function of both θ and δ and will be denoted by $R(\delta, \theta)$. It follows that

$$R(\delta, \theta) = \int_Y L\left[\delta(Y), \theta\right] f(Y/\theta)\, dY$$

Risk or average loss provides a rather general criterion of estimator performance. Of course, risk is always a function of the particular loss function, which includes squared-error ϵ^2; therefore, risk includes mean-squared error ϵ^2 as a particular case. In general, it is desirable to minimize the risk involved in estimating θ by δ. Thus, for a particular loss $L(\delta, \theta)$ it is desirable to select an estimator δ which minimizes the expected loss or risk in estimating θ. This leads to a class of estimators referred to as minimum risk estimators which are defined as follows.

2.3.4.5.1 Minimum Risk Estimators

An estimator is defined as a "minimum risk" estimator for the loss function $L(\delta, \theta)$ if the expected loss or risk $R(\delta, \theta)$ is a minimum for all θ.

There is no argument concerning the desirability of minimum risk estimators; however, difficulty frequently arises in finding such estimators. It is usually possible to find an estimator which has minimum risk for certain Θ but which does not have minimum risk for other Θ. A single minimum risk estimator does not exist. Such situations are untenable since Θ is unknown beforehand. Thus, two alternate approaches are taken in minimizing risk. One approach is that of finding an estimator which minimizes the maximum risk. The other approach is that of finding the estimator which minimizes the expected value of risk over all Θ. This leads to the following two classes of estimators:

Minimum-Maximum Risk Estimators (Minimax)

An estimator is defined as a "minimax" estimator if the minimum risk is a minimum.

Minimum Expected Risk Estimators (Bayes)

An estimator is defined as a "Bayes" estimator if the expected risk over all Θ is a minimum.

The use of minimum expected risk was first introduced by Bayes and, therefore, such estimators are referred to as Bayes estimators. The expected risk is the expected value of $R(\delta, \Theta)$ with respect to Θ, i.e.,

$$E_{\Theta}\left[R(\delta, \Theta)\right] = \bar{R}\left[\delta(Y)\right] = \int_{\Theta} R(\delta, \Theta)\, f(\Theta)\, d\Theta$$

where $f(\Theta)$ is the pdf of $\underline{\Theta}$. From the equation for $R(\delta, \Theta)$ it follows that

$$\bar{R}\left[\delta(Y)\right] = \int_{\Theta}\int_{Y} L\left[\delta(Y), \Theta\right] f(Y/\Theta)\, f(\Theta)\, dY\, d\Theta$$

Thus, a Bayes estimator is one which minimizes $\bar{R}\left[\delta(Y)\right]$.

A general preference from minimax or Bayes estimators cannot be given and could vary from one problem to another. However, there is an obvious disadvantage with minimax estimators in that the Θ of maximum risk (referred as the least favorable Θ) for which the estimator δ minimizes risk can occur with very small probability; thus, on the average, the minimax estimator can perform very poorly. This disadvantage is overcome in the Bayes estimator. Thus, the Bayes estimator is generally more desirable. That is, the Bayes estimator makes use of the a'priori knowledge of Θ available in the pdf of Θ to arrive at an estimator of minimum average risk. Bayes estimation will be discussed in further detail in Section 2.3.5.2.

2.3.4.5.2 Estimator Properties Based on Risk

Two properties of estimators are based upon estimator risk. These are estimator "efficiency" and "consistency." Estimator efficiency is a relative measure of estimator risk while estimator consistence is a property of rick as sample size increases. These two properties are discussed below.

Estimator Efficiency (Relative)

Consider two estimators δ_1 and δ_2 with associated risks $R(\delta_1, \theta)$ and $R(\delta_2, \theta)$ with respect to a loss function $L(\delta, \theta)$. The relative efficiency, $r(\delta_1,\delta_2)$ of δ_1 to δ_2 is the inverse ratio of their risks, i.e.,

$$r(\delta_1, \delta_2) = \frac{R(\delta_2, \theta)}{R(\delta_1, \theta)}$$

If $r(\delta_1, \delta_2) > 1$, then δ_1 is considered a better estimator than δ_2 and vice versa. However, it should be noted that $r(\delta_1, \delta_2)$ is a function of θ; therefore, $r(\delta_1, \delta_2)$ can vary for different θ. That is, δ_1 can be better than δ_2 for certain θ and the opposite true for other values of θ.

Estimator Consistency

One fundamental property of a good estimator is that as sample size increases the risk should decrease and in the limit the risk should approach zero as sample size m increases indefinitely; i.e.,

$$\lim_{m \to \infty} R_m(\delta, \theta) = 0 \quad \text{for all } \theta$$

An alternate statement of estimator consistency is

$$\lim_{m \to \infty} \text{PROB}\left[\theta - \Delta < \delta_m < \theta + \Delta\right] = 1 \qquad \text{for all } \theta$$

for Δ arbitrarily small. This statement is referred to as simple consistency. If the loss function is squared error, $L(\epsilon) = \epsilon^2$, then $R_m(\delta, \theta) = \overline{\epsilon^2}_m$, and the estimator, δ, is termed squared-error consistent if

$$\lim_{m \to \infty} \overline{\epsilon^2}_m = 0 \quad \text{for all } \theta$$

Since $\overline{\epsilon^2} = V(\delta) + (\theta - \bar{\delta})^2$, squared-error consistency implies that both the variance and bias of δ approach zero for indefinitely large sample size. Estimator consistencies, in the previous context, concern behavior for indefinitely large samples. However, estimator consistence for finite sample size is equally important. This fact leads to the definition of a uniformly consistent estimator as one for which the risk decreases uniformly for increases in sample size, i.e., if $R_{m+1}(\delta, \theta) < R_m(\delta, \theta)$, for all θ, then δ is uniformly consistent.

2.3.4.6 Sufficient Statistics

In general, an estimator δ is a random variable which is used to determine estimations of a particular θ from a random sample. The underlying principle is that the statistical regularity of a random process, which is characterized by the parameter θ, is demonstrated in a random sample of the process; thus, information concerning θ resides in the random sample. The purpose of the estimator is to extract this available information from the random sample. It should be apparent from the foregoing sections that various criteria do not directly measure the degree of utilization of the information available in the random sample. Of course, it is desirable that all of the available information is utilized by the estimator for θ. It could be strongly argued that estimators which fulfill certain of the criteria discussed

must utilize all available information; however, such arguments are largely hueristic, and some question remains. Fortunately, this question can be satisfactorily resolved by considering sufficient statistics.

The basic argument is that in order to obtain information about Θ sample data must be taken from a process which is characterized by Θ. Sample data taken from a process which is independent of Θ are useless in determining Θ. Thus, if the pdf of a random sample is a function of Θ, then the random sample is useful in determining Θ; further, an additional random sample whose pdf is independent of Θ provides no additional information for determining Θ. By the same argument, if the conditional pdf of a set of data Y_1, given a set of data Y_2, is not a function of Θ, then the data Y_1 provides no additional information of Θ. That is, the data Y_1, given the data Y_2, provides no additional information about Θ if $f(Y_1/Y_2)$ is independent of Θ. This leads to the following definition for a sufficient statistic.

Definition: (Sufficient Statistic)

Let Y be a random sample from a process with pdf dependent on Θ, and let T_S be a statistic of Y (a known function of Y). Let T be any other statistic of Y which is independent of T_S. If, for each T, the conditional pdf of T, given T_S, is independent of Θ, then T_S is a sufficient statistic for Θ.

Thus, it follows by the definition of a sufficient statistic, that the information available from the sufficient statistic cannot be increased by a statistic whose conditional pdf, given the sufficient statistic, is independent of Θ. That is, since $\partial/\partial\theta f(T/T_S) = 0$, no additional information is obtained from T, given T_S.

The final conclusion is that estimators which utilize all available information in a random sample should be functions of sufficient statistics. (This is the reason for expressing estimators as a function of statistics in previous sections.) Indeed, estimators are desired which are actually sufficient statistics. It becomes apparent that if sufficient statistics can be determined, then they should contain the desired estimators. Fortunately, it is not too difficult to determine sufficient statistics by making use of the following theorem.

If T_S is a sufficient statistic of the random sample Y, then the joint pdf of Y can be factored as follows:

$$f(Y,\theta) = h(T_S, \theta)K(Y)$$

where K(Y) is independent of Θ. The theorem can be proved by the method of contradiction in the following manner. From the set of sample data Y of m elements, construct the set T(Y) of m independent statistics. Let each member of the statistics set have a single-valued inverse (i.e., $Y = T^{-1}(Y)$) and let J be the Jacobian of $T^{-1}(Y)$. Now let f(T) = joint pdf of T and f(Y) = joint pdf of the set Y; then $f(T) = f[T^{-1}(Y)]\,J$.

Assume that T_S is not a sufficient statistic and that $f(Y) = h(T_S, \Theta)K(Y)$. Let T_i be any other statistic in the set T. It follows that

$$f(T_S) = \int_{(T-T_S)} f(T)\,dT = \int_{(T-T_S)} f\left[T^{-1}(Y)\right] J\,dT$$

$$f(T_i, T_S) = \int_{[T-(T_i+T_S)]} f\left[T^{-1}(Y)\right] J\,dT$$

and

$$f(T_i/T_s) = \frac{f(T_i, T_s)}{f(T_s)}$$

Now, if

$$f(Y,\theta) = h(T_s,\theta)\,K(Y)$$

then,

$$\begin{aligned} f(Y) &= \int_\theta f(Y,\theta)\,d\theta \\ &= K(Y)\int_\theta h(T_s,\theta)\,d\theta \\ &= K(Y)\,H(T_s,\theta) \end{aligned}$$

it follows that,

$$f(T_i/T_s) = \frac{\int_{T-T_s} H(T_s,\theta)K(T)J\,dT}{\int_{T-(T_i+T_s)} H(T_s,\theta)K(T)J\,dT}$$

$$= \frac{H(T_s,\theta)\,J}{H(T_s,\theta)\,J}\,\frac{\int_{T-T_s} K(Y)\,dT}{\int_{T-(T_i+T_s)} K(Y)\,dt}$$

$$f(T_i/T_s) = \frac{\int_{T-T_s} K(Y)\,dt}{\int_{T(T_i+T_s)} K(Y)\,dt}$$

Therefore, f (T_i/T_S) is independent of Θ. This observation contradicts the assumption that T_S is not a sufficient statistic. Therefore, T_S is a sufficient statistic if f $(Y) = h\,(T_S, \Theta)\,K(Y)$. Thus, it is often possible to determine sufficient statistics by inspection from the joint pdf of the random sample set Y.

2.3.5 Determination of Estimators

In the previous section, various criteria are discussed which can be used as measures of good estimators. These criteria can be generally classified as minimum variance and minimum risk. In this section, particular attention is given to methods of determining estimators of minimum variance and risk.

2.3.5.1 Minimum Variance Estimators

2.3.5.1.1 Via Sufficient Statistics and Complete pdf's

In Section 2.3.4.6 it was concluded that, in order to utilize all available information in a random sample, an estimator should be a function of a sufficient statistic. As a corollary, it should follow that estimators which are functions of sufficient statistics possess certain criteria of good estimators. Indeed, it can be shown that, generally, an unbiased estimator which is a function of a sufficient statistic possesses smaller variance than an unbiased estimator which is not a function of a sufficient statistic. That is, let T_S be a sufficient statistic for Θ and let δ_1 be an unbiased estimator of Θ which is not a function of a sufficient statistic. Then, it can be shown that there exists some function of T_S, $\delta(T_S)$, which is an unbiased estimator of Θ and which possesses smaller variance than the estimator δ_1. Therefore, in order to determine minimum variance unbiased estimators, it is only necessary to consider estimators which are functions of sufficient statistics. This is shown in the proof of the following theorem.

THEOREM - Let Y be a random sample from a process with pdf dependent on Θ and let T_S be a sufficient statistic for Θ. Let another statistic T be an unbiased estimator of Θ; i.e., $E(T) = \Theta$. Then

(a) $E(T/T_S)$ is independent of Θ and is a statistic.

(b) $E\left[E(T/T_S)\right] = \Theta$

(c) $V\left[E(T/T_S)\right] < V(T)$.

Before presenting the proof for this theorem, a few comments concerning its meaning are in order. Part (a) states that the conditional expectational of T, given T_S, is a statistic for Θ. Part (b) states that this statistic is also an unbiased estimator for Θ, and Part (c) states that the variance for this estimator is smaller than that of T. Thus, for the criterion of minimum variance, the latter estimator is superior to the first. The proofs are as follows. Part (a): That $E(T/T_S)$ is independent of Θ follows directly from the definition of a sufficient statistic. That is, if T_S is a sufficient statistic, then, by definition, the conditional pdf of T, given T_S, is independent of Θ. Moreover, $f(T/T_S)$ is a function of the random sample Y and the conditional expectational $E(T/T_S)$ is a known function of Y; thus, it is a statistic. Part (b): Since T is an unbiased estimator of Θ; i.e.,

$E(T) = \Theta$, it follows that

$$E(T) = \int_T T f(T;\Theta)\, dT = \Theta$$

$$= \iint_{T_S} T f(T, T_S;\Theta)\, dT\, dT_S$$

Now, since T_S is a sufficient statistic

$$f(T, T_S;\Theta) = f(T/T_S)\, h(T_S;\Theta)$$

where $h(T_S;\Theta)$ is the marginal pdf of T_S. Thus,

$$E(T) = \int_{T_S}\int_T T f(T/T_S) h(T_S;\Theta)\, dT\, dT_S$$

$$= \int_{T_S} \left[\int_T f(T/T_S)\, dT\right] h(T_S;\Theta)\, dT_S$$

$$= \int_{T_S} \left[E(T/T_S)\right] h(T_S;\Theta)\, dT_S$$

$$E(T) = E\left[E(T/T_S)\right] = \Theta$$

Therefore, the conditional expectation of T, given T_S, is an unbiased estimator of Θ if T is an unbiased estimator of Θ. Part (c): To prove this part, let δ denote the unbiased estimator of Θ which is $E(T/T_S)$; i.e., $\delta = E(T/T_S)$.

$$V(T) = E\left[(T-\Theta)^2\right] = E\left[(T-\delta) + (\delta-\Theta)\right]^2$$

$$= E\left[(\delta-\Theta)^2\right] + E\left[(T-\delta)^2\right] + 2E\left[(T-\delta)(\delta-\Theta)\right]$$

It can be shown that the last term is zero, to wit,

$$E\left[(T-\delta)(\delta-\theta)\right] = \int_{T_s}\int_{T}(T-\delta)(\delta-\theta)f(T,T_s;\theta)\,dT\,dT_s$$

$$= \int_{T_s}\left[\int_{T}(T-\delta)f(T/T_s)\,dT\right](\delta-\theta)\,h(T_s;\theta)\,dT_s$$

However, the term in brackets is zero since $\delta = E(T/T_s)$, thus $E\left[(T-\delta)(\delta-\theta)\right] = 0$ and

$$V(T) = E\left[(\delta-\theta)^2\right] + E\left[(T-\delta)^2\right]$$

$$V(T) = V(\delta) + E\left[(T-\delta)^2\right]$$

Since $T \not\equiv \delta$, the second term is always positive, and it follows that

$$V(T) > V(\delta)$$

This theorem establishes the following: The set S of unbiased estimators can be divided into two subsets, S_1 and S_2, which contain, respectively, estimators which are functions of sufficient statistics and those which are not. Estimators which are minimum variance unbiased estimators are found in S_1, only. Of course, there can exist several estimators in S_1, and a problem exists concerning the determination of the estimator with minimum variance in S_1. However, if the pdf of the sufficient statistic is complete (see Section 2.3.2.6), then the unbiased estimator is unique, and the set S_1 contains a single element which is the minimum variance unbiased estimator. This can be seen from part (b) of the previous theorem (where $E(T/T_s)$ is shown to be an unbiased estimator of θ). The expectation of $E(T/T_s)$ with respect to T_s is

$$E\left[E(T/T_s)\right] = \int_{T_s} E(T/T_s)\,h(T_s;\theta)\,dT_s = \theta \quad .$$

If the pdf of T_s, $h(T_s;\theta)$, is a complete pdf, then $E(T/T_s)$ is the unique unbiased estimator of θ, which is a function of the sufficient statistic T_s. And, by part (c) above, $E(T/T_s)$ is the unique minimum variance unbiased estimator of θ.

The foregoing will be illustrated with the following case. Let the random process of Section 2.3.3 be the particular case of $y = \theta + e$, where e is Gaussian with zero mean and variance 1. Let an independent random sample of size m be taken for estimating θ. Thus, the joint pdf of Y becomes

$$f(Y) = (2\pi)^{-\frac{m}{2}} \exp\left[-\frac{1}{2}\sum_{i=1}^{m}(y_i-\theta)^2\right]$$

$$= (2\pi)^{-\frac{m}{2}} \exp\left[-\frac{1}{2}\left(\sum_{i=1}^{m} y_i^2 - 2T\theta + m\theta^2\right)\right]$$

where T is a statistic, i.e.,

$$T = \sum_{i=1}^{m} y_i$$

Now T can be shown to be sufficient since $f(Y) = h(T_s, \theta)\, K(Y)$ (see Section 2.3.2.5); to wit,

$$f(Y) = (2\pi)^{-\frac{m}{2}} \exp\left[+\left(T - \frac{m}{2}\theta\right)\theta\right] \exp\left[-\frac{1}{2}\sum_{i=1}^{m} y_i^2\right]$$

It immediately follows that T is a sufficient statistic T_s for θ. The pdf of T_s is also complete, since it is Gaussian with mean value $m\theta$ and variance m. Therefore, it follows that the unbiased estimator of θ, which is a function of $T_s = \sum_{i=1}^{m} y_i$, is the minimum variance unbiased estimator of θ. Of course, $\delta = 1/m\, T_s$ is this estimator since $E(\delta) = \theta$ and is a function of T_s.

2.3.5.1.2 Minimum Variance Via Least Squares

The method of least squares is one of the most common methods of estimation providing linear estimations of parameters for a linear system (by minimizing the sum of residuals). If the random sample set is independent, then the linear least squares estimations become linear minimum variance estimators. This can be shown as follows:

Let $\underline{y} = A\,\underline{\theta} + \underline{e}$ where an estimator δ for $\underline{\theta}$ is desired, given $\underline{y}$. The least squares estimator $\underline{\delta}_{LS}$ minimizes the scalar product of the residual vector $\underline{v} = \underline{y} - A\,\underline{\delta}_{LS}$; i.e.,

$$\frac{\partial}{\partial \underline{\delta}_{LS}}\left(\underline{v}^T \underline{v}\right) = \frac{\partial}{\partial \underline{\delta}_{LS}}\left(\underline{y} - A\,\underline{\delta}_{LS}\right)^T\left(\underline{y} - A\,\underline{\delta}_{LS}\right) = 0$$

where the superscript T denotes transpose. Alternately,

$$A^T\left(\underline{y} - A\,\underline{\delta}_{LS}\right) = 0$$

$$A^T A\,\underline{\delta}_{LS} = A^T \underline{y}$$

Thus, the least squares estimator is

$$\underline{\delta}_{LS} = (A^T A)^{-1} A^T \underline{y} = Q \underline{y}$$

where

$$Q = (A^T A)^{-1} A^T .$$

The expected value of δ_{LS} is

$$E(\underline{\delta}_{LS}) = QA\underline{\theta} + QE(\underline{e})$$

$$= (A^T A)^{-1} A^T A \underline{\theta} + QE(\underline{e})$$

$$E(\underline{\delta}_{LS}) = \underline{\theta} + QE(\underline{e}) = \bar{\underline{\delta}}_{LS} .$$

Thus, if the error vector $\underline{e}$ is unbiased, then the least squares estimator $\underline{\delta}_{LS}$ is an unbiased estimator of $\underline{\theta}$. The covariance matrix of $\underline{\delta}_{LS}$ is given by

$$COV(\delta_{LS}) = (\underline{\delta}_{LS} - \bar{\underline{\delta}}_{LS})(\delta_{LS} - \bar{\underline{\delta}}_{LS})^T$$

$$(\delta_{LS} - \bar{\delta}_{LS}) = (Q\underline{y} - \underline{\theta} - Q\bar{\underline{e}})$$

$$= Q(\underline{e} - \bar{\underline{e}}) .$$

Thus,

$$COV(\underline{\delta}_{LS}) = Q \, COV(\underline{e}) \, Q^T$$

If the random sample is independent, then COV $(\underline{e}) = \sigma^2 I$ where σ^2 is the variance of the vector $\underline{e}$. In this case, COV $(\underline{\delta}_{LS}) = \sigma^2 Q Q^T = \sigma^2 (A^T A)^{-1}$. Consider another estimator $\underline{\delta}_{MV}$ which has minimum variance and is not the least squares estimator $\underline{\delta}_{LS}$; i.e., let

$$\underline{\delta}_{MV} = C\underline{y}$$

$$= CA\underline{\theta} + C\underline{e}$$

where $CA = I$ and C is the linear minimum variance estimator for θ. Thus, COV $(\underline{\delta}_{MV}) = \sigma^2 CC^T$. Now consider the difference of Q and C as follows:

$$(C-Q)(C-Q)^T = CC^T - QC^T - CQ^T + QQ^T$$

$$= CC^T - (A^TA)^{-1}A^TC^T - CA(A^TA)^{-1} + QA(A^TA)^{-1}$$

$$= CC^T - (A^TA)^{-1} - (A^TA)^{-1} + (A^TA)^{-1}$$

$$= CC^T - (A^TA)^{-1}$$

Thus,

$$CC^T = (A^TA)^{-1} + (C-Q)(C-Q)^T$$

and

$$COV(\underline{\delta}_{MV}) = \sigma^2 CC^T = \sigma^2(A^TA)^{-1} + \sigma^2(C-Q)(C-Q)^T$$

$$= COV(\underline{\delta}_{LS}) + \sigma^2(C-Q)(C-Q)^T$$

Now, in order to minimize the variance of $\underline{\delta}_{MV}$, which is the trace of COV ($\underline{\delta}_{MV}$), it is necessary to take C = Q, therefore,

$$\underline{\delta}_{MV} = \underline{\delta}_{LS}$$

Thus, if the random sample is independent, the least squares estimator is a minimum variance estimator; also, if the vector $\underline{e}$ is unbiased, then the least squares estimator is a minimum variance unbiased estimator.

2.3.5.1.3 A Lower Bound for Estimator Variance

In the previous sections it was tacitly assumed that unbiased estimators are desirable, which is not necessarily a valid assumption; therefore, it is informative to investigate the effect of estimator bias on a general basis, if possible. This investigation is possible through an inequality

of estimator variance as a function of estimator bias. Let δ be an estimator for θ which has the bias $b(\theta)$; i.e.,

$$E(\delta) = \theta + b(\theta)$$

If $\delta(Y)$ is a random variable of an independent random sample Y, then it can be shown that

$$V(\delta,\theta) \geq \frac{\left[1 + \frac{\partial}{\partial\theta} b(\theta)\right]^2}{m E\left[\frac{\partial}{\partial\theta} \log f(y/\theta)\right]^2}$$

where $f(y/\theta)$ is the pdf of the random process and m is the number of elements in Y. The estimator variance inequality follows from taking the partial derivative of $E(\delta)$ with respect to θ, to wit,

$$E(\delta) = \theta + b(\theta)$$

$$= \int_Y \delta(Y) f(Y/\theta)\, dY$$

$$\frac{\partial E(\delta)}{\partial \theta} = 1 + \frac{\partial}{\partial\theta} b(\theta)$$

$$= \frac{\partial}{\partial\theta} \int_Y \delta(Y) f(Y/\theta)\, dY$$

$$\frac{\partial}{\partial\theta} E(\delta) = \int_Y \delta(Y) \frac{\partial}{\partial\theta} f(Y/\theta)\, dY \quad ,$$

assuming that the interchange of order is permissible. Since, for the random sample, Y,

$$f(Y/\theta) = \prod_{i=1}^{m} f(y_i/\theta) \quad ,$$

it follows that

$$\frac{\partial}{\partial\theta} f(Y/\theta) = S(Y) f(Y/\theta)$$

where

$$S(Y) = \sum_{i=1}^{m} \frac{1}{f(y_i/\theta)} \frac{\partial}{\partial \theta} f(y_i/\theta)$$

$$= \sum_{i=1}^{m} \frac{\partial}{\partial \theta} \log f(y_i/\theta) = \sum_{i=1}^{m} \Delta_i(Y) \quad .$$

Thus,

$$1 + \frac{\partial}{\partial \theta} b(\theta) = \int_Y \delta(Y) S(Y) f(Y/\theta)\, dY$$

$$= E(\delta S) \quad .$$

However,

$$E\left[\Delta_i(Y)\right] = E\left[\frac{\partial}{\partial \theta} \log p(y_i/\theta)\right]$$

$$= \int_{y_i} \frac{\partial}{\partial \theta} f(y_i/\theta)\, dy_i$$

$$= \frac{\partial}{\partial \theta}(1) = 0$$

Thus, E(δ, S) is the covariance of δ (Y) and S(Y); it follows that

$$1 + \frac{\partial}{\partial \theta} b(\theta) = cov\left[\delta(Y), S(Y)\right]$$

Using the inequality of covariance, i.e.,

$$\left\{cov\left[\delta(Y), S(Y)\right]\right\}^2 \leq V\left[\delta(Y)\right] V\left[S(Y)\right]$$

it follows that

$$V(\delta) \geq \frac{\left[1 + \frac{\partial}{\partial \theta} b(\theta)\right]^2}{V\left[S(Y)\right]}$$

But, V [S(Y)] is simply the sum of variances for each $s_i(Y)$; thus,

$$V\left[S(Y)\right] = mE\left[\Delta_i^2(Y)\right]$$

$$V\left[S(Y)\right] = mE\left[\frac{\partial}{\partial\theta}\log f(y/\theta)\right]^2 .$$

Therefore,

$$V(\delta,\theta) \geq \frac{\left[1+\frac{\partial}{\partial\theta}b(\theta)\right]^2}{mE\left[\frac{\partial}{\partial\theta}\log f(y/\theta)\right]^2} .$$

For constant estimator bias, independent of θ, the estimator variance bound becomes

$$V(\delta,\theta) \geq \frac{1}{mE\left[\frac{\partial}{\partial\theta}\log f(y/\theta)\right]^2} .$$

It is generally desirable to achieve the lower bound for estimator variance. It is important to note that since the variance bound was derived from the covariance of δ (Y) and S(Y), the lower bound occurs for δ (Y), (δ (Y) is some linear function of S(Y)).

At this point, it is illustrative to check the minimum variance unbiased estimator derived in Section 2.3.5.1.1. The pdf for y was

$$f(y/\theta) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-\theta)^2}$$

hence,

$$\log f(y/\theta) = \left[-\tfrac{1}{2}\log(2\pi) - \tfrac{1}{2}(y-\theta)^2\right]$$

$$\frac{\partial}{\partial\theta}\log f(y/\theta) = y-\theta$$

$$E\left[\frac{\partial}{\partial\theta}\log f(y/\theta)\right]^2 = E(y-\theta)^2$$

$$= V(y) = 1 ,$$

Thus,

$$V(\delta) \geq \frac{1}{m} .$$

Of course, this is the variance of

$$\delta = \frac{1}{m} \sum_{i=1}^{m} y_i$$

for $f(y/\theta)$ with variance 1. It is easy to show that if the variance of is σ^2 instead of 1, then the estimator variance bound becomes

$$V(\delta) \geq \frac{\sigma^2}{m}$$

And the lower bound is again achieved by the minimum variance unbiased estimator

$$\delta = \frac{1}{m} \sum_{i=1}^{m} y_i .$$

2.3.5.2 Minimum Expected Risk Estimators - Bayes

2.3.5.2.1 Bayes Function - A'Posteriori Risk

The Bayes estimator, δ for θ, is the estimator which minimizes the expected value of risk for some loss function. (See Section 2.3.4.5.) That is, a Bayes estimator minimizes $\bar{R}[\delta(Y)]$ where

$$\bar{R}[\delta(Y)] = \int_{\theta}\int_{Y} L[\delta(Y), \theta] f(Y/\theta) f(\theta) dY d\theta$$

$$= \int_{Y} \left[\int_{\theta} L[\delta(Y), \theta] f(Y/\theta) f(\theta) d\theta \right] dY .$$

Of course, the estimator which minimizes the inner integral is the Bayes estimator. The inner integral can be reduced further in the following manner. The produce $f(Y/\theta)\ f(\theta)$ is the joint pdf of Y and θ; i.e., $f(Y, \theta) = f(Y/\theta)\ f(\theta)$. The marginal pdf of Y, $f(y)$, is given by

$$f(Y) = \int_{\theta} f(Y, \theta) d\theta$$

The conditional pdf of Θ, given Y, $f(\Theta/Y)$ becomes

$$f(\theta/y) = \frac{f(Y,\theta)}{f(Y)} = \frac{f(Y/\theta)\,f(\theta)}{f(Y)}$$

Thus, the expected risk $\bar{R}\left[\delta(Y)\right]$ can be written as

$$\bar{R}\left[\delta(Y)\right] = \int_{Y}\left[\int_{\theta} L\left[\delta(Y),\theta\right] f(\theta/Y)\,d\theta\right] f(Y)\,dY$$

$$= \int_{Y} B\left[\delta(Y)\right] f(Y)\,dY$$

where $B\left[\delta(Y)\right]$ is referred to as the Bayes function or the a'posteriori risk. This function is given by

$$B\left[\delta(Y)\right] = \int_{\theta} L\left[\delta(Y),\theta\right] f(\theta/Y)\,d\theta \quad .$$

Thus, a Bayes estimator minimizes the Bayes function or the a'posteriori risk. That is, a Bayes estimator minimizes the risk in estimating Θ given a particular sample set Y. The minimized average risk is the expected value of $B\ [\ \delta(Y)\ ]$ over all Y.

The procedure for determining a Bayes estimator is, generally, that of finding the estimator $\delta(Y)$ which minimizes the integrand of $B\ [\ \delta\ (Y)]$ for all Θ, i.e., the Bayes estimator minimizes the product $L\ [\ \delta\ (Y),\ \Theta]$ $f(\Theta/Y)$ for all Θ. Of course, a Bayes estimator is dependent on the pdf's of the random process and the parameter Θ; hence, a Bayes estimator utilizes the a priori knowledge of Θ. Moreover, a particular Bayes estimator is dependent on the particular loss function used. The determination of a particular Bayes estimator is illustrated in the following example.

EXAMPLE - Bayes Estimator

Consider a scalar random process similar to that considered in Section (2.3.5.1.1); i.e.,

$$y = \theta + e$$

where e is Gaussian with zero mean and variance of unity. Thus,

$$f(y/\theta) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}(y-\theta)^2}$$

Further, let Θ be Gaussian with zero mean and variance σ^2, i.e.,

$$f(\theta) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{\theta^2}{2\sigma^2}}$$

Now, let an independent random sample Y of m elements be used to estimate Θ. The following results are obtained for $f(Y/\Theta)$, $f(Y)$ and $f(\Theta/Y)$.

$$f(Y/\theta) = (2\pi)^{-\frac{m}{2}} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{m}(y_i-\theta)^2\right]$$

$$= (2\pi)^{-\frac{m}{2}} \exp\left[-\tfrac{1}{2}\left(\sum_{i=1}^{m} y_i^2 - 2\theta\sum_{i=1}^{m} y_i + m\theta^2\right)\right]$$

where

$$\equiv C^m \exp\left[-\frac{1}{2}(S - 2\theta + m\theta^2)\right]$$

$$C = (2\pi)^{-1/2}\,,\quad S = \sum_{i=1}^{m} y_i^2 \quad \text{and} \quad T = \sum_{i=1}^{m} y_i$$

thus,

$$f(Y,\theta) = f(Y/\theta)\,f(\theta) = \frac{C^{m+1}}{\sigma} \exp\left\{-\frac{1}{2}\left[S - 2\theta T + (m + \tfrac{1}{\sigma^2})\theta^2\right]\right\}$$

$$= \frac{C^{m+1}}{\sigma} \exp\left[-\frac{1}{2}S\right] \exp\left\{-\frac{1}{2}\left[(m+\tfrac{1}{\sigma^2})\theta^2 - 2\theta\sum_{1}^{m} y_i\right]\right\}$$

$$= \frac{C^{m+1}}{\sigma} \exp\left\{-\frac{1}{2}\left[S - \frac{T^2}{K}\right]\right\} \exp\left\{-\frac{1}{2}(K)\left[\theta - \frac{T}{K}\right]^2\right\}$$

$$K = m + \frac{1}{\sigma^2}$$

Also,

$$f(Y) = \int_{\theta} f(Y,\theta)\,d\theta$$

$$= \frac{cm}{\sigma} \exp\left\{-\frac{1}{2}\left[S - \frac{T}{K}\right]\right\} \exp\left\{-\frac{1}{2}(K)\left[\theta - \frac{T}{K}\right]^2\right\}$$

$$f(Y) = \frac{cm}{\sigma} \exp\left\{-\frac{1}{2}\left[S - \frac{T}{K}\right]\right\} (K)^{-\frac{1}{2}}$$

$$f(\theta/Y) = \frac{f(Y/\theta)}{f(Y)} = \frac{(K)^{\frac{1}{2}}}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(K)\left[\theta - \frac{T}{K}\right]^2\right\}$$

Considering the loss function as squared error, i.e., $L(\delta, \theta) = (\delta - \theta)^2$, $B[\delta(Y)]$ becomes

$$B[\delta(Y)] = \int_{\theta} (\delta-\theta)^2 f(\theta/Y)\,d\theta$$

$$= \frac{(K)^{1/2}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\delta-\theta)^2 \exp\left\{-\frac{1}{2}(K)\left[\theta - \frac{T}{K}\right]\right\} d\theta .$$

Expanding and integrating,

$$B[\delta(Y)] = \delta^2 - \frac{2\delta T}{K} + \frac{1}{K} + \left(\frac{T}{K}\right)^2$$

Using the calculus of variations,

$$\frac{\partial B}{\partial \delta} = 2\delta - 2\frac{T}{K} = 0 \quad .$$

Thus, the Bayes estimator δ_B is

$$\delta_B = \frac{T}{K} \equiv \frac{1}{m + \frac{1}{\sigma^2}} \sum_{i=1}^{m} y_i$$

$$\delta_B = \frac{\sigma^2}{1 + m\sigma^2} \sum_{i=1}^{m} y_i \quad ,$$

This estimator is seen to be a function of the sufficient statistic T. Further, the Bayes estimator is seen to be different from the minimum variance unbiased estimator derived in Section 2.3.5.1.1. The two estimators are listed for comparison.

$$\delta_B = \frac{\tau^2}{1 + m\tau^2} \sum_{i=1}^{m} y_i \qquad \text{Bayes}$$

$$\delta_M = \frac{1}{m} \sum_{i=1}^{m} y_i \qquad \text{minimum variance unbiased}$$

It is seen that the two estimators can be significantly different for small samples, though for indefinitely large samples they are approximately equal. Strictly speaking, if $m\sigma^2 >> 1$, then $\delta_B \simeq \delta_m$. In general,

$$\delta_B = \frac{m\sigma^2}{1 + m\sigma^2} \delta_M$$

Thus,

$$\delta_B < \delta_m$$

For the case where $m\sigma^2 = 1$, $\delta_B = \frac{1}{2} \delta_m$. It should not be surprising that the Bayes estimator tends to be smaller than the minimum variance unbiased estimator since use has been made of the pdf of θ in deriving the Bayes estimator.

It should be noted that the Bayes estimator is a minimum mean-squared error estimator since a squared error loss function was used. It is also noted that the Bayes estimator is biased since the conditional expectational of δ_B is not θ, i.e.,

$$E(\delta_B/\theta) = \left(\frac{m\sigma^2}{1 + m\sigma^2}\right)\theta$$

Finally, it should be noted that for $\sigma^2 \longrightarrow \infty$, i.e., a nearly uniform distribution for θ, $\delta_B = \delta_M$ regardless of the sample size m.

2.3.5.2.2 Minimum Mean-Squared-Error

In the previous section a Bayes estimator was derived for a particular case by minimizing the a'posteriori risk for a squared-error loss function. This procedure determined the minimum mean-squared error estimator. It would be desirable to determine the Bayes estimator in a more general form. Fortunately, it is possible to determine a more general solution to the minimum mean-squared-error estimator which precludes some of the detailed steps involved in the method used previously.

Consider the Bayes function for the squared-error loss function $L(\delta(Y), \theta) = [\delta(Y) - \theta]^2$; to wit,

$$B\left[\delta(Y)\right] = \int_{\theta}\left[\delta(Y) - \theta\right]^2 f(\theta/Y)\, d\theta$$

Expanding the integrand, B [δ (Y)] becomes

$$B\left[\delta(Y)\right] = \int_{\theta}\left[\delta^2(Y) - 2\delta(Y)\theta + \theta^2\right] f(\theta/Y)\, d\theta$$

$$B\left[\delta(Y)\right] = \delta^2(Y)\int_{\theta} f(\theta/Y)\, d\theta - 2\delta(Y)\int_{\theta}\theta\, f(\theta/Y)\, d\theta + \int_{\theta}\theta^2 f(\theta/Y)\, d\theta$$

Now, differentiating with respect to δ (Y) and equating the result to zero yields

$$\frac{\partial B\left[\delta(Y)\right]}{\partial\, \delta(Y)} = \delta(Y)\int_{\theta} f(\theta/Y)\, d\theta - \int_{\theta}\theta\, f(\theta/Y)\, d\theta = 0$$

Thus, the minimum mean-squared error estimator is given by

$$\delta(Y) = \frac{\int_{\theta}\theta\, f(\theta/Y)\, d\theta}{\int_{\theta} f(\theta/Y)\, d\theta} \equiv \int_{\theta}\theta\, f(\theta/Y)\, d\theta$$

However, the right-hand member is precisely the conditional expectation of Θ, given Y; i.e.,

$$\delta(Y) = E(\theta/Y)\ .$$

This form for the Bayes estimator makes it rather convenient to determine the Bayes estimator from the conditional pdf of Θ, given Y, if its mean value can be recognized.

Consider the particular case of the previous section wherein

$$f(\theta/Y) = \frac{K^{1/2}}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(K)\left[\theta - \frac{T}{K}\right]^2\right\}$$

$$f(\theta/Y) = \frac{1}{\sqrt{2\pi\, V(\theta/Y)}} \exp\left\{-\frac{1}{2\,V(\theta/Y)}\left[\theta - E(\theta/Y)\right]\right\}$$

where

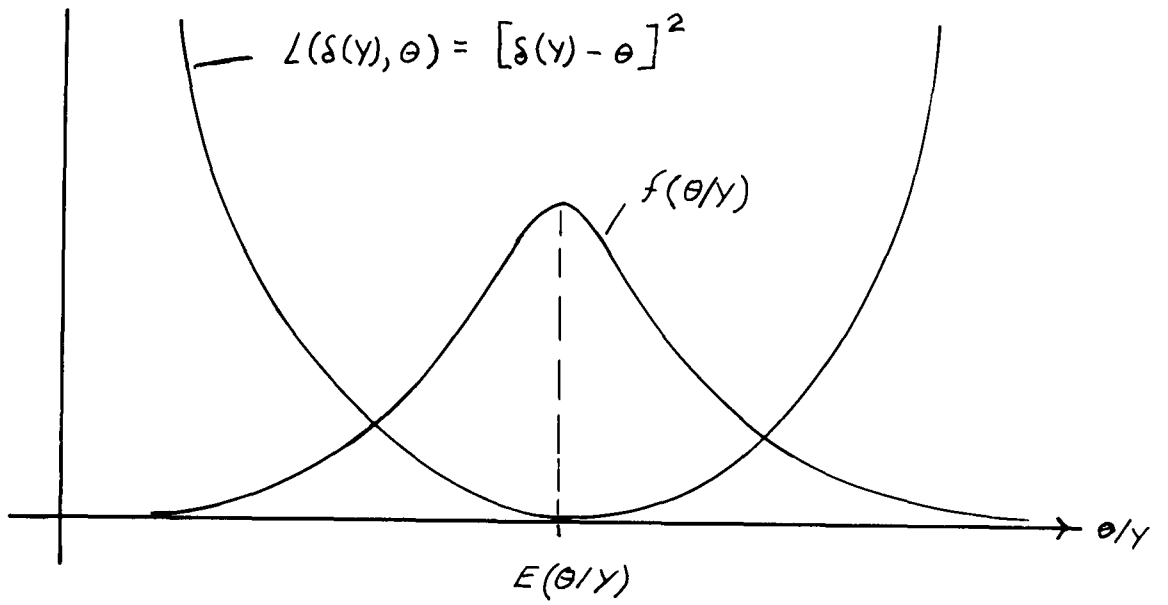
$$V(\theta/Y) = K = \frac{\sigma^2}{1 + m\sigma^2}$$

$$E(\theta/Y) = \frac{\sigma^2}{1 + m\sigma^2} \sum_{i=1}^{m} y_i \quad .$$

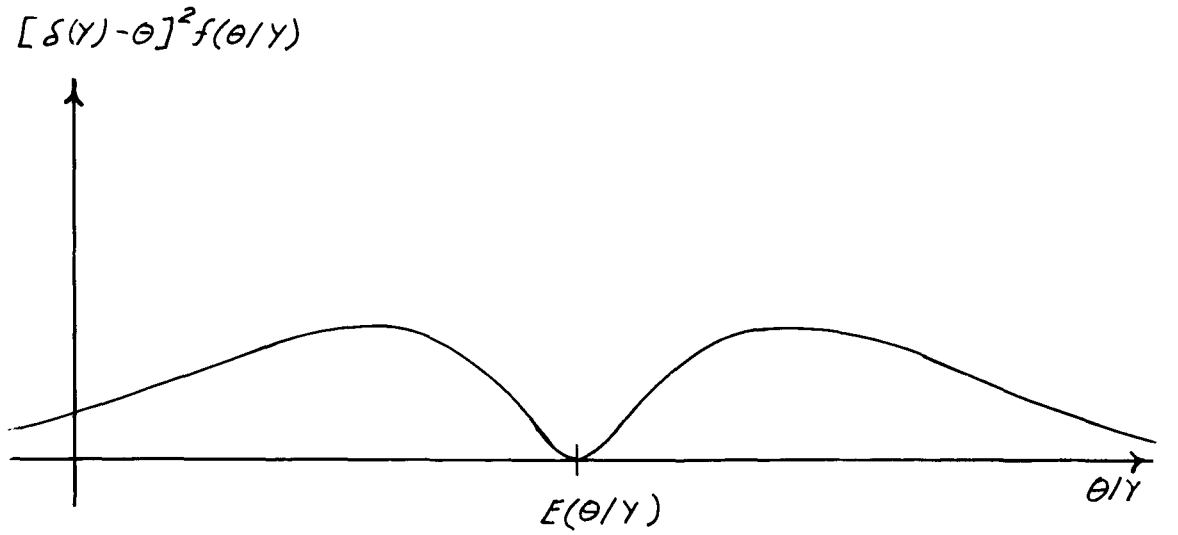
Thus, $f(\Theta/Y)$ is Gaussian with conditional variance $V(\Theta/Y)$ and conditional mean $E(\Theta/Y)$. Of course, $E(\Theta/Y)$ is the Bayes estimator as derived previously.

It is informative to consider the Bayes estimator on a geometrical basis. For the case considered above, $L\ [\delta\ (Y), \Theta]$ and $f(\Theta/Y)$ are shown in Figure 2.3.4(a), and their product, which is the integrand of $B\ [\delta\ (Y)]$, is shown in Figure 2.3.4 (b). For this case it is seen that the Bayes estimator places the minimum of $L\ [\delta\ (Y), \Theta]$ in coincidence with the maximum of $f(\Theta/Y)$.

In general, the Bayes estimator minimizes the area under the product of $L\ [\delta\ (Y), \Theta]$ and $f(\Theta/Y)$, since $B\ [\delta\ (Y)]$ is a minimum. It is seen that for squared-error loss the minimum area occurs for the minimum of the loss function being to coincidence with the conditional expectation of Θ, given Y, or the mean of $f(\Theta/Y)$. However, this occurrence does not necessarily have to be coincident with the maximum of $f(\Theta/Y)$, as it does in the Gaussian case.



(a) Squared-error loss and the Gaussian pdf



(b) Integrand of Bayes a'posteriori risk

2.3.5.2.3 Bayes Estimator for Convex Loss Functions

In the previous section, it was shown that the conditional expectation of Θ, given Y, is the Bayes estimator for the squared-error loss function. Of course, other loss functions are possible. Thus, some question arises concerning the Bayes estimator for some arbitrary loss function. It could be conjectured that the Bayes estimator for squared-error loss should possess desirable properties for loss functions similar to the squared-error loss function. Indeed, it can be shown that the squared-error loss function is a single member of a set of loss functions which has the same Bayes estimator.

Consider the set of loss functions which are defined as follows:

$$L(\epsilon) = L_o \quad \text{for } \epsilon = 0$$

$$L(\epsilon_2) = L(\epsilon_1) \quad \text{for } \epsilon_2 = \epsilon_1$$

$$L(\epsilon_2) > L(\epsilon_1) \text{ for } |\epsilon_2| > |\epsilon_1|$$

$$L(\epsilon) = L(-\epsilon)$$

The loss functions are continuous symmetrical convex functions with symmetry about their minimums L_o. Let the conditional pdf $f(\Theta/Y)$ be factorable into a function of Y, $f_1(Y)$, and a symmetrical function of Θ and Y such that

$$f(\Theta/Y) = f_1(Y)\, f_2[\Theta - g(Y)]$$

where

$$f_2[\theta - g(Y)] = f_2[g(Y) - \theta]$$

Now the a'posteriori risk becomes

$$B[\delta(Y)] = \int_\theta L(\delta - \theta)\, f_1(Y)\, f_2[\theta - g(Y)]\, d\theta$$

Taking the partial of B $[\delta(Y)]$ with respect to $\delta(Y)$ yields

$$\frac{\partial B[\delta(Y)]}{\partial \delta(Y)} = f_1(Y) \int_\theta f_2[\theta - g(Y)] \left[\frac{\partial L}{\partial \delta}(\delta - \theta)\right] d\theta$$

Since $L(\delta - \Theta)$ is an even function of δ, its derivatives is an odd function; thus, it is possible to set the first partial of B $[\delta(Y)]$ identically to zero by setting $\delta(Y) = g(Y)$, i.e.,

$$\int_{\theta} f_2[\theta - g(Y)]\left\{\frac{\partial}{\partial \delta} L[g(Y) - \theta]\right\} d\theta \equiv 0$$

Therefore, the extremum of B [δ (Y)] occurs for

$$\delta(Y) = g(Y)$$

If B [δ (Y)] has a unique extreme, then, of course, δ (Y) = g(Y) is the Bayes estimator. Thus, if f_2 [Θ - g(Y)] is unimodal, i.e., has a single maximum value at Θ = g(Y), then δ (Y) = g(Y) is the Bayes estimator for all convex functions L(δ - Θ) defined above. Of course, for f(Θ/Y) as defined, g(Y) = E(Θ/Y).

2.3.5.2.4 Determination of Bayes Risk

A Bayes estimator minimizes the a'posteriori risk B [δ (Y)], given a random sample Y. This, in turn, minimizes the expected risk R [δ (Y)] over all Y. The minimum expected risk for a Bayes estimator is referred to as the Bayes risk. The Bayes risk can be determined from the expected value of the a'posteriori risk as a function of the Bayes estimator, to wit,

$$\bar{R}_B = \int_Y B(\delta_B) f(Y) dY$$

where $\bar{R}_B$ is the Bayes risk and δ_B is a Bayes estimator for some loss function. Of course, $\bar{R}_B$ is dependent on the loss function.

This equation provides the general means of determining the Bayes risk; however, there are two cases of particular interest. The first case is that for constant a'posteriori risk; i.e., quite often B (δ_B) is independent of the random sample Y. In this case, the expected value of B(δ_B) is simply B(δ_B); i.e.,

$$\bar{R}_B = \int_Y B(\delta_B) f(Y) dY$$

$$= B(\delta_B) \int_Y f(Y) dY$$

$$\bar{R}_B = B(\delta_B)$$

Thus, for a'posteriori risk (independent of Y) the Bayes risk is simply the a'posteriori risk B(δ_B).

The second case is that of a squared-error loss function. In this case, the Bayes risk can be expressed in terms of the conditional variance

of Θ, given Y. That is,

$$\bar{R}_B = \iint_{Y\,\Theta} \left[\delta_B(Y) - \theta\right]^2 f(\theta/Y) f(Y)\, d\theta\, dY$$

However, for squared-error loss, the Bayes estimator is $E(\Theta/Y)$ (see Section 2.3.5.2.2), and the a'posteriori risk $B[\delta_B(Y)]$ is the variance of Θ, given Y; to wit,

$$B\left[\delta_B(Y)\right] = \int_{\theta} \left[\theta - E(\theta/Y)\right]^2 f(\theta/Y)\, d\theta$$

$$B\left[\delta_B(Y)\right] = V(\theta/Y)$$

It follows that

$$\bar{R}_B = \int_{Y} V(\theta/Y)\, f(Y)\, dY$$

If $V(\Theta/Y)$ is independent of Y, then, as in the first case,

$$\bar{R}_B = V(\theta/Y) = B\left[\delta_B(Y)\right]$$

where $V(\Theta/Y)$ is independent of Y.

Consider the case previously discussed wherein $Y = \Theta + e$ with Θ and e Gaussian with zero mean values and variances σ^2 and 1, respectively. The conditional pdf of Θ, given Y, is

$$f(\theta/Y) = \left(\frac{m + \frac{1}{\sigma^2}}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(m + \frac{1}{\sigma^2}\right)\left[\theta - \frac{\sum_{i=1}^{m} y_i}{m + \frac{1}{\sigma^2}}\right]^2\right\}$$

For squared-error loss, the Bayes estimator δ_B is $E(\Theta/Y)$, to wit,

$$\delta_B = E(\theta/Y) = \frac{\sigma^2}{1 + m\sigma^2} \sum_{i=1}^{m} y_i$$

The conditional pdf $f(\Theta/Y)$ is Gaussian with mean $E(\Theta/Y)$ and variance $\sigma^2/(1+m\sigma^2)$ which is the conditional variance of Θ, given Y. The a'posteriori risk becomes

$$B\left[\delta_B(Y)\right] = V(\theta/Y) = \frac{\sigma^2}{1 + m\sigma^2}$$

It is seen that B [$\delta_B(Y)$] is independent of Y, therefore, the Bayes risk for squared-error loss is

$$\bar{R}_B = V(\theta/Y) = \frac{\sigma^2}{1+m\sigma^2}$$

It is seen that

$$\lim_{m\to\infty} \bar{R}_B$$

and

$$\bar{R}_B(m+1) < \bar{R}_B(m) \quad .$$

Therefore, for the problem considered, the Bayes estimator is uniformly consistent. (See Section 2.3.4.5.)

2.3.5.3 A Comparison of Minimum Variance and Bayes Estimators

In the previous sections, two general classes of estimators are discussed. These two classes are: (1) minimum variance and (2) minimum expected risk or Bayes. The class of minimum variance estimator contains the set of minimum variance unbiased estimators while the class of Bayes estimators contains the set of minimum mean-squared-error estimators. In general, these two sets can be disjoint since a minimum mean-squared-error estimator is not necessarily a minimum variance unbiased estimator. That is, biased estimators can exist which possess smaller variance than a minimum variance unbiased estimator. The situation can be clearly stated in the following manner.

In Section 2.3.4.3 it was shown that the mean squared-error for an estimator generally exceeds its variance, i.e.,

$$E(\epsilon^2/\theta) = V(\delta/\theta) + E^2(\epsilon/\theta)$$

thus

$$\overline{\epsilon^2(\delta)} \geq V(\delta)$$

with equality if $E(\delta/\theta = \theta$ for all θ. In Section 2.3.5.1.3 it was shown that a lower bound exists for the variance of an estimator; i.e.,

$$V(\delta/\theta) \geq \frac{\left[\frac{\partial}{\partial\theta} E(\delta/\theta)\right]^2}{mE\left[\frac{\partial}{\partial\theta} \log f(y/\theta)\right]^2}$$

Immediately, a question arises concerning the effect of estimator bias.

It is seen that the lower bound for estimator variance can be smaller for a biased estimator than for an unbiased estimator. That is, for an unbiased estimator, the numerator in the variance bound is unity. However, for a biased estimator, the numerator in the variance bound can be either greater or less than unity for positive or negative bias, respectively. This follows since the denominator of the variance bound is unaltered by a particular estimator and is primarily dependent on the pdf f(Y/Θ) of the random process and the sample size m.

Thus, if it is possible to achieve the variance lower bound with each of two estimators δ_1 and δ_2 which have negative and zero bias, respectively, then the biased estimator δ_1 will possess smaller variance than the unbiased estimator δ_2. Therefore, although the mean squared-error of δ_1 is greater than its variance, it is still possible that the mean squared error of δ_1 is less than that of δ_2, which is unbiased. That is, let $E(\delta_1/\theta) = K\theta$ and $E(\delta_2/\theta) = \theta$, then

$$V(\delta_1/\theta) \geq \frac{K^2}{mD}$$

$$V(\delta_2/\theta) \geq \frac{1}{mD}$$

where

$$D = E\left[\frac{\partial}{\partial\theta} \log f(y/\theta)\right]^2$$

If δ_1 has negative bias $K < 1$ and if δ_1 and δ_2 both achieve their lower variance bounds, then

$$V(\delta_2) > V(\delta_1)$$

Also,

$$\overline{\epsilon^2(\delta_2)} = V(\delta_2)$$

$$\overline{\epsilon^2(\delta_1)} > V(\delta_1)$$

From the foregoing it is entirely possible that

$$\overline{\epsilon^2(\delta_2)} > \overline{\epsilon^2(\delta_1)}$$

Indeed, this can occur, as can be seen in the particular case considered in the previous sections wherein

$$f(y/\theta) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(y-\theta)^2\right]$$

$$f(\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\theta^2}{2\sigma^2}\right]$$

$$D = 1$$

$$\delta_B = \frac{z}{1+m\sigma^2} \sum_{i=1}^{m} y_i$$

$$\delta_M = \frac{1}{m} \sum_{i=1}^{m} y_i$$

where δ_B = Bayes estimator and δ_m = minimum variance unbiased estimator. It is easily shown that

$$E(\delta_B/\theta) = \frac{m\sigma^2}{1+m\sigma^2}\theta$$

$$= K\theta$$

where

$$K = \frac{m\sigma^2}{1+m\sigma^2} = \frac{1}{1+\frac{1}{m\sigma^2}} < 1$$

Therefore, the Bayes estimator has negative bias. The variance bounds are

$$V(\delta_B/\theta) \geq \frac{K^2}{m}$$

$$V(\delta_m/\theta) \geq \frac{1}{m}$$

It was shown in Section 2.3.5.1.3 that δ_m achieves its lower bound. But since δ_m is unbiased, it follows that

$$\overline{\epsilon^2(\delta_m)} = V(\delta_M) = \frac{1}{m}$$

On the other hand, the Bayes risk for δ_B is its mean squared error, i.e.,

$$\overline{\epsilon^2(\delta_B)} = \overline{R_B} = \frac{\sigma^2}{1+m\sigma^2} = \frac{1}{m+\frac{1}{\sigma^2}}.$$

(See Section 2.3.5.2.4.)

It follows that

$$\overline{\epsilon^2(\delta_B)} < \overline{\epsilon^2(\delta_M)} \quad .$$

Thus, the Bayes estimator which has negative bias has smaller mean squared error than the minimum variance unbiased estimator δ_m. Therefore, the variance of δ_B must be smaller than that for δ_m. Indeed, δ_B also achieves its variance lower bound. This can be shown as follows.

$$V(\delta_B/\theta) = E\left\{\left[\delta_B - E(\delta_B/\theta)\right]^2/\theta\right\}$$

$$= E\left\{\left[\frac{1}{m}K\sum_{i=1}^{m} y_i - K\theta\right]^2/\theta\right\}$$

$$= \frac{K^2}{m^2} E\left\{\left[\sum_{i=1}^{m} y_i - m\theta\right]^2/\theta\right\}$$

$$= \frac{K^2}{m^2} E\left\{\left[\left(\sum_{i=1}^{m} y_i\right)^2 - 2m\theta\sum_{i=1}^{m} y_i + m^2\theta^2\right]/\theta\right\}$$

$$= \frac{K^2}{m^2}\left[E\left(\sum_{i=1}^{m} y_i\right)^2 - 2m^2\theta^2 + m^2\theta^2\right]$$

$$= \frac{K^2}{m^2}\left[m(1+m\theta^2) - m^2\theta^2\right]$$

$$V(\delta_B/\theta) = \frac{K^2}{m}$$

Thus, the Bayes estimator achieves its variance lower bound.

The foregoing clearly demonstrates that a biased estimator can possess both variance and mean squared error which are less than those for a minimum variance unbiased estimator!

The foregoing does not categorically settle the question of selecting a good estimator. If mean squared error is an indisputable criterion, then the Bayes estimator is preferential. On the other hand, if a situation exists where bias is highly deleterious, then a minimum variance unbiased estimator would be dictated. However, consideration should also be given to other criteria of evaluation; i.e., consistency, relative efficiency and the effect of sample size.

For the case considered above, both estimators are uniformly squared-error consistent, i.e., the mean squared-error decreases uniformly with sample size m. The relative efficiency, using mean squared-error is

$$r(\delta_B, \delta_M) = \frac{\frac{1}{m}}{\frac{\sigma^2}{1+m\sigma^2}} = \frac{1+m\sigma^2}{m\sigma^2} = 1 + \frac{1}{m\sigma^2}$$

The relative efficiency demonstrates the effect of sample size m. Of course, if sample size is large, such that $m\sigma^2 >> 1$ then there is a small difference in efficiency. However, for small sample size and/or small σ^2 such that $m\sigma^2 < 1$, then a significant difference results.

The foregoing demonstrates the improved performance for small sample size for the Bayes estimator over the minimum variance unbiased estimator. This improvement is derived from the use of the a'priori information in terms of the pdf of Θ in the Bayes estimator (this information tends to bias the estimations of Θ toward its average value) which was zero in the case considered. The improved performance tends to overcome the lack of fidelity in statistical regularity demonstrated in small samples. That is, the statistical regularity of a random process is not faithfully demonstrated in small samples. The utilization of any available information can improve performance significantly. On the other hand, if a large number of samples are available, the statistical regularity is more reliably demonstrated and the use of additional information not as effective or critical.

2.3.5.4 Minimum Risk Estimation

In general, risk is the expected loss in estimating a parameter Θ by an estimator. It is apparent that risk is a measure of estimator performance, which provides a general criterion of minimization in determining estimators. A particular case of interest is that of squared-error loss wherein risk becomes mean squared-error and an estimator of minimum risk is a minimum mean squared-error estimator. Of course, minimum risk estimators for loss functions other than squared error are equally important, and general properties of minimum risk estimators are of considerable interest. There are two general properties of minimum risk estimators which are of particular significance. These properties are: (1) generality of loss function and (2) dependence on sufficient statistics. That is, for a general class of convex loss functions, a minimum risk estimator must be a function of sufficient statistic. These properties are established below.

2.3.5.4.1 Convex Loss Functions

A convex function is illustrated in Figure 2.3.5 and can be described in the following manner. Let $l(x)$ be a line which intersects $L(x)$ at $x = A$ and B.

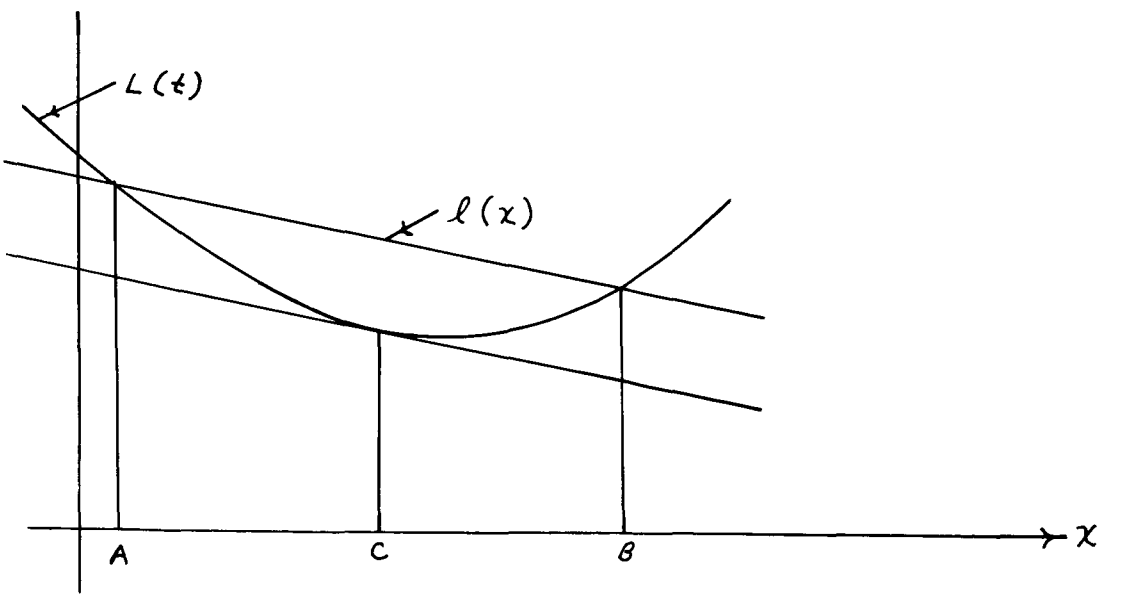


Figure 2.3.5

If, for all A and B, $L(x) < l(x)$ for $A < x < B$ and if, for $A = B = C$, $L(x) > l(x)$ for x, except $x = C$ where $L(x) = l(x)$ then $L(x)$ is a convex function. Alternatively, through any point on a continuous convex function $L(x)$ there passes a line $l_s(x)$ which lies everywhere below $L(x)$. The line $l_s(x)$ is referred to as a "supporting" line for $L(x)$.

An important property of a convex function of a random variable is that its expectation generally exceeds its value at the expectation of the random variable; i.e., if L(x) is a convex function and x is a random variable then

$$E\left[L(X)\right] \geq L\left[E(X)\right]$$

This can be shown as follows: Let $l_s(x)$ be a supporting line of L(x) at the point E(x). Thus

$$L(X) \geq \ell_s(X)$$

and

$$E\left[L(X)\right] \geq E\,\ell_s(X)$$

However, since $l_s(x)$ is linear in x $E\left[\,l_s(x)\,\right] = l_s\left[\,E(x)\right]$ thus,

$$E\left[L(X)\right] \geq \ell_s\left[E(X)\right]$$

Moreover, since $l_s(x)$ supports L(x) at E(x), $L\left[\,E(x)\right] = l_s\left[\,E(x)\right]$ thus,

$$E\left[L(X)\right] \geq L\left[E(X)\right] .$$

As a particular case consider x as estimation error ϵ and $L(\epsilon)$ as squared error loss; i.e.,

$$L(\epsilon) = \epsilon^2$$

It follows that

$$E\left[L(\epsilon)\right] = E(\epsilon^2) \geq \left[E(\epsilon)\right]^2$$

This result agrees with that of Section 2.3.4.3 wherein it was determined that

$$E(\epsilon^2) = V(\epsilon) + E^2(\epsilon)$$

on

$$E(\epsilon^2) \geq E^2(\epsilon)$$

with equality for $V(\epsilon) = 0$ which implies a constant ϵ.

The preceding inequality of convex function expectation is fundamental

to showing the necessity of sufficient statistics for minimum risk estimators, which is considered below.

2.3.5.4.2 Minimum Risk Via Sufficient Statistics

In Section 2.3.5.1.1 it was shown that if T and Ts are an unbiased estimator and a sufficient statistic for Θ, respectively, then $E\,[\,E(T/Ts)]$ is an unbiased estimator for Θ; and the variance of $E(T/Ts)$ is less than the variance of T. These results can be generalized in two respects. First, any function of Θ, $U(\Theta)$, can be considered; and, second, variance can be replaced by risk for convex loss functions. These results will be proved in the following theorem, which is an extremely important extension of the theorem of Section 2.3.5.1.1 concerning sufficient statistics and minimum variance estimators. In effect, this theorem demonstrates the fundamental importance of sufficient statistics in determining estimators.

THEOREM: Let Y be a random sample from a process with pdf dependent on Θ, and let Ts be a sufficient statistic for Θ. Let T be another statistic which is an unbiased estimator for any function of Θ, $U(\Theta)$; i.e., $E(T) = U(\Theta)$. Then

(a) $E(T/Ts)$ is independent of Θ and is a statistic

(b) $E\,[E(T/Ts)] = U(\Theta)$

(c) $R(T, \Theta) \geq R(\delta s, \Theta)$

where the risk is for convex loss functions and $\delta_{LS} = E(T/Ts)$.

Before proving this theorem, two comments are in order. First, it should be noted that biased estimators for Θ are included since an unbiased estimator of $U(\Theta)$ includes a biased estimator of Θ; i.e., the case $U(\Theta) = k\Theta$ where $k \neq 1$ is included. Second, δ_{LS} is an unbiased estimator of $U(\Theta)$, or a biased estimator of Θ, and δ_{LS} is an estimator which has less risk than T for all convex loss functions.

The proofs of parts (a) and (b) of the theorem are direct extensions of the proofs given in Section 2.3.5.1.1. Part (a) is an obvious extension and part (b) follows.

$$E(T) = u(\theta)$$

$$= \int_T T f(T;\theta)\, dT$$

$$= \int_{T_s}\int_T T f(T, T_s;\theta)\, dT\, dT_s$$

$$= \int_{T_s}\int_T T f(T/T_s) h(T_s;\theta)\, dT\, dT_s$$

$$= \int_{T_s} \left[E(T/T_s)\right] h(T_s;\theta)\, dT_s$$

$$= E\left[E(T/T_s)\right] = u(\theta)$$

Part (c) follows from the property of expectation of convex functions as shown in the previous section. Let $L(\delta, \theta)$ be a convex loss function with respect to the unbiased estimator δ for $U(\theta)$; i.e., for fixed θ, $L(\delta, \theta)$ is a convex function of δ. Taking the conditional expectation of $L(\delta, \theta)$, given Ts, it follows that

$$E\left[L(T,\theta)/T_s\right] \geq L\left[E(T/T_s), \theta\right]$$

$$\geq L(\delta_s, \theta)$$

Taking the expectation over all Ts it is found that

$$E\left[L(T,\theta)\right] \geq E\left[L(\delta_s, \theta)\right]$$

Therefore,

$$R(T,\theta) \geq R(\delta_s, \theta)$$

The foregoing is extremely important since it establishes that minimum risk estimators are functions of sufficient statistics and, thus, only sufficient statistics need be considered in determining minimum risk estimators. Furthermore, if the pdf of Ts is complete, then the unbiased estimator δ_s of $U(\theta)$ which is a function of Ts is unique and the estimator of minimum

risk for the class of convex loss functions. These results represent extremely important extensions of the significance of sufficient statistics in determining estimators.

2.3.5.5 Maximum Likelihood Estimators

In general, maximum likelihood estimators are determined by maximizing a pdf. This method of estimation does not explicitly seek to satisfy criteria of estimation, rather, the method is based upon the premise of the most probable occurrence being observed most frequently. In effect, it is assumed that a random sample is always one of high relative probability. Nonetheless, the method is often equivalent to other methods of determining estimators, and similar results are often obtained. The method is discussed herein to show its significant similarities and differences with other methods.

2.3.5.5.1 Principle of Maximum Likelihood

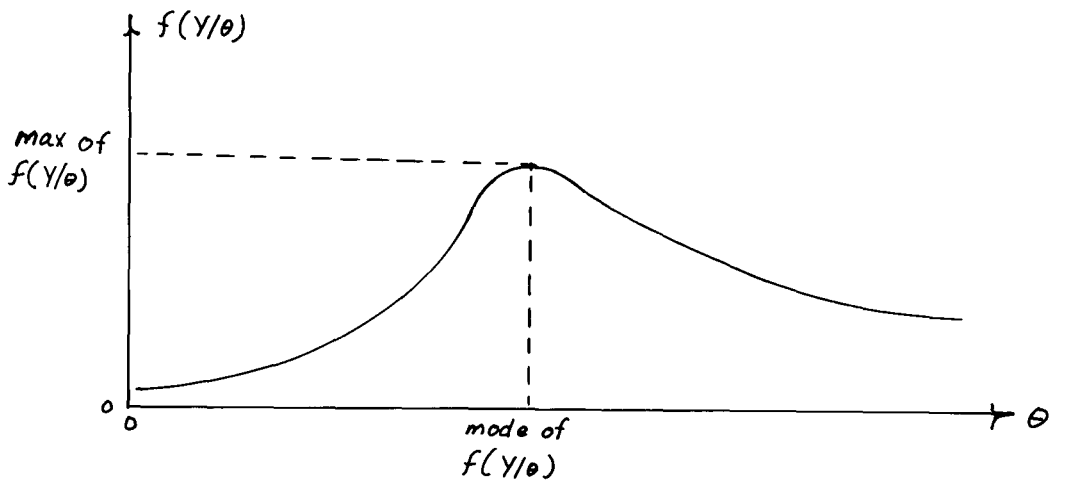
The principle of maximum likelihood is predicated upon a "most likely occurrence" of a set of variables. In general, a pdf is a relative measure of the probability of occurrence of its variable. The particular set of variables for which the pdf is a maximum can be considered as a maximum likelihood set. Any pdf associated with a set of variables is considered to be a likelihood function for the set of variables; e.g., the pdf's $f(\Theta/Y)$ and $f(Y/\Theta)$ are likelihood functions for the variables Θ and Y. If the likelihood function possesses a maximum value, then the particular set of variables for which the maximum value occurs is the set of most likely occurrence or maximum likelihood. This set of variables is referred to as the "mode" of the likelihood function or pdf of the set of variables.

The principle of maximum likelihood leads to the following method of estimation, which is based on maximizing a pdf as a function of a set of parameters Θ.

2.3.5.5.2 Maximum-Likelihood Estimator

Let the likelihood function for a random sample Y be some pdf for Y as a function of a set of parameters Θ. If the likelihood function has a maximum value for Θ (Θ is some function, $\delta(Y)$, of the random sample, Y), then $\delta(Y)$ is the maximum-likelihood estimator of Θ. Thus, the maximum-likelihood estimator for Θ is the one for which the random sample Y occurs with maximum likelihood.

The method of maximum-likelihood estimation is depicted in Figure 2.3.6.



It is easily seen that a maximum likelihood estimator will, in general, be different from a Bayes estimator. However, for a simple loss function the methods yield equivalent results, as shown below.

2.3.5.5.3 Bayes Estimator for a Simple Loss Function

A simple loss function is defined as one which has zero loss for zero estimation errors and constant loss for all non-zero estimation error; to wit,

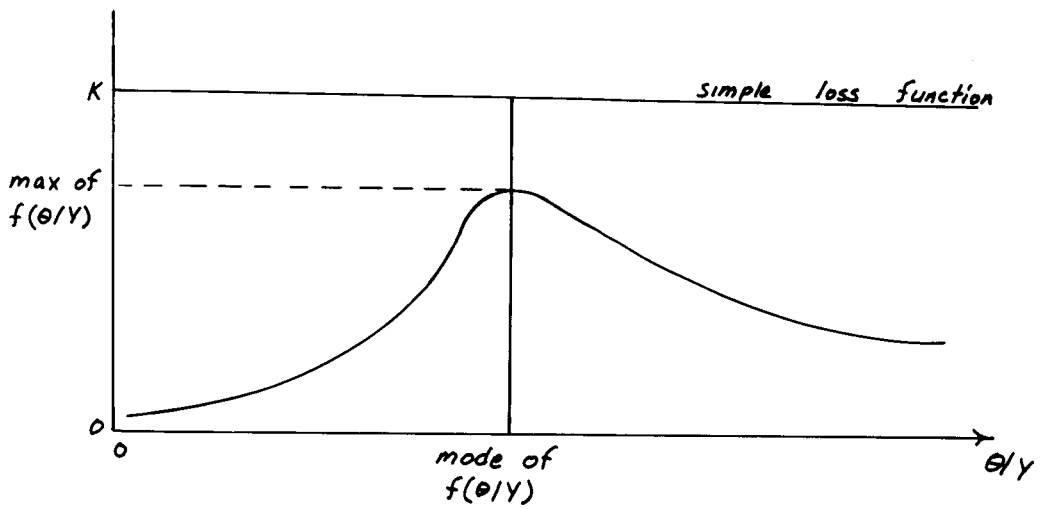
$$L[\delta(Y), \theta] = 0 \qquad \delta(Y) = \theta$$

$$= K > 0 \qquad \delta(Y) \neq \theta$$

The Bayes function is

$$B[\delta(Y)] = \int_{\theta} L[\delta, \theta] f(\theta/Y) d\theta$$

Hence, B [$\delta(Y)$] is minimized by maximizing $f(\Theta/Y)$ at $\delta = \Theta$. The integrand of B [$\delta(Y)$] is depicted in Figure 2.3.7. Of course, this is exactly the maximum-likelihood estimator for $f(\Theta/Y)$.



Thus, it can be stated that the Bayes estimator for a simple loss function is a maximum-likelihood estimator. However, it should not be concluded that maximum-likelihood estimators are Bayes estimators. The difference lies in the likelihood function which is maximized by the maximum-likelihood estimator and the loss function used. That is, the Bayes estimator for a simple loss function maximizes $f(\Theta/Y)$ which is also a maximum-likelihood estimator. However, a maximum likelihood estimator can also maximize the conditional pdf $f(Y/\Theta)$, which is not a Bayes estimator. It should be noted that a Bayes estimator for simple loss can be the same as that for squared-error loss. That is, if the maximum value or mode of $f(\Theta/Y)$ occurs for $E(\Theta/Y)$, then the Bayes estimator is the same for both simple loss and squared-error loss functions. Therefore, if the mode and $E(\Theta/Y)$ of $f(\Theta/Y)$ are coincident, the maximum likelihood estimator for $f(\Theta/Y)$ is the Bayes estimator for squared-error loss, also. However, this is only true if $E(\Theta/Y)$ is equal to the mode of $f(\Theta/Y)$.

Consider the particular case of Section 2.3.5.2.1 wherein $f(\Theta/Y)$ was Gaussian with

$$E(\Theta/Y) = \frac{\sigma^2}{1+m\sigma^2} \sum_{i=1}^{m} y_i \quad .$$

The maximum of $f(\Theta/Y)$ occurs at $E(\Theta/Y)$, therefore, the Bayes estimator is the same for both simple and squared-error loss. It is easily shown that the maximum likelihood estimator for both $f(\Theta/Y)$ and $f(Y, \Theta)$ are $E(\Theta, Y)$; however, if the likelihood function $f(Y/\Theta)$ is used, then the maximum likelihood estimator is the same as the minimum variance unbiased estimator, $1/m \sum y_i (i=1,m)$, determined in Section 2.3.5.2.1.

In general, if the likelihood function is dependent on the a priori pdf for Θ, then the maximum likelihood estimator is the same as the Bayes estimator for a simple loss function, and both estimators utilize the a'priori information available concerning Θ. If, on the other hand, the likelihood function is independent of the pdf for Θ, the a'priori information concerning Θ is not utilized. However, it should not be concluded that Bayes estimators can generally be derived by the method of maximum likelihood.

2.3.6 Application of Bayes Estimation

2.3.6.1 Introduction

The discussions of the previous sections have established the criteria of value in defining estimators and measures of the "goodness" which results. This section is intended to conclude the discussion of estimation by applying the most general of the previously developed estimators (Bayes) to the general non-linear estimation problems. This application will, however, fall short of providing computational algorithms which can be utilized in analyses since the pdf's for the variables involved and the functional relationships between them must be specified. Thus, the special case of the general linear system of equations and Gaussian statistics will be developed.

2.3.6.2 Non-Linear Case

The general case of interest is defined by the following vector equation: (See Section 2.3.3.)

$$\underline{y} = \underline{F}(\underline{\theta}) + \underline{e}$$

The parameter vector $\underline{\theta}$ contains n elements, and an estimator for each element is required. Thus, an estimator is denoted by the vector $\underline{\delta}$. Estimation error is likewise a vector defined by

$$\underline{\epsilon} = (\underline{\delta} - \underline{\theta}) .$$

A loss function for the estimation error vector $\underline{\epsilon}$ is defined as a scalar function of some positive definite measure of the modulus of $\underline{\epsilon}$; e.g., sum-squared-error, SSE, is a particular loss function which is defined by

$$SSE = \sum_{i=1}^{n} e_i^2$$

or

$$L(\underline{\epsilon}) = \underline{\epsilon}^T \underline{\epsilon}$$

where T denotes transpose.

A more general loss function is defined in terms of the squared modulus of a linear transformation of $\underline{\epsilon}$; i.e.,

$$L(\underline{\epsilon}) = \epsilon^T B^T B \underline{\epsilon}$$

$$= \epsilon^T M \epsilon$$

where B is some linear transformation and M is a real symmetrical matrix. In this case, $L(\underline{\epsilon})$ is a generalized quadratic loss function which is a positive semi-definite function with $L(\underline{\epsilon}) \equiv 0$ if, and only if, $\underline{\epsilon} \equiv 0$ for $B \neq 0$. Of course, for $M = I$, $L(\underline{\epsilon}) = SSE = \underline{\epsilon}^T \underline{\epsilon}$.

In general, $L(\underline{\epsilon})$ is a function of $\underline{\delta}$ and $\underline{\theta}$; i.e., $L(\epsilon) = L(\underline{\delta}, \underline{\theta})$. For generalized quadratic loss, $L(\underline{\delta}, \underline{\theta})$ becomes

$$L(\underline{\delta}, \underline{\theta}) = (\underline{\delta} - \underline{\theta})^T M (\underline{\delta} - \underline{\theta})$$

$$= \underline{\delta}^T M \underline{\delta} - 2\underline{\delta}^T M \underline{\theta} + \underline{\theta}^T M \underline{\theta} .$$

Note. This form exists because each of the terms involved is a scalar.

The Bayes estimator $\underline{\delta}_B$ for generalized quadratic loss is the vector of estimators which minimize the expected risk or the a'posteriori risk, as discussed in Section 2.3.5.2. The Bayes function for this loss is

$$B[\underline{\delta}(\underline{y})] = \int_{\theta} L(\underline{\delta}, \theta) f(\underline{\theta}/\underline{y}) \, d\theta$$

$$= \int_{\theta} (\underline{\delta}^T M \underline{\delta} - 2\underline{\delta}^T M \underline{\theta} + \underline{\theta}^T M \underline{\theta}) f(\underline{\theta}/\underline{y}) d\theta .$$

The Bayes estimator $\underline{\delta}_B$ can be found by differentiating $B[\underline{\delta}(\underline{y})]$ with respect to $\underline{\delta}$ and equating to zero, to wit,

$$\frac{\partial}{\partial \underline{\delta}} B\left[\underline{\delta}(\underline{y})\right] = \frac{\partial}{\partial \underline{\delta}} \int_{\theta} (\underline{\delta}^T M \underline{\delta} - 2\underline{\delta}^T M \underline{\theta} + \theta^T M \theta) f(\underline{\theta}/\underline{y}) d\theta$$

$$= \int_{\theta} (2M\underline{\delta} - 2M\theta) f(\underline{\theta}/\underline{y}) d\theta$$

$$= 2\int_{\theta} M(\underline{\delta} - \underline{\theta}) f(\underline{\theta}/\underline{y}) d\theta$$

$$\frac{\partial}{\partial \delta} B\left[\underline{\delta}(\underline{y})\right] = 2M\left[\underline{\delta}\int_{\theta} f(\underline{\theta}/\underline{y}) d\theta - \int_{\theta} \theta f(\underline{\theta}/\underline{y}) d\theta\right] .$$

The resulting expression can be made identically zero for all n if

$$\underline{\delta}\int_{\theta} f(\underline{\theta}/\underline{y}) d\theta = \int_{\theta} \theta f(\underline{\theta}/\underline{y}) d\theta$$

or

$$\underline{\delta} = \int_{\theta} \theta f(\underline{\theta}/\underline{y}) d\theta$$

since

$$\int_{\theta} f(\underline{\theta}/\underline{y}) d\theta \equiv 1$$

Thus, it is seen that for a generalized quadratic loss function the Bayes estimator $\underline{\delta}_B$ is the conditional expectation of $\underline{\theta}$, given $\underline{y}$; i.e.,

$$\underline{\delta}_B = \int_{\theta} \underline{\theta} f(\underline{\theta}/\underline{y}) d\theta = E(\underline{\theta}/\underline{y}) .$$

This result is equivalent to that determined in Section (2.3.5.2.2) for the explicit case of a single parameter θ.

The Bayes risk $\bar{R}_B$ for $\underline{\delta}_B$ is the expected value of $B(\underline{\delta}_B)$ over all $\underline{y}$, i.e.,

$$\bar{R}_B = E_{\underline{y}}\left[B(\underline{\delta}_B)\right]$$

$$= E_{\underline{y}}\left\{E\left[L(\underline{\delta}_B, \theta)/\underline{y}\right]\right\}$$

For generalized quadratic loss, the Bayes risk becomes

$$\bar{R}_B = E_y\left[\int_\theta (\underline{\delta}_B^T M \underline{\delta}_B - 2\underline{\delta}_B^T M \underline{\theta} + \underline{\theta}^T M \underline{\theta}) f(\underline{\theta}/\underline{y})\, d\theta\right]$$

$$= E_y\left[\int_\theta \underline{\theta}^T M \underline{\theta} f(\underline{\theta}/\underline{y})\, d\theta - \underline{\delta}_B^T M \underline{\delta}_B\right]$$

$$= E_y\left\{E\left[(\underline{\theta}^T M \underline{\theta})/\underline{y}\right] - \underline{\delta}_B^T M \delta_B\right\}$$

$$\bar{R}_B = E_y\left\{E\left[(\underline{\theta}^T M \underline{\theta})/\underline{y}\right]\right\} - E_y(\underline{\delta}_B^T M \underline{\delta}_B)$$

If $B(\underline{\delta}_B)$ is independent of $\underline{y}$, then

$$\bar{R}_B = B(\underline{\delta}_B) = E\left[L(\underline{\delta}_B, \underline{\theta})/\underline{y}\right]$$

$$\bar{R}_B = E\left[(\underline{\theta}^T M \underline{\theta})/\underline{y}\right] - \underline{\delta}_B^T M \underline{\delta}_B \ .$$

It should be emphasized at this time that the results given above are not restricted to the linear case since the explicit form of the process does not enter into the determination of the Bayes estimator as the conditional expectation of $\underline{\theta}$, given $\underline{y}$. However, it should be noted that the form of the process, as determined by $\underline{F}(\underline{\theta})$, definitely enters into the determination of the $E(\underline{\theta}/\underline{y})$; and, therefore, the Bayes estimator is dependent on the form of the process. That is, the conditional expection of $\underline{\theta}$, given $\underline{y}$, is the Bayes estimator for a generalized quadratic loss function for any functional relationship of $\underline{\theta}$ and $\underline{y}$. On the other hand, the specific form of the Bayes estimator is dependent on the form of $\underline{F}(\underline{\theta})$. Further, the Bayes estimator can be a non-linear function of the observation vector $\underline{y}$. Moreover, the Bayes estimator for a particular case is dependent on the statistics involved; i.e., the pdf's $f(\underline{\theta})$, and $f(\underline{e})$. Thus, it can be concluded that for a generalized quadratic loss function the conditional expectation of $\underline{\theta}$, given $\underline{y}$, is the Bayes estimation which will minimize the expected risk. In a particular case, the specific form of the Bayes estimator is determined by the form of the process, $\underline{F}(\theta)$, and the statistical nature of $\underline{\theta}$ and $\underline{e}$ as reflected in their pdf's.

The determination of a Bayes estimator for the general case can be formulated by determining the conditional pdf, $f(\underline{\theta}/\underline{y})$, and its conditional expectation. This procedure will be outlined for the case that the statistical nature of $\underline{\theta}$ and $\underline{e}$ must be known; i.e., the pdf's $f(\underline{\theta})$ and $f(\underline{e})$ are assumed available.

First, the conditional pdf of $\underline{y}$, given $\underline{\theta}$, is determined by considering $\underline{F}(\underline{\theta})$ as a constant in the process $\underline{y} = \underline{F}(\underline{\theta}) + \underline{e}$. Thus, $f(\underline{y}/\underline{\theta})$ is derived from $f(\underline{e})$ by including the additional term $\underline{F}(\underline{\theta})$ in the expected value of $\underline{e}$.

Second, the joint pdf of $\underline{y}$ and $\underline{\theta}$, $f(\underline{y}/\underline{\theta})$ is formed. For the most common situation of statistically independent $\underline{\theta}$ and $\underline{e}$, $f(\underline{y}/\underline{\theta})$ $f(\underline{\theta})$.

Third, the marginal pdf of $\underline{y}$, $f(\underline{y})$, is determined by integrating $f(\underline{y}, \underline{\theta})$ over all $\underline{\theta}$, to wit,

$$f(\underline{y}) = \int_{\theta} f(\underline{y}/\theta) f(\underline{\theta})\, d\theta$$

Fourth, the conditional pdf of $\underline{\theta}$, given $\underline{y}$, is formed by the standard form for conditional pdf's, to wit,

$$f(\underline{\theta}/\underline{y}) = \frac{f(\underline{y}, \theta)}{f(\underline{y})} = \frac{f(\underline{y}/\theta) \cdot f(\theta)}{f(\underline{y})}$$

Fifth, the conditional expectation of $\underline{\theta}$, given $\underline{y}$, is determined from $f(\underline{\theta}, \underline{y})$. Once $f(\underline{\theta}/\underline{y})$ is determined, it is often possible to recognize the $E(\underline{\theta}/\underline{y})$ by inspection; and, hence, the Bayes is directly determined. Otherwise, the general approach is to determine the conditional expectation by integrating, to wit,

$$\delta_B = E(\underline{\theta}/\underline{y}) = \int_{\theta} \underline{\theta} f(\underline{\theta}/\underline{y})\, d\theta$$

The degree of difficulty in determining a Bayes estimator is directly related to the nature and degree of complexity of the functional form of $\underline{F}(\underline{\theta})$ which significantly affects the determination of $f(\underline{y})$. Moreover, the form of the pdf's $f(\underline{e})$ and $f(\underline{\theta})$ affect the determination of a Bayes estimator; and often the task can be somewhat difficult. However, the approach provides a general method of determining the optimum estimator using minimum expected risk as primary estimation criterion for generalized quadratic loss.

In general, it is to be expected that the Bayes estimator will be a non-linear function of $\underline{y}$, even in the case for a linear form of $\underline{F}(\underline{\theta}) = A\underline{\theta}$, depending on the statistical nature of $\underline{\theta}$ and $\underline{e}$. However, quite often it is highly desirable to utilize a linear estimator; and under such a constraint, the resulting estimator will not be a Bayes estimator though it is possible to determine the linear estimator which will minimize the expected risk. Such estimators would be sub-optimal in the sense that the Bayes estimator would possess uniformly smaller risk. The use of sub-optimal linear estimators must be considered on the basis of particular problems if the utilization of a non-linear estimator presents a difficult situation in terms of data processing an estimator mechanization in other respects.

On the other hand, the Bayes estimator can be a linear function of $\underline{y}$; and in such cases, no particular problem arises. This is particularly true for linear systems with Gaussian statistics. This case is of general interest and is discussed in detail below.

2.3.6.3 Linear Case, Gaussian Statistics

The general linear case is defined by $\underline{F}(\underline{\theta}) = A\underline{\theta}$, i.e., $\underline{y} = A\underline{\theta} + \underline{e}$, where A is a known matrix independent of $\underline{\theta}$. However, A can be time dependent or, generally, non-stationary. For the case of Gaussian statistics, it can be shown that the Bayes estimator is a linear function of $\underline{y}$. This particular case is of general interest and is derived in detail herein.

The pdf's for $\underline{\theta}$ and $\underline{e}$ are assumed to be n and m dimensional multivariate Gaussian pdf's with the following specification:

$$E(\underline{\theta}) = \underline{m}_\theta$$

$$V(\theta) = E(\underline{\theta} - \underline{m}_\theta)(\underline{\theta} - \underline{m}_\theta)^T$$

$$E(\underline{e}) = \underline{m}_e$$

$$V(e) = E(\underline{e} - \underline{m}_e)(\underline{e} - \underline{m}_e)^T$$

where, of course, $V(\theta)$ and $V(e)$ are the covariance matrices of $\underline{\theta}$ and $\underline{e}$, respectively.

Further, the pdf's $f(\underline{y}/\underline{\theta})$, $f(\underline{y}, \underline{\theta})$, $f(\underline{y})$ and $f(\underline{\theta}/\underline{y})$, where $\underline{\theta}$ and $\underline{e}$ are assumed as statistically independent are:

$$f(\underline{y}/\underline{\theta}) = c^m |V(e)|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\underline{x}^T V^{-1}(e)\underline{x}\right]$$

where

$$c = \frac{1}{\sqrt{2\pi}}$$

$$\underline{x} = \underline{y} - A\underline{\theta} - \underline{m}_e = \underline{y} - A\underline{\theta} + A\underline{m}_\theta - \underline{m}_y = (\underline{y} - \underline{m}_y) - A(\theta - \underline{m}_\theta)$$

$|V(e)|$ = determinant of $V(e)$

$$f(\underline{y}, \theta) = f(\underline{y}/\underline{\theta}) \cdot f(\theta)$$

Rather than integrating $f(\underline{y}, \underline{\theta})$ over all $\underline{\theta}$ to obtain $f(\underline{y})$, it is easier to use the fact that $\underline{y}$ is the linear sum of two independent Gaussian random variables, hence,

$$E(\underline{y}) = \underline{m}_y = A\,\underline{m}_\theta + m_e$$

$$V(\underline{y}) = E(\underline{y} - \underline{m}_y)(\underline{y} - \underline{m}_y)^T$$

$$= A\,V(\theta)A^T + V(e)$$

It follows that

$$f(\underline{y}) = c^m |V(\underline{y})|^{-1/2} \exp\left[-\frac{1}{2}(\underline{y} - \underline{m}_y)^T V^{-1}(\underline{y})(\underline{y} - \underline{m}_y)\right]$$

Now

$$f(\underline{\theta}/\underline{y}) = \frac{f(\underline{y}, \underline{\theta})}{f(\underline{y})} = \frac{f(\underline{y}/\underline{\theta}) \cdot f(\underline{\theta})}{f(\underline{y})}$$

Hence

$$f(\underline{\theta}/\underline{y}) \equiv c^m \frac{|V(\underline{y})|^{1/2}}{\left[|V(e)| \cdot |V(\theta)|\right]^{1/2}} \exp\left[-\frac{1}{2} Q(\underline{\theta}, \underline{y})\right]$$

where

$$Q(\theta, \underline{y}) = (\underline{W} - A\underline{V})^T V^{-1}(e)(\underline{W} - A\underline{V}) + \underline{V}^T V^{-1}(e)\,\underline{V} - \underline{W}^T V^{-1}(\underline{y})\,\underline{W}$$

and

$$\underline{W} = \underline{y} - \underline{m}_y$$

$$\underline{V} = \underline{\theta} - \underline{m}_\theta$$

The term $Q(\underline{\theta}, \underline{y})$ specifies the conditional Gaussian pdf $f(\underline{\theta}/\underline{y})$, and its general form is

$$Q(\underline{\theta}, \underline{y}) = \left[\theta - E(\underline{\theta}, \underline{y})\right] V^{-1}(\theta/\underline{y}) \left[\underline{\theta} - E(\theta, \underline{y})\right]$$

$$= (\underline{\theta} - \underline{\delta}_B)^T V^{-1}(\theta/\underline{y})(\underline{\theta} - \underline{\delta}_B)$$

where $V(\Theta/y)$ is the conditional covariance matrix of $\underline{\Theta}$, given $\underline{y}$. Thus, in order to determine the Bayes estimator, it is only necessary to determine the reduced quadratic form for $Q(\underline{\Theta}, \underline{y})$.

Expanding $Q(\underline{\Theta}, \underline{y})$ in terms of $\underline{w}$ and $\underline{v}$, it is found that

$$Q(\underline{\theta}, \underline{y}) = \underline{V}^T\left[V^{-1}(\theta) + A^T V^{-1}(e) A\right]\underline{V} - 2\underline{V}^T A^T V^{-1}(e)\underline{V} + \underline{W}^T\left[V^{-1}(e) - V^{-1}(y)\right]\underline{W}$$

$$= (\underline{V} - K\underline{W})^T V^{-1}(\theta/y)(\underline{V} - K\underline{W})$$

where K and $V^{-1}(\Theta/y)$ must be determined. It should be noted that the term K essentially determines the Bayes estimator since

$$(\underline{V} - K\underline{W}) = \left[\underline{\theta} - \underline{m}_\theta - K(\underline{y} - \underline{m}_y)\right]$$

$$= \left[\underline{\theta} - E(\underline{\theta}/\underline{y})\right]$$

$$(\underline{V} - K\underline{W}) = \left[\underline{\theta} - \underline{\delta}_B\right]$$

Thus, the Bayes estimator is

$$\underline{\delta}_B = E(\underline{\theta}/\underline{y})$$

$$\underline{\delta}_B = \underline{m}_\theta + K(\underline{y} - \underline{m}_y).$$

The term K can be determined by expanding the second expression for $Q(\underline{\Theta}, \underline{y})$

$$Q(\underline{\theta}, \underline{y}) = \underline{V}^T V^{-1}(\theta/y)\underline{V} - 2\underline{V}^T V^{-1}(\theta/y) K\underline{W} + W^T K^T V^{-1}(\theta/y) K\underline{W}$$

and by equating terms in the two products

$$V^{-1}(\theta/y) = \left[V^{-1}(\theta) + A^T V^{-1}(e) A\right]$$

$$V^{-1}(\theta/y) K = A^T V^{-1}(e)$$

$$K^T V^{-1}(\theta/y) K = V^{-1}(e) - V^{-1}(y)$$

Thus, it is found that two solutions exist for K. The second equation yields

$$K_1 = V(\theta/y)\, A^T V^{-1}(e)$$

Substituting the second equation into the third, it is seen that

$$K^T A^T = V(e)\left[V^{-1}(e) - V^{-1}(y)\right]$$

$$= I - V(e)V^{-1}(y)$$

Taking the transpose and substituting $V(e) = V(y) - AV(\Theta)A^T$, it is found that

$$AK = I - \left[V(y) - AV(\theta)A^T\right]V^{-1}(y)$$

$$= AV(\theta)A^T V^{-1}(y)$$

Thus,

$$K_2 = V(\theta)\, A^T V^{-1}(y)$$

The two solutions for K generally imply two Bayes estimators for the case being considered; however, it will be shown later that the Bayes estimator is unique and that K_1 and K_2 represent two equivalent forms of the same estimator.

Several comments concerning the Bayes estimator for this case are in order. It is seen that the Bayes estimator for this case is a linear function of $\underline{y}$; thus, $\underline{\delta}_B$ is a linear estimator.

The matrix K is a "gain" matrix which, in effect, specifies the weights in estimating $\underline{\theta}$ that are given to the deviations in the observed $\underline{y}$ from its mean value $\underline{m}_y$. If the mean value $\underline{m}_\theta$ is actually observed, then the Bayes estimate for $\underline{\theta}$ is the mean value of $\underline{\theta}$; i.e., for $\underline{y} = \underline{m}_y$, $\underline{\delta}_B = \underline{m}$. On the other hand, if the observed $\underline{y}$ is identically zero, then the Bayes estimate for $\underline{\theta}$ weights the mean value of $\underline{y}$ by the gain matrix K to form the estimate of $\underline{\theta}$; i.e., for $\underline{y} \equiv 0$

$$\underline{\delta}_B = \underline{m}\theta - K\underline{m}\, y$$

In general, the deviations of observed $\underline{y}$ from $\underline{m}_y$ are weighted by K to form the Bayes estimate $\underline{\delta}_B$ as shown. It is easily shown that $\underline{\delta}_B$ is a biased estimator, to wit,

$$E(\underline{\delta}_B/\underline{\theta}) = \underline{m}_\theta + K\left[E(\underline{y}/\underline{\theta}) - \underline{m}_y\right]$$

$$= \underline{m}_\theta + K\left[A\underline{\theta} + \underline{m}_e - \underline{m}_y\right]$$

$$E(\underline{\delta}_B/\underline{\theta}) = \underline{m}_\theta + KA\left[\underline{\theta} - \underline{m}_\theta\right]$$

Of course, if KA = I, then $E(\underline{\delta}_B/\underline{\theta}) = \underline{\theta}$; and $\underline{\delta}_B$ would be unbiased. It is interesting to note that if $V(\underline{e}) = 0$, which implies a constant $\underline{e} = \underline{m}_e$, then $V^{-1}(y) = (AV(\theta)\,A^T)^{-1}$ and $KA = V(\theta)A^T\,(AV(\theta)A^T)^{-1}A = I$, where A^{-1} exists. However, in this case

$$\underline{\delta}_B = \underline{m}_\theta + V(\theta)A^T\left[AV(\theta)A^T\right]^{-1}\left[\underline{y} - \underline{m}_y\right] = m_\theta + A^{-1}(\underline{y} - \underline{m}_y)$$

or

$$\underline{\delta}_B = \underline{m}_\theta + \theta + A^{-1}\underline{m}_e - A^{-1}\underline{m}_y = \theta$$

This is to be expected since, if $\underline{e}$ is a known constant, then $\underline{\theta}$ can be determined exactly.

In general, $V(e) \neq 0$, and an error $\underline{\epsilon}_B$ exists in the Bayes estimate. The Bayes error $\underline{\epsilon}_B$ is

$$\underline{\epsilon}_B = \underline{\delta}_B - \underline{\theta}$$

$$= \underline{m}_\theta + K(\underline{y} - \underline{m}_y) - \underline{\theta}$$

$$= \underline{m}_\theta - \underline{\theta} + K(\underline{y} - \underline{m}_y)$$

$$\underline{\epsilon}_B = (I - KA)(\underline{\theta} - \underline{m}_\theta) + K(\underline{e} - \underline{m}_e)$$

The conditional expectation of $\underline{\epsilon}_B$, given $\underline{\theta}$, is thus

$$E(\underline{\epsilon}_B/\underline{\theta}) = (\underline{m}_\theta - \underline{\theta}) + K\left[E(\underline{y}/\theta) - \underline{m}_y\right]$$

$$= (\underline{m}_\theta - \underline{\theta}) + K\left[A\underline{\theta} + \underline{m}_e - \underline{m}_y\right]$$

$$= (\underline{m}_\theta - \underline{\theta}) + KA(\underline{\theta} - \underline{m}_\theta)$$

$$E(\underline{\epsilon}_B/\theta) = (I - KA)(\underline{\theta} - \underline{m}_\theta) .$$

The total expectation of $\underline{\epsilon}_B$ is zero; i.e.,

$$E(\underline{\epsilon}_B) = E\left[E(\underline{\delta}_B/\theta)\right]$$

$$= \left[I - KA\right]\left[\underline{m}_\theta - E(\underline{\theta})\right]$$

$$= \left[I - KA\right](\underline{m}_\theta - \underline{m}_\theta)$$

$$E(\underline{\epsilon}_B) = 0$$

The covariance matrices for $\underline{\delta}_B$ and $\underline{\epsilon}_B$ can be readily derived and one found to be equal, to wit,

$$V(\delta_B/\theta) = E\left\{\left[\underline{\delta}_B - E(\underline{\delta}_B/\underline{\theta})\right]\left[\underline{\delta}_B - E(\underline{\delta}_B/\underline{\theta})\right]^T/\theta\right\}$$

$$V(\epsilon_B/\theta) = E\left\{\left[\underline{\epsilon}_B - E(\underline{\epsilon}_B/\underline{\theta})\right]\left[\underline{\epsilon}_B - E(\underline{\epsilon}_B/\underline{\theta})\right]^T/\theta\right\}$$

where

$$\underline{\delta}_B - E(\underline{\delta}_B/\theta) = K\left[A\underline{\theta} - A\underline{m}_\theta - \underline{y} - \underline{m}_y\right]$$

$$= K\left[\underline{e} - \underline{m}_e\right]$$

and, likewise,

$$\underline{\epsilon}_B - E(\underline{\epsilon}_B/\underline{\theta}) = K\left[\underline{e} - \underline{m}_e\right] .$$

Thus, $\underline{\delta}_B - E(\underline{\delta}_B/\Theta) = \underline{\epsilon}_B - E(\underline{\epsilon}_B/\underline{\Theta})$, and it follows that $V(\delta_B) = V(\epsilon_B)$. The covariance matrices for δ_B and ϵ_B can be determined in explicit form by using the conditional expectation of $\underline{\epsilon}_B$, given $\underline{Y}$, to wit,

$$V(\epsilon_B) = E\left\{E\left[(\underline{\epsilon}_B \underline{\epsilon}_B^T)/\underline{y}\right]\right\}$$
$$= E\left\{V(\underline{\epsilon}_B/\underline{y})\right\}$$

Now,

$$V(\epsilon_B/y) = E\left[(\underline{\delta}_B - \underline{\theta})(\underline{\delta}_B - \underline{\theta})^T/\underline{y}\right]$$
$$= E\left\{\left[\underline{\theta} - E(\underline{\theta}/\underline{y})\right]\left[\underline{\theta} - E(\underline{\theta}/\underline{y})\right]^T/\underline{y}\right\}$$
$$= V(\theta/y)$$

$$V(\epsilon_B/y) = \left(V^{-1}(\theta) + A^T V^{-1}(e) A\right)^{-1}.$$

Thus,

$$V(\epsilon_B) = E\, V(\epsilon_B/y)$$
$$V(\epsilon_B) = \left(V^{-1}(\theta) + A^T V^{-1}(e) A\right)^{-1}.$$

Thus, the covariance matrices for $\underline{\delta}_B$ and $\underline{\epsilon}_B$ are actually independent of K and are essentially dependent on the intrinsic aspects of the problem; i.e., the covariance matrices of the parameters, $\underline{\theta}$; the measurement errors, $\underline{e}$; and the transformation, A, of parameters $\underline{\theta}$ into observables, $\underline{y}$.

The Bayes risk can now be determined as follows: first, the generalized quadratic loss function is constructed

$$L(\underline{\epsilon}) = \underline{\epsilon}^T M \underline{\epsilon}$$
$$= (\underline{\delta}_B - \underline{\theta})^T M (\underline{\delta}_B - \underline{\theta})$$
$$L(\underline{\epsilon}) = \left[\underline{\theta} - E(\underline{\theta}/\underline{y})\right]^T M \left[\underline{\theta} - E(\underline{\theta}/\underline{y})\right].$$

Next, the Bayes risk is defined

$$\bar{R}_B = E\left\{E\left[L(\underline{e})/\underline{y}\right]\right\} \quad .$$

Finally, the weighting matrix (M) must be selected. Consider the special case (of particular interest) where the risk desired is the sum-squared-error loss; i.e., $M = I$ and $L(\underline{\epsilon}) = \underline{\epsilon}^T\underline{\epsilon} = SSE$. For this case, R_B is the minimum mean-sum-squared error, MMSSE. The SSE can be expressed as

$$\begin{aligned} SSE &= \epsilon_B^T \underline{\epsilon}_B = (\underline{\delta}_B - \theta,)^T(\underline{\delta}_B - \theta) \\ &= \underline{\delta}_B^T \underline{\delta}_B - 2\,\underline{\delta}_B^T \underline{\theta} + \underline{\theta}^T\underline{\theta} \quad . \end{aligned}$$

The Bayes risk becomes

$$\bar{R}_B = MMSSE = E\left\{E\left[(\underline{\epsilon}_B^T \underline{\epsilon}_B)/\underline{\theta}\right]\right\} \quad .$$

However, it is seen from the equations for $L(\underline{\epsilon})$ that the MMSSE can be determined from the trace of the covariance matrix for $\underline{\epsilon}_B$; i.e.,

$$\begin{aligned} \bar{R}_B &= MMSSE \\ &= TRACE\left[V(\epsilon_B)\right] \\ &= TRACE\left[V^{-1}(\theta) + A^T V^{-1}(e) A\right]^{-1} \end{aligned}$$

From the foregoing it is seen that the Bayes risk and the covariances of the Bayes estimator and its error are not explicitly dependent on K. However, this does not mean that the minimum expected risk is obtained independent of K. Rather, the Bayes estimator as a function of K achieves the minimum expected risk, which is explicitly dependent on $V(\Theta)$, $V(e)$ and A only. That is, the trace of $V(\epsilon_B)$ represents a minimum-mean sum squared error which is intrinsic to the problem, and the Bayes estimator achieves the MMSSE. Thus, for the linear case since K has two solutions, there exist two equivalent forms for the Bayes estimator.

The equivalence of K_1 and K_2 can be shown in the following manner. By inspection, it is seen that K_1 can be derived from K_2 by a series of elementary transformations (i.e., row and column operations of interchange, multiplications by scalar and additions); i.e., $K_2 = PK_1Q$ where P and Q must be non-singular square matrices of proper order. Since

$$K_1 = V(\theta/y)\, A^T V^{-1}(e)$$

$$K_2 = V(\theta)\, A^T V^{-1}(y) \quad ,$$

P and Q can be selected by inspection as

$$P = V(\theta)\, V^{-1}(\theta/y)$$

$$Q = V(e)\, V^{-1}(y) \quad .$$

Since P and Q are non-singular, $K_1 = P^{-1}K_2Q^{-1}$. Thus, it is established that K_1 and K_2 are equivalent matrices; i.e., one can be derived from the other by a set of elementary transformations. In particular, it can be seen that K_2 and K_1 transform directly by two identity matrices P and Q; i.e., let

$$P = V^{-1}(\theta/y)\, V(\theta/y) = I$$

$$Q = I \quad .$$

Then,

$$K_1 = V^{-1}(\theta/y)\, V(\theta/y)\, K_2\, I$$

$$= V^{-1}(\theta/y)\left[V^{-1}(\theta) + A^T V^{-1}(e)\right] V(\theta)\, A^T V^{-1}(y) \cdot I$$

$$= V^{-1}(\theta/y)\left[A^T V^{-1}(y) + A^T V^{-1}(e)\, A\, V(\theta)\, A^T V^{-1}(y)\right] \cdot I$$

$$= V^{-1}(\theta/y)\, A^T V^{-1}(e)\left[A\, V(\theta)\, A^T V^{-1}(y) + V(e)\, V^{-1}(y)\right] \cdot I$$

$$K_1 = K_2\left[\left(V(y) - V(e)\right) V^{-1}(y) + V(e)\, V^{-1}(y)\right] \cdot I$$

$$= K_2\left[I - V(e)\, V^{-1}(y) + V(e)\, V^{-1}(y)\right] \cdot I$$

$$K_1 = K_2 \cdot I \cdot I$$

Thus, K_1 and K_2 are identically equivalent. Moreover, consider the two non-singular linear transformations represented by the square matrices K_1A and K_2A. These transformations represent a mapping of parameter space into estimator space, since

$$\underline{\delta}_B = K\underline{y}$$

where, without loss of generality, the mean values of $\underline{G}$ and $\underline{e}$ are assumed zero. Hence, for $\underline{y} = A\underline{\Theta} + \underline{e}$

$$\delta_B = KA\underline{\Theta} + K\underline{e}.$$

2.3.6.3.1 Limiting Cases for Bayes Estimation in the Linear Case

In general, the Bayes estimator utilizes the available information concerning both the parameter to be estimated and the uncertainty involved in the observations of a process. In determining the Bayes estimator, the pdf's of $\underline{\Theta}$ and $\underline{e}$ are generally required which represents the a'priori information utilized in the Bayes estimator. It is informative to consider the Bayes estimator for limiting cases where the a'priori information is unavailable. In such cases, it can be shown that the Bayes estimator becomes equivalent to other estimators, such as least squares and minimum variance.

Two types of limiting cases will be considered which will be referred to as parameter and observation uncertainty limiting. These limiting cases can be representing by null matrices for $V(e)$, $V^{-1}(e)$, $V(\Theta)$ and $V^{-1}(\Theta)$, to wit,

$\text{Lim } V(e) = 0$	MINIMUM OBSERVATION UNCERTAINTY
$\text{Lim } V^{-1}(e) = 0$	MAXIMUM OBSERVATION UNCERTAINTY
$\text{Lim } V(\Theta) = 0$	MINIMUM PARAMETER UNCERTAINTY
$\text{Lim } V^{-1}(\Theta) = 0$	MAXIMUM PARAMETER UNCERTAINTY

That is, in the limit a null covariance matrix represents a constant, whereas a null inverse covariance matrix represents a uniformly distributed random variable. For these limiting cases, the following results are obtained.

For the case of minimum observation uncertainty, it is found that the parameters are determined exactly with zero risk. This is seen by considering the limit of $\underline{\delta}_B$ for the limit of $V(e)$, to wit,

$$\lim_{V(e)\to 0} \underline{\delta}_B = \underline{M}_\theta + \lim_{V(e)\to 0} K(\underline{y} - \underline{M}_y)$$

$$= \underline{M}_\theta + \lim_{V(e)\to 0} K(A\underline{\theta} + \underline{e} - A\underline{M}_\theta - \underline{M}e)$$

$$= \underline{M}_\theta + \lim_{V(e)\to 0} KA(\underline{\theta} - \underline{M}_\theta) + \lim_{V(e)\to 0} K(\underline{e} - \underline{M}_e)$$

$$= \underline{M}_\theta + \lim_{V(e)\to 0}\left[V(\theta)A^T V(\vec{y})\ A(\theta - \underline{M}_\theta)\right] + \lim_{V(e)\to 0}\left(V(\theta)A^T V(\vec{y})\right)\lim_{V(e)\to 0}(\underline{e} - \underline{M}_e)$$

$$= \underline{M}_\theta + V(\theta)A^T \lim_{V(e)\to 0}\left(V(\vec{y})\ A\right)(\underline{\theta} - \underline{M}_\theta) + \lim_{V(e)\to 0}\left(V(\theta)A^T V(\vec{y})\right)\cdot 0$$

$$= \underline{M}_\theta + V(\theta)A^T(A\,V(\theta)A^T)^{-1}A(\underline{\theta} - \underline{M}_\theta) + 0$$

$$= \underline{M}_\theta + V(\theta)A^T A^{-T} V(\vec{\theta})\ A^{-1}A(\underline{\theta} - \underline{M}_\theta)$$

$$= \underline{M}_\theta + I(\underline{\theta} - \underline{M}_\theta)$$

$$\lim_{V(e)\to 0} \underline{\delta}_B = \underline{\theta}$$

For the case of maximum observation uncertainty, it is found that the observations are ignored and the Bayes estimator is the mean value of the parameters; and the Bayes risk is the sum of the variances of the parameters, to wit,

$$\lim_{V^{-1}(e)\to 0} \underline{\delta}_B = \underline{M}_\theta + \lim_{V(e)\to 0} K(\underline{y} - \underline{M}_y)$$

$$= \underline{M}_\theta + \lim_{V(e)\to 0} V(\theta/Y)A^T V(\vec{e})\ (\underline{y} - \underline{M}_y)$$

$$\lim_{V^{-1}(e)\to 0} \underline{\delta}_B = \underline{M}_\theta + 0$$

$$\lim_{V^{-1}(e)\to 0} \overline{R}_B = \lim_{V^{-1}(e)\to 0} TRACE\ V(\theta/Y)$$

$$= TRACE \lim_{V^{-1}(e)\to 0} V(\theta/Y)$$

$$= \text{TRACE} \lim_{V^{-1}(e) \to 0} \left(V^{-1}(\theta) + A^T V^{-1}(e) A\right)^{-1}$$

$$= \text{TRACE}\ V(\theta)$$

$$\lim_{V^{-1}(e) \to 0} \bar{R}_B = E\left[(\underline{\theta} - \underline{M}_\theta)^T (\underline{\theta} - \underline{M}_\theta)\right]$$

For the case of minimum parameter uncertainty, it is again found that the observations are ignored and the Bayes is again the mean value of the parameters; however, the Bayes risk is zero since the parameters are constants equal to their mean values, to wit,

$$\lim_{V(\theta) \to 0} \underline{\delta}_B = \underline{M}_\theta + \lim_{V(\theta) \to 0} K(\underline{y} - \underline{M}_y)$$

$$= \underline{M}_\theta + \lim_{V(\theta) \to 0} V(\theta) A^T V^{-1}(y) (\underline{y} - \underline{M}_y)$$

$$= \underline{M}_\theta + 0$$

$$\lim_{V(\theta) \to 0} \underline{\delta}_B = \underline{M}_\theta$$

$$\lim_{V(\theta) \to 0} \underline{R}_B = \text{TRACE} \lim_{V(\theta) \to 0} \left(V^{-1}(\theta) + A^T V^{-1}(e)\ A\right)^{-1}$$

$$= \text{TRACE} \lim_{V(\theta) \to 0} \left[V^{-1}(\theta)\ \right]^{-1}$$

$$\lim_{V(\theta) \to 0} \bar{R}_B = 0$$

For the case of maximum parameter uncertainty, it is found that the Bayes estimator reduced in form to the methods of minimum variance, weighted least-squares and least-squares dependent on the characteristics of the observation errors. Without loss of generality, the mean values of $\underline{\theta}$ and $\underline{e}$ will be taken as zero to show the direct equivalence of the Bayes estimator for maximum parameter uncertainty and the estimators derived in Section 2.2.2, to wit,

$$\lim_{V^{-1}(\theta)\to 0} \hat{\theta}_B = \lim_{V^{-1}(\theta)\to 0} K(y)$$

$$= \lim_{V^{-1}(\theta)\to 0} V(\theta/Y)\, A^T V^{-1}(e)\, y$$

$$= \lim_{V^{-1}(\theta)\to 0} \left(V(\theta)^{-1} + A^T V^{-1}(e)\, A\right)^{-1} A^T V \qquad y$$

$$\lim_{V^{-1}(\theta)\to 0} \hat{\theta}_B = \left(A^T V^{-1}(e) A\right)^{-1} A^T V^{-1}(e)\, y$$

$$\lim_{V^{-1}(\theta)\to 0} \bar{R}_B = TRACE\, (A^T V^{-1}(e)\, A)^{-1}$$

If the observation errors are statistically independent and stationary, i.e., $V(e) = \sigma_e^2 I$, then the Bayes estimator for maximum parameter uncertainty is the same as the least-squares estimator as derived in Section 2.2.2.1. The Bayes risk for this case is σ_e^2 TRACE $(A^TA)^{-1}$.

If the observation errors are statistically independent and non-stationary; i.e., V(e) is a diagonal matrix, then the Bayes estimator for maximum parameter uncertainty is the same as the weighted-least-squares estimator as derived in Section 2.2.2.2.

If the observation errors are generally statistically dependent and non-stationary, then the Bayes estimator for maximum parameter uncertainty is the same as the minimum-variance estimator as derived in Section 2.2.2.3.

From the results presented here, it follows that the estimators derived in Section 2.2.2 are minimum-mean-sum-squared-error estimators if the observation errors are Gaussian and if the parameters are uniformly distributed over the parameter space, which represents no a'priori information available concerning the parameters.

2.3.6.3.2 Single Parameter Estimation

The Bayes estimator given in the previous section is a generalization of the linear case considered in Section 2.3.5.2.1 for which $y = \theta + e$ where θ and e were Gaussian with zero mean values and variances σ and 1, respectively. This previous example can, however, be considered as a special case of single parameter estimation. In order to demonstrate some of the general characteristics of Bayes estimation, this example will be considered on a more general basis. For single parameter estimation, the set of observations can be

written as

$$\underline{y} = a\underline{1}\theta + \underline{e}$$

where $\underline{1}$ is the unity column vector of order m whose components are each unity; i.e., 1. The components of $\underline{e}$ are considered as the errors of observations, i.e., e_i in the error in the observation y_i. The parameter θ is Gaussian with mean value μ_θ and variance σ_θ^2. The error vector is also Gaussian with mean value vector $\underline{\mu}_e$ and covariance matrix V(e) as follows:

$$V(e) = E\left[(\underline{e} - \underline{\mu}_e)(\underline{e} - \underline{\mu}_e)^T\right]$$

The Bayes estimator δ_B for θ is given by

$$\delta_B = \mu_\theta + K(\underline{y} - \underline{\mu}_y)$$

where

$$\underline{\mu}_y = (a\,\underline{1}\,\mu_\theta + \underline{\mu}_e)$$

$$K = a\,V(\theta/Y)\,\underline{1}^T V^{-1}(e)$$

$$V(\theta/Y) = \left[\frac{1}{\sigma_\theta^2} + a^2\underline{1}^T V^{-1}(e)\underline{1}\right]$$

The Bayes risk for sum-squared-error loss is minimum-mean-squared-error, MMSE, since only a single parameter exists. Moreover, $V(\theta/Y)$ is a scalar, and MMSE = R_B = TRACE $V(\theta/Y)$ becomes

$$\bar{R}_B = MMSE = TRACE\ V(\theta/Y) = V(\theta/Y)$$

$$= \left[\frac{1}{\sigma_\theta^2} + a^2\underline{1}^T V^{-1}(e)\underline{1}\right]$$

The Bayes estimator can thus be written in the following form:

$$\delta_B = \mu_\theta - a\,V(\theta/Y)\left(\underline{1}^T V^{-1}(e)\underline{\mu}_y\right) + a\,V(\theta/Y)\left(\underline{1}^T V^{-1}(e)\underline{y}\right)$$

Consider the case of independent observations with equal mean values and variances for the observation errors; i.e., $V(e) = e^2 I$ and $\underline{\mu}_y = (a\mu_\theta + \mu_e)\cdot \underline{1}$. Substituting and noting that $\underline{1}^T\underline{1} = m$, it is found that

$$V(\theta/Y) = \left[\frac{1}{\sigma_\theta^2} + a^2 \underline{1}^T\left(\frac{I}{\sigma_e^2}\right)1\right]^{-1}$$

$$= \left[\frac{1}{\sigma_\theta^2} + a^2\left(\frac{\underline{1}^T\underline{1}}{\sigma_e^2}\right)\right]^{-1}$$

$$= \left[\frac{1}{\sigma_\theta^2} + \frac{a^2 M}{\sigma_e^2}\right]^{-1}$$

$$= \left(\frac{\sigma_e^2 + Ma^2\sigma_\theta^2}{\sigma_\theta^2\,\sigma_e^2}\right)^{-1}$$

$$V(\theta/Y) = \left(\frac{\sigma_\theta^2\,\sigma_e^2}{\sigma_e^2 + Ma^2\sigma_\theta^2}\right)$$

and

$$\underline{1}^T V^{-1}(e) = \underline{1}^T \cdot \frac{1}{\sigma_e^2} I$$

$$= \frac{1}{\sigma_e^2}\underline{1}^T$$

The Bayes estimator is thus

$$\delta_B = \mu_\theta - a\,V(\theta/Y)\left(\frac{1}{\sigma_e^2}\right)\underline{1}^T\underline{\mu}_y + a\,V(\theta/Y)\left(\frac{1}{\sigma_e^2}\right)\underline{1}^T\underline{y}$$

$$= \mu_\theta - K(\underline{1}^T\underline{1})(a\mu_\theta + \mu_e) + K\,\underline{1}^T y$$

$$\delta_B = \mu_\theta - MK(a\mu_\theta + \mu_e) + K\sum_{i=1}^{m} y_i$$

where $K = a\,V(\theta/Y)\left(\frac{1}{\sigma_e^2}\right)$

$$= \frac{a\,\sigma_\theta^2}{\sigma_e^2 + Ma^2\sigma_\theta^2}$$

$$= \frac{1}{a\left[M + \frac{\sigma_e^2}{a^2\sigma_\theta^2}\right]}$$

The corresponding Bayes risk is

$$R_B = MMSE$$

$$= \frac{\sigma_e^2}{a} K$$

$$= \frac{\sigma_e^2}{a^2\left[M + \frac{\sigma_e^2}{a^2\sigma_\theta^2}\right]}$$

For $\mu_\theta = \mu_e = 0$ and $\sigma e^2 = a = 1$, δ_B and $\overline{R}_B$ reduce to the results obtained previously, to wit,

with $$\delta_B = \frac{1}{\left(M + \frac{1}{\sigma_\theta^2}\right)} \sum_{i=1}^{m} y_i$$

$$\overline{R}_B = \frac{1}{\left(M + \frac{1}{\sigma_\theta^2}\right)} \quad .$$

For the more general case, the following comments are in order which illustrate some general characteristics of Bayes estimation. It is seen that the Bayes risk, or MMSE, is proportional to σ_e^2 and inversely proportional to a^2. However, as σ_e^2 becomes large without limit, the Bayes risk approaches a definite limit, i.e.,

$$\lim_{\sigma_\theta^2 \to \infty} \overline{R}_B = \sigma_\theta^2$$

The reason for this limit is found in the limit of K and δ_B as $\sigma_e^2 \to \infty$, i.e.,

$$\lim_{\sigma_e^2 \to \infty} K = 0$$

and

$$\lim_{\sigma_e^2 \to \infty} \delta_B = \mu_\Theta$$

Thus, it is seen that as σ_e^2 becomes indefinitely large, the observations are essentially ignored and the Bayes estimate of Θ is taken as the mean value of Θ. The risk in this case is σ_Θ^2 which is simply the uncertainty in Θ as measured by σ_Θ^2.

The results given above can be readily extended to the case of non-stationary observations where, in the variance and mean value of observation, error varies. In this case $V(e) \neq \sigma_e^2 I$ and the Bayes estimator must be modified accordingly. Consider the case of independent observations with the mean value and variance of the error in each observation being μ_i and σ_i^2. In this case

$$V(e) = \begin{vmatrix} \sigma_1^2 & 0 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 & 0 \\ 0 & 0 & \cdot & 0 & 0 \\ 0 & 0 & 0 & \cdot & 0 \\ 0 & 0 & 0 & 0 & \sigma_m^2 \end{vmatrix}$$

and

$$V^{-1}(e) = \begin{vmatrix} \sigma_1^{-2} & 0 & 0 & 0 \\ 0 & \sigma_2^{-2} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_m^{-2} \end{vmatrix}$$

The term $V(\Theta/Y)$ becomes

$$V(\Theta/Y) = \left(\frac{1}{\sigma_\Theta^2} + a^2 \sum_{i=1}^{m} \frac{1}{\sigma_i^2} \right)^{-1}$$

Also,

$$\underline{1}^T V^{-1}(e)\underline{\mu}_y = \underline{1}^T V^{-1}(e)\left(a\,\underline{1}\mu_\theta + \underline{\mu}_e\right)$$

$$= a\mu_\theta \sum_{i=1}^{m} \frac{1}{\sigma_i^2} + \sum_{i=1}^{m} \frac{\mu_i}{\sigma_i^2}$$

$$\underline{1}^T V(e)\,\underline{\mu}_y = \sum_{i=1}^{m}\left(\frac{\mu_i + a\mu_\theta}{\sigma_i^2}\right)$$

and

$$\underline{1}^T V^{-1}(e)\,\underline{y} = \sum_{i=1}^{m}\left(\frac{y_i}{\sigma_i^2}\right)$$

The Bayes estimator becomes

$$\delta_B = \mu_\theta + a\,V(\theta/Y)\left[\underline{1}^T V^{-1}(e)\,\underline{y} - \underline{1}^T V^{-1}(e)\,\underline{\mu}_y\right]$$

$$= \mu_\theta + a\,V(\theta/Y)\left[\sum_{i=1}^{m}\left(\frac{y_i}{\sigma_i^2}\right) - \sum_{i=1}^{m}\left(\frac{\mu_i + a\mu_\theta}{\sigma_i^2}\right)\right]$$

$$\delta_B = \mu_\theta + a\,\frac{\left[\sum_{i=1}^{m}\left(\frac{y_i - \mu_i - a\mu_\theta}{\sigma_i^2}\right)\right]}{\left[\frac{1}{\sigma_\theta^2} + a^2\sum_{i=1}^{m}\left(\frac{1}{\sigma_i^2}\right)\right]}$$

In this case it is seen that in the Bayes estimator for Θ each observation y_i is "unbiased" by substracting $\mu_i + a\mu_\theta$ from y_i and the resulting unbiased observation is weighted inversely proportional to the variance of the error in the observation. The unbiased, weighted observations are then summed and again weighed, multiplied by $aV(\Theta/Y)$ and added to the mean value of Θ, μ_θ, to obtain an estimate of Θ. In this manner, the observations which have error of large variance are essentially de-emphasized in the estimation of $\underline{\Theta}$. The Bayes estimator can thus be considered as a selective filter of observation data.

2.3.6.3.3 Recursive Bayes Estimators

For reasons outlined in Section 2.2.2.4 and 2.2.2.5, it is frequently desired that a new estimate of the parameters be generated from a previous estimate and information acquired since this lost estimate. It might be argued that this step can be performed simply by utilizing the last estimate of the parameters and the covariance matrix of the estimation errors as the a'priori information or initial conditions for the continued process. This argument is predicated on the observation that an optimum estimate must be locally optimum to be optimum in the large. However, to demonstrate that this situation does in fact result in the optimum estimate it is essential that the risk for the estimate still be minimum. The following paragraphs were prepared in way of proof.

Let the random sample Y of m observations be grouped in terms of subsets of m observations each, i.e.,

$$Y = (\underline{y}_1, \underline{y}_2, \cdots, \underline{y}_j, \cdots, \underline{y}_\ell)$$

where

$$m = \sum_{j=1}^{\ell} m_j$$

Each subset $\underline{y}_j$ of observations is related to the parameter vector $\underline{\theta}$ as follows:

$$\underline{y}_j = A_j \underline{\theta} + \underline{e}_j$$

The total observations can be written as a function of the parameter vector in partitioned form. In this manner the system $\underline{y} = A\underline{\theta} + \underline{e}$ becomes

$$\begin{vmatrix} \underline{y}_1 \\ \underline{y}_2 \\ \vdots \\ \underline{y}_j \\ \vdots \\ \underline{y}_\ell \end{vmatrix} = \begin{vmatrix} A_1 \\ A_2 \\ \vdots \\ A_j \\ \vdots \\ A_\ell \end{vmatrix} \underline{\theta} + \begin{vmatrix} \underline{e}_1 \\ \underline{e}_2 \\ \vdots \\ \underline{e}_j \\ \vdots \\ \underline{e}_\ell \end{vmatrix}$$

It becomes apparent that a Bayes estimator for $\underline{\theta}$ can be determined for any particular subset or collection of subsets of observations. The objective of this effort, however, is to show that the Bayes estimator $\underline{\delta}_B^K$ for the first K subsets of observations, $Y_K = (\underline{y}_1, \underline{y}_2, \ldots\ldots, \underline{y}_K)$ is an explicit function of $\underline{\delta}_B^{K-1}$ only. That is, the Bayes estimator can be determined in a "single-step" recursive form. This objective is readily accomplished for the case of observations which are statistically dependent within subsets but which are statistically independent between subsets. In this case the covariance matrix of the observation errors becomes a partitioned diagonal matrix; i.e.,

$$V(e) = \begin{vmatrix} V(e)_1 & 0 & 0 & 0 & 0 \\ 0 & V(e)_2 & 0 & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & V(e)_j & \\ 0 & 0 & 0 & 0 & V(e)_\ell \end{vmatrix}$$

where $V(e)_j$ is the covariance matrix for the observation error of the jth subset of observations.

The Bayes estimator $\underline{\delta}_B$ for the total set of observations Y is

$$\underline{\delta}_B = M_\theta + \left(V^{-1}(\theta) + A^T V^{-1}(e)\ A\right)^{-1} A^T V^{-1}(e)\ \left(\underline{y} - \underline{M}_y\right)$$

Adding $-M_\theta$ and premultiplying by $(V^{-1}(\theta) + A^T V^{-1}(e)\ A)$ it is found that

$$\left(V^{-1}(\theta) + A^T V^{-1}(e)\ A\right)\left(\underline{\delta}_B - \underline{M}_\theta\right) = A^T V^{-1}(e)\ \left(\underline{y} - \underline{M}_y\right)$$

Substituting for A and V(e), the previous equation becomes

$$\left(V^{-1}(\theta) + \sum_{j=1}^{\ell} A_j^T V_j^{-1}(e)\ A_j\right)\left(\underline{\delta}_B - \underline{M}_\theta\right) = \sum_{j=1}^{\ell} A_j^T V_j^{-1}(e)\ \left(\underline{y}_j - \underline{M}_j\right)$$

In a similar manner, it is found that

$$V^k(\theta/Y)\ \left(\underline{\delta}_B^k - \underline{M}_\theta\right) = \sum_{j=1}^{k} A_j^T V_j^{-1}(e)\ \left(\underline{y}_j - \underline{M}_{y_j}\right)$$

where

$$V^{k}(\theta/Y) = \left(V^{-1}(\theta) + \sum_{j=1}^{k} A_j^T V_j^{-1}(e) A_j \right)$$

and $\underline{\delta}_B^K$ is the Bayes estimator for the first K subsets of observations. Of course, $\underline{\delta}_B^{K=1} = \underline{\delta}_B^1 = \underline{\delta}_B$. Similarly,

$$V^{k-1}(\theta/Y) \left(\underline{\delta}_B^{k-1} - M_\theta \right) = \sum_{j=1}^{k-1} A_j V_j^{-1}(e) \left(y_j - \underline{M} y_j \right) .$$

By adding $A_K V_K^{-1}(e)$ $(\underline{Y}_K - \underline{M}_{Y_K})$ to the right hand member,

$$V^{k}(\theta/Y) \left(\underline{\delta}_B^{k} - \underline{M}_\theta \right) = V^{k-1}(\theta/Y) \left(\underline{\delta}_B^{k-1} - \underline{M}_\theta \right) + A_k V_k^{-1}(e) \left(y_k - \underline{M} y_k \right)$$

Similarly,

$$V^{k}(\theta/Y) = V^{k-1}(\theta/Y) + A_k^T V_k^{-1}(e) A_k$$

By substituting $V^{K-1}(\Theta/Y)$ in terms of V^K (Θ/Y) it follows that

$$V^{k}(\theta/Y) \left(\underline{\delta}_B^{k} - \underline{M}_\theta \right) = V^{k}(\theta/Y) \left(\underline{\delta}_B^{k-1} - \underline{M}_\theta \right) + A_k V_k^{-1}(e) \left[y_k - \underline{M} y_k - A_k \left(\underline{\delta}_B^{k-1} - \underline{M}_\theta \right) \right]$$

Premultiplying by the inverse of $V^K(\Theta/Y)$ it is determined that

$$\underline{\delta}_B^{k} = \underline{\delta}_B^{k-1} + \left[V^{-1}(\theta) + \sum_{j=1}^{k} \left(A_j V_j^{-1}(e) A_j \right) \right]^{-1} A_k V_k^{-1}(e) \left[y_k - A_k \underline{\delta}_B^{k-1} - \underline{M} e_k \right] .$$

From the foregoing it is seen that the Bayes estimator can be determined as a "single-step" recursive process in which the observations are processed sequentially provided the errors in the observables are not sequentially correlated. In this manner, if the Bayes estimate $\underline{\delta}_B^K$ and V^K (Θ/Y) are known and an additional set of observations $\underline{Y}_{K+1}$ are available for which A_{K+1}, $V_{K+1}(e)$, and $\underline{M}_{e_{K+1}}$ are known, then the Bayes estimate $\underline{\delta}_B^{K+1}$ can be determined recursively as follows:

$$V^{k+1}(\theta/Y) = V^{k}(\theta/Y) + A_{k+1}^T V_{k+1}^{-1}(e) A_{k+1}$$

$$\underline{\delta}_B^{k+1} = \underline{\delta}_B^k + \left(V^{k+1}(\theta/Y)\right)^{-1} A_{k+1}' V_{k+1}^{-1}(e)\left(\underline{y}_{k+1}' - A_{k+1}\underline{\delta}_B^k\right)$$

where

$$\underline{y}_{k+1}' = \underline{y}_{k+1} - \underline{M}e_{k+1} \cdot$$

The initial conditions for $V^K(\theta/Y)$ and $\underline{\delta}_B^K$ are $V^o(\theta/Y) = V^{-1}(\theta)$ and $\underline{\delta}_B^o = \underline{M}_\theta$ since $\underline{\delta}_B'$ is given by

$$\underline{\delta}_B' = M_\theta + \left(V^{-1}(\theta) + A_1' V_1^{-1}(e) A_1\right)^{-1} A_1' V_1^{-1}(e)\left(\underline{y}_1 - \underline{M}y_1\right)$$

$$\underline{\delta}_B' = \underline{\delta}_B^o + \left(V^o(\theta/y) + A_1' V_1^{-1}(e) A_1\right)^{-1} A_1' V_1^{-1}(e)\left(\underline{y}_1 - \underline{M}y_1\right)$$

These initial conditions follow from the fact that $(V(\theta/Y)^{-1}$ is the covariance matrix of θ, given y. For the case of no observations, $V(\theta/Y) = V(\theta)$ and the Bayes estimate is simply $\underline{M}_\theta$.

2.4 THE STATE TRANSITION MATRIX

The navigational filters described in sections 2.2 and 2.3 required knowledge of the relationships between the state at various epochs so as to allow all observations to be referred to a given epoch or so as to allow for the generation of an a'priori estimate of the state at the epoch of the observation. Thus, this matrix is an explicit part of the navigational logic and will be discussed as such. It is noted before initiating these discussions, however, that there are many other applications for this matrix which will be explored in future monographs of this series (Midcourse Guidance).

Consider the following nonlinear system of equations (which define the motion of the vehicle being considered)

$$\frac{d}{dt}(\vec{x}) = \vec{f}(\vec{x}, t) + \vec{u}(t) \tag{4.1}$$

where $\vec{X}$ is the "state" vector for the system (commonly, $\left\{\vec{r}, \vec{v}\right\}^T$) and where $\vec{u}$ denotes the control applied to the system. Now consider a neighboring trajectory defined by the same system of equations and the vector displacement from a "nominal" solution

$$\vec{\delta}(t) = \vec{x} - \vec{x}_o \ . \tag{4.2}$$

Combining these two notations, the time rate of change of the displacement is obtained as

$$\dot{\vec{\delta}}(t) = \left[\vec{f}(\vec{x}, t) - \vec{f}_o(\vec{x}_o, t)\right] + \left[\vec{u}(t) - \vec{u}_o(t)\right] \tag{4.3}$$

but since $\vec{\delta}(t)$ is assumed to be small, $\vec{f}(\vec{x}, t)$ can be represented by a linear combination of two functions, i.e.,

$$\vec{f}(\vec{x}, t) = \vec{f}_o(\vec{x}, t) + \vec{\beta}(\vec{x}, t) \ . \tag{4.4}$$

Now, the vector difference $\left[\vec{f}_o(\vec{x}, t) - \vec{f}(\vec{x}_o, t)\right]$ can be expanded in a Taylor series as

$$\vec{f}_o(\vec{x}, t) - \vec{f}_o(\vec{x}_o, t) = \frac{\partial \vec{f}_o}{\partial \vec{x}}\, \vec{\delta} \ . \tag{4.5}$$

Similarly, the vector function $\vec{\beta}(\vec{x}, t)$ can be obtained from a Taylor series in terms of the state along the nominal trajectory

$$\vec{\beta}(\vec{x}, t) = \vec{\beta}(\vec{x}_o, t) + \frac{\partial \vec{\beta}}{\partial \vec{x}}\, \vec{\delta} \ . \tag{4.6}$$

Thus, the differential equations for the system become

$$\dot{\vec{\delta}} = \left[\frac{\partial \vec{f}}{\partial \vec{x}} + \frac{\partial \vec{\beta}}{\partial \vec{x}}\right]_{\vec{x} = \vec{x}_o} \vec{\delta} + \left[\vec{\beta}(\vec{x}_o, t) + \Delta \vec{u}(t)\right] \tag{4.7a}$$

$$\equiv A(t)\vec{\delta} + \vec{F}(t) . \tag{4.7b}$$

But, the trajectory for the motion under consideration itself is expressible (analytically and/or numerically) as a unique function of the position and velocity at some epoch, and as a function of the accelerations experienced relative to a nominal path (differences in the gravitational accelerations and in any applied thrusts). This set of relationships is represented by

$$\vec{r} = \vec{r}(\vec{r}_o + \delta\vec{r}_o, \dot{\vec{r}}_o + \delta\dot{\vec{r}}_o, \delta\vec{a}, t) \tag{4.8a}$$

$$\dot{\vec{r}} = \dot{\vec{r}}(\vec{r}_o + \delta r_o, \dot{\vec{r}}_o + \delta\dot{\vec{r}}_o, \delta\vec{a}, t). \tag{4.8b}$$

Thus, by a Taylor's expansion

$$\vec{r} = \vec{r}_n + \frac{\partial \vec{r}}{\partial r_o}\delta\vec{r}_o + \frac{\partial \vec{r}}{\partial \dot{\vec{r}}_o}\partial\dot{\vec{r}}_o + C\vec{f}(\delta\vec{a}) \tag{4.9}$$

and

$$\dot{\vec{r}} = \dot{\vec{r}}_n + \frac{\partial \dot{\vec{r}}}{\partial \vec{r}_o}\delta\vec{r}_o + \frac{\partial \dot{\vec{r}}}{\partial \dot{\vec{r}}_o}\partial\dot{\vec{r}}_o + C\vec{f}(\delta\vec{a}) \tag{4.10}$$

where $\vec{f}(\delta\vec{a})$ is a set of linear functions which additively (due to the small assumed magnitude of the deviations) represent the effects of the differences in the perturbing forces, and where B and C are 3 x 6 matrices. These linear equations can, in turn, be expressed as

$$\begin{Bmatrix} \vec{r} - \vec{r}_n \\ \dot{\vec{r}} - \dot{\vec{r}}_n \end{Bmatrix} \equiv \vec{\delta} = \begin{bmatrix} \frac{\partial \vec{r}}{\partial \vec{r}_o} & \frac{\partial \vec{r}}{\partial \dot{\vec{r}}_o} \\ \frac{\partial \dot{\vec{r}}}{\partial \vec{r}_o} & \frac{\partial \dot{\vec{r}}}{\partial \dot{\vec{r}}_o} \end{bmatrix} \vec{\delta}_o + \begin{bmatrix} B \\ C \end{bmatrix} \vec{f}(\delta\vec{a}) \tag{4.11a}$$

$$= \varphi(t, t_o)\vec{\delta}_o + \Delta(t, t_o)\vec{f}(\delta\vec{a}) \tag{4.11b}$$

showing that, to the first order, the dynamics of the problem can be represented by the two matrices $\varphi\ (t, t_o)$ and $\Delta(t, t_o)$.

Now, comparing the two equations for $\vec{\delta}$ and $\dot{\vec{\delta}}$ [equations (4.7) and (4.11)], it becomes apparent that the fundamental solution matrix [the state transition matrix, the matrix denoted by $\varphi\ (t, t_o)$] is the solution to the homogeneous equation obtained for $\vec{F} = 0$. In order to define the nature of the term $\Delta(t, t_o)\ \vec{f}\ (\delta\vec{a}\ )$ it is, however, necessary to consider the solution to $\vec{\delta}\ (t)$.

Let $\vec{\epsilon}\ (t)$ denote the homogeneous part of the solution for $\vec{\delta}\ (t)$. Then

$$\dot{\vec{\epsilon}}(t) = A(t)\,\vec{\epsilon}(t)$$

But

$$\vec{\epsilon}(t) = \varphi(t, t_o)\,\vec{\epsilon}(t_o)$$

Thus,

$$\dot{\varphi}(t, t_o) = A(t)\,\varphi(t, t_o) \tag{4.12}$$

At this point, introduce the system of differential equations adjoint to this family

$$\dot{\Lambda}(t, t_o) = -\Lambda(t, t_o)\ A(t) \tag{4.13}$$

and note that

$$\frac{d}{dt}\left[\Lambda(t, t_o)\ \delta(t)\right] = \Lambda(t, t_o)\ \vec{F}(t).$$

Performing the substitution yields

$$\frac{d}{dt}\left[\Lambda(t, t_o)\,\delta(t)\right] = \dot{\Lambda}(t, t_o)\,\vec{\delta} + \Lambda(t, t_o)\,\dot{\vec{\delta}} \tag{4.14}$$

which, upon integration, can be written as

$$\Lambda(t, t_o)\,\delta(t) = \Lambda(t_o, t_o)\,\delta(t_o) + \int_{t_o}^{t} \Lambda(\tau, t_o)\,\vec{F}(\tau)\,d\tau \tag{4.15a}$$

$$\equiv \delta(t_o) + \int_{t_o}^{t} \Lambda(\tau, t_o)\,\vec{F}(\tau)\,d\tau \tag{4.15b}$$

Multiplying by $\Lambda^{-1}(t, t_o)$ now reduces this result to

$$\delta(t) = \Lambda^{-1}(t, t_o)\,\delta(t_o) + \Lambda^{-1}(t, t_o) \int_{t_o}^{t} \Lambda(\tau, t_o)\,\vec{F}(\tau)\,d\tau \tag{4.16}$$

Finally, comparing the form of this equation to that given earlier [equation (4.11)] for $\vec{\delta}(t)$, it is apparent that

$$\varphi(t, t_o) = \Lambda^{-1}(t, t_o) \tag{4.17}$$

and that

$$\Delta(t, t_o)\,\vec{f}(\delta\vec{a}) = \varphi(t, t_o) \int_{t_o}^{t} \varphi^{-1}(\tau, t_o)\,\vec{F}(\tau)\,d\tau \tag{4.18}$$

The preceding material clearly establishes the mathematical nature of the state transition matrix and, at the same time, introduces the adjoint equations as one means which might be employed to compute this matrix. The sections which follow will explore these avenues in some detail for the purpose of providing the state transition matrix as required in the navigational filter.

2.4.1 Generating The State Transition Matrix

2.4.1.1 By Direct Integration

The introductory paragraphs to this section of the report showed that the state transition matrix obeys the differential equation (4.12)

$$\dot{\varphi}(t, t_o) = A(t)\, \varphi(t, t_o)$$

and the boundary conditions

$$\varphi(t_o, t_o) = I$$

Thus, the desired matrix can be obtained directly by numerical integration of each element of the array. This procedure is completely adequate for most applications in which the task is concerned only with the reduction of observed sightings (fixed values of t_o).

However, there are applications (midcourse corrections, for example) in which the state transition matrix relating the state of the present epoch and a fixed future epoch (e.g., the epoch of rendezvous) is desired. For these applications, it is unnecessarily complex to require that the preceding integration be reperformed at each epoch (i.e., a new t_o at each epoch). Rather, it is convenient to generate the desired solutions as follows:

$$\vec{\delta}(t) = \varphi(t, t_o)\, \vec{\delta}_o \qquad (4.19a)$$

$$\vec{\delta}(T) = \varphi(T, t_o)\, \vec{\delta}_o \quad . \qquad (4.19b)$$

Thus,

$$\vec{\delta}(T) = \varphi(T, t_o)\, \varphi^{-1}(t, t_o)\, \vec{\delta}(t). \qquad (4.20)$$

This technique requires that two integrations and an inversion be accomplished. First, the transition matrix relating the arbitrary epoch (t_o) and the fixed end time (T) is generated. [This integration need be performed only once, providing that $\vec{\delta}$ (t) never leaves the neighborhood of the nominal trajectory about which the partials are evaluated.] Secondly, the transition matrix relating the arbitrary epoch (t_o) and the epoch of the present state must be generated from the previous transition matrix, the differential equation, and the change in the time. Finally, the latter matrix must be inverted.

The first two of the three steps outlined are relatively simple. However, the third involves a considerable amount of effort if normal matrix inversion techniques are employed. Thus, it is instructive to present a means of developing the inverse in an analytic manner from the transition matrix itself. This procedure, if employed, will avoid roundoff and loss of significance problems inherent in numerical techniques, in addition to

drastically simplifying the mechanization logic.

Consider a linear system (expressed in cartesian coordinates) described by the following equation

$$\dot{\vec{\delta}} = A(t)\,\vec{\delta}(t) \tag{4.21}$$

where $\vec{\delta}$ (t) is an even-ordered state vector composed of a set of output variables and their derivatives; and A(t) is a coefficient array for the system composed of square, symmetric, even-ordered subarrays of the following form

$$A(t) = \left[\begin{array}{c|c} 0 & A_{12} \\ \hline A_{21} & 0 \end{array}\right] \tag{4.22a}$$

This form for A(t) is representative of motion in conservative force fields. To be specific

$$A_{12} = I \tag{4.22b}$$

$$A_{21} = \frac{\partial \vec{g}}{\partial \vec{r}} = \frac{\mu}{r^3}\left[I - \frac{3\hat{r}\hat{r}^T}{r^2}\right] \tag{4.22c}$$

for motion in the vicinity of

$$\vec{\delta}(t) = \varphi(t, t_o)\,\vec{\delta}(t_o) \equiv \left[\begin{array}{c|c} \varphi_{11} & \varphi_{12} \\ \hline \varphi_{21} & \varphi_{22} \end{array}\right] \vec{\delta}(t_o). \tag{4.23}$$

If equations (4.22) and (4.23) are substituted back into equation (4.21), the following identity results

$$\begin{bmatrix} \dot{\varphi}_{11} & \dot{\varphi}_{12} \\ \dot{\varphi}_{21} & \dot{\varphi}_{22} \end{bmatrix} = \begin{bmatrix} 0 & I \\ \frac{\partial \vec{g}}{\partial \vec{r}} & 0 \end{bmatrix} \begin{bmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{bmatrix}, \tag{4.24}$$

and upon expansion, equation (4.24) yields

$$\dot{\varphi}_{11} = \varphi_{21} \tag{4.25a}$$

$$\dot{\varphi}_{12} = \varphi_{22} \tag{4.25b}$$

$$\dot{\varphi}_{21} = \frac{\partial \vec{g}}{\partial r}\varphi_{11} \tag{4.25c}$$

$$\dot{\varphi}_{22} = \frac{\partial \vec{g}}{\partial r}\varphi_{12}\,. \tag{4.25d}$$

Equations (4.25) may now be operated on to produce an equivalent set of differential equations. This operation is performed as follows:

$$\Phi_{21}^T \dot{\Phi}_{11} = \Phi_{21}^T \Phi_{21}$$

$$\dot{\Phi}_{21}^T \Phi_{11} = \Phi_{11}^T \frac{\partial \hat{\dot{q}}}{\partial r}^T \Phi_{11} = \Phi_{11}^T \frac{\partial \hat{\dot{q}}}{\partial \dot{r}} \Phi_{11}$$

$$\Phi_{11}^T \dot{\Phi}_{21} = \Phi_{11}^T \frac{\partial \hat{\dot{q}}}{\partial \dot{r}} \Phi_{11}$$

$$\dot{\Phi}_{11}^T \Phi_{21} = \Phi_{21}^T \Phi_{21}$$

or

$$\frac{d}{dt}\left(\Phi_{21}^T \Phi_{11} - \Phi_{11}^T \Phi_{21}\right) = 0 \tag{4.26a}$$

$$\Phi_{21}^T \Phi_{11} - \Phi_{11}^T \Phi_{21} = C_1 \quad . \tag{4.26b}$$

Similarly

$$\Phi_{22}^T \Phi_{11} - \Phi_{12}^T \Phi_{21} = C_2 \tag{4.27}$$

$$\Phi_{22}^T \Phi_{12} - \Phi_{12}^T \Phi_{22} = C_3 \tag{4.28}$$

$$\Phi_{11}^T \Phi_{22} - \Phi_{21}^T \Phi_{12} = C_4 \, . \tag{4.29}$$

Finally, the results of the integration can be restated in matrix notation as

$$\begin{bmatrix} \Phi_{22}^T & -\Phi_{12}^T \\ -\Phi_{22}^T & \Phi_{11}^T \end{bmatrix} \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{21} & \Phi_{22} \end{bmatrix} = \begin{bmatrix} C_2 & C_3 \\ -C_1 & C_4 \end{bmatrix} \tag{4.30}$$

and the constant arrays resulting from these integrations may now be evaluated by substituting the initial conditions

CR-1004

$$\Phi_{11}(0) = \Phi_{22}(0) = I$$

$$\Phi_{12}(0) = \Phi_{21}(0) = 0 \, .$$

This step produces

$$C_1 = O \tag{4.31a}$$

$$C_2 = I \tag{4.31b}$$

$$C_3 = O \tag{4.31c}$$

$$C_4 = I \tag{4.31d}$$

and reduces equation (4.30) to the identity matrix. But, the only matrix which can be utilized to reduce an arbitrary square matrix to I is its inverse. Thus,

$$\Phi^{-1}(t, t_o) = \begin{bmatrix} \Phi_{22}^T & -\Phi_{12}^T \\ -\Phi_{21}^T & \Phi_{11}^T \end{bmatrix} . \tag{4.32}$$

Equation (4.32) is important for general linear systems in that it provides an analytic means of constructing the inverse transition matrix directly from the elements of the known transition matrix by rearrangement of terms and the change of a few signs. In conclusion, it is noted that the true meaning of the terms A_{12} and A_{21} [equation (4.12)] was never employed, and only symmetry is required. Thus, there is an immediate generalization providing that an arbitrary system can be described by equation (4.22). (In general, this formulation is possible only in terms of inertial coordinates.)

2.4.1.2 By Integration of the Adjoint Equations

The differential equation for the adjoint matrix was shown to be

$$\dot{\Lambda}(t, t_o) = -\Lambda(t, t_o) A(t) \tag{4.33}$$

with boundary conditions

$$\Lambda(t_o, t_o) = I$$

Further, the inverse of the adjoint matrix was shown to be the transition matrix. Thus, the adjoint equation can be integrated and the result inverted by employing the analytic inverse property developed in the previous paragraphs to yield the transition matrix.

This technique offers little in the problem where observed sightings are being sequentially reduced (fixed t_o, variable T). However, if the problem also involves the generation of midcourse guidance commands (fixed T, variable t_o) so that both φ (t, t_o) and φ (T, t) are required, an adaptation of the adjoint technique is equally as suited to the problem as the more straightforward approach.

Define the time to go as

$$t_{go} = T - t \tag{4.34}$$

so that

$$dt_{go} = -dt$$

and

$$\dot{\tilde{\delta}}(t) = \frac{d\tilde{\delta}(t)}{dt_{go}} \frac{dt_{go}}{dt} \equiv -\mathring{\tilde{\delta}}(T - t_{go}). \tag{4.35}$$

Now, substituting into

$$\dot{\tilde{\delta}}(t) = A(t)\,\tilde{\delta}(t)$$

yields

$$\mathring{\tilde{\delta}}(T - t_{go}) = -A(T - t_{go})\,\tilde{\delta}(T - t_{go}). \tag{4.36}$$

But,

$$\tilde{\delta}(t) = \varphi^{-1}(T, t)\,\tilde{\delta}(T) \equiv \varphi(t, T)\,\tilde{\delta}(T).$$

Thus,

$$\mathring{\varphi}(T - t_{go}, T) = -A(T - t_{go})\,\varphi(T - t_{go}, T) \tag{4.37}$$

and the system of adjoint equations is

$$\mathring{\Lambda}(T - t_{go}, T) = \Lambda(T - t_{go}, T)\,A(T - t_{go}). \tag{4.38}$$

But at the epoch $t_{go} = 0$, φ (T, T) and thus φ^{-1}(T, T) or Λ (T, T) is the identity matrix

$$\Lambda(T, T) = I \quad .$$

Therefore, the adjoint equation expressed in terms of the time-to-go can be integrated from $t_{go} = 0$ (backward in real time) to the epoch in question to yield φ (t, T).

Note is made before leaving that this process involves successively small values of t_{go}. Thus, it is either necessary to store these successive values of Λ (t_i, T) or to construct these successive matrices as follows:

$$\hat{\delta}(t) = \varphi(t,T)\ \hat{\delta}(T)$$

$$\hat{\delta}(t_i) = \varphi(t_i,t)\ \delta(t)$$

Thus, $\hat{\delta}(t_i) = \varphi(t_i,t)\,\varphi(t,T)\,\hat{\delta}(T)$

$\equiv \varphi(t_i,T)\ \hat{\delta}(T)$

2.4.2 The State Transition Matrix For Conic Motion

The preceding paragraphs have presented several means of determining the state transition matrix for the true motion. However, the amount of effort required is considerably in excess of the amount normally spent in preliminary analysis, or in those cases where only moderate precision is desired. For this reason, an approximate logic for constructing this matrix will be developed based on the two-body solution of the equations of motion. This development is justified since the deviations from the true and conic trajectories will agree to a higher degree than the trajectories themselves (i.e., the error introduced by neglecting the variations in the perturbing forces is small compared to the magnitude of the variation in the state of the system produced by erroneous initial conditions).

The development of the matrices of partial derivatives to relate the first order variations (from the nominal conic trajectory) in the state at two arbitrary epochs will be accomplished in four steps:

1. Circular orbits (rotating and inertial coordinates)
2. Elliptic and hyperbolic orbits (rotating coordinates)
3. Elliptic and hyperbolic orbits (inertial coordinates)
4. Approximate method of including the effects of trajectory perturbation

As will be apparent, the material of these discussions is related and sometimes overlaps. However, as will also be noted, there are computational problems associated with some of the functions, and significant differences in the coordinate systems employed in the analyses which combine to suit a given formulation to a given task to a higher degree than is possible with any single step.

2.4.2.1 Circular Motion (Rotating and Inertial Coordinates)

Consider the perturbation equation

$$\vec{r} = \vec{r}_o + \Delta\vec{r}$$

and its second derivative with respect to time

$$\ddot{\vec{r}} = \ddot{\vec{r}}_o + \Delta\ddot{\vec{r}}$$

under the substitution of central force field dynamics.

$$\ddot{\vec{r}}_o + \Delta\ddot{\vec{r}} = -\frac{\mu}{r_o^3}\,\frac{r_o^3}{r^3}\,\vec{r}$$

$$= -\frac{\mu}{r_o^3}\,\frac{\vec{r}}{\left[1+3\left(\frac{\vec{r}_o\cdot\Delta\vec{r}}{r_o^2}\right)\right]} \tag{4.39}$$

$$= -\frac{\mu}{r_o^3}\left[1-3\left(\frac{\vec{r}_o\cdot\Delta\vec{r}}{r_o^2}\right)\right]\left(\vec{r}_o+\Delta\vec{r}\right)$$

$$= -\frac{\mu}{r_o^3}\left[\Delta\vec{r}-3\left(\hat{r}_o\cdot\Delta\vec{r}\right)\hat{r}_o\right]-\frac{\mu}{r_o^3}\,\vec{r}_o$$

Thus,

$$\ddot{\vec{\delta}} = -\frac{\mu}{r_o^3}\left[\Delta\vec{r}-3\left(\hat{r}_o\cdot\vec{\delta}\right)\hat{r}_o\right]$$

$$= -\frac{\mu}{r_o^3}\left[I-3\,\hat{r}_o\,\hat{r}_o^{T}\right]\Delta\vec{r} \tag{4.40}$$

where I (the identity matrix) and $\hat{r}_o\,\hat{r}_o^{T}$ are 3 by 3 matrices.

Now assuming that the eccentricity is small (i.e., $\dot{\theta}\equiv\omega=\frac{2\pi}{T}=\left(\frac{\mu}{r_o^3}\right)^{1/2}$) produces the equation

$$\Delta\ddot{\vec{r}}(t) = -\omega^2\left[I-3\,\hat{r}_o(t)\,\hat{r}_o^{T}(t)\right]\Delta\vec{r}(t) \tag{4.41}$$

But, as has been shown, this equation possesses the solution

$$\begin{Bmatrix} \Delta\vec{r} \\ \Delta\vec{v} \end{Bmatrix}_t = \varphi(t, t_o) \begin{Bmatrix} \Delta\vec{r} \\ \Delta\vec{v} \end{Bmatrix}_o$$

Thus, an explicit solution for φ (t, t_o) can be obtained (for circular motion) by noting (from the previous equation) that $\Delta\vec{r}$ can be expressed as

$$\Delta\vec{r}(t) = F_1(t, t_o)\,\hat{X} + F_2(t, t_o)\,\hat{Y} + F_3(t, t_o)\,\hat{Z}$$

by substituting this representation back into the differential equation (4.41). First, however, a simplifying assumption will be made in that the coordinates selected will be centered at the nominal position and rotating with the satellite so that

$$\hat{X} = \frac{\vec{r}_o(t)}{r_o(t)}$$
$$\hat{Z} = \vec{r}(0) \times \dot{\vec{r}}_o(0)$$
$$\hat{Y} = \hat{Z} \times \hat{X}$$

and

$$\dot{\hat{X}} = \omega\hat{Y}$$
$$\dot{\hat{Y}} = -\omega\hat{X}$$
$$\dot{\hat{Z}} = 0$$

At this point, the derivatives of $\Delta\vec{r}$ (t) are formed and substitution into the differential equation will be accomplished.

$$\begin{aligned}
\Delta\dot{\vec{r}} &= \dot{F}_1\hat{X} + F_1\dot{\hat{X}} + \dot{F}_2\hat{Y} + F_2\dot{\hat{Y}} + \dot{F}_3\hat{Z} + F_3\dot{\hat{Z}} \\
&= (\dot{F}_1 - \omega F_2)\hat{X} + (\dot{F}_2 + \omega F_1)\hat{Y} + \dot{F}_3\hat{Z} \\
\Delta\ddot{\vec{r}} &= (\ddot{F}_1 - \omega\dot{F}_2)\hat{X} + (\dot{F}_2 + \omega\dot{F}_1)\hat{Y} + \dot{F}_3\hat{Z} \\
&\quad + (\dot{F}_1 - \omega F_2)\omega\hat{Y} - (\dot{F}_2 + \omega F_1)\omega\hat{X} \\
&= (\ddot{F}_1 - 2\omega\dot{F}_2 - \omega^2 F_1)\hat{X} \\
&\quad + (\ddot{F}_2 + 2\omega\dot{F}_1 - \omega^2 F_2)\hat{Y} + \ddot{F}_3\hat{Z}.
\end{aligned} \tag{4.42}$$

Now

$$3\hat{r}_0\hat{r}_0^T = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and

$$-\omega^2(I-3\hat{r}_0\hat{r}_0^T) = \begin{bmatrix} +2\omega^2 & 0 & 0 \\ 0 & -\omega^2 & 0 \\ 0 & 0 & -\omega^2 \end{bmatrix}. \tag{4.43}$$

So, equating the components of equations (4.42) and (4.43) yields

$$\ddot{F}_1 - 2\omega\dot{F}_2 - \omega^2 F_1 = 2\omega^2 F_1$$

$$\ddot{F}_2 + 2\omega\dot{F}_1 - \omega^2 F_2 = -\omega^2 F_2$$

$$\ddot{F}_3 = -\omega^2 F_3$$

or

$$\ddot{F}_1 - 3\omega^2 F_1 - 2\omega\dot{F}_2 = 0 \tag{4.44a}$$

$$\ddot{F}_2 + 2\omega\dot{F}_1 = 0 \tag{4.44b}$$

$$\ddot{F}_3 + \omega^2 F_3 = 0 \quad . \tag{4.44c}$$

These three second order scalar equations define the time dependent coefficients of $\Delta\vec{r}$.

Note that the third of these equations is uncoupled from the first two, and thus may be integrated directly. The solution is:

$$F_3(t) = S\cos\omega t + T\sin\omega t. \tag{4.45a}$$

But

$$F_3(t) = F_3(0) \quad , \quad t = 0$$
$$\dot{F}_3(t) = \dot{F}_3(0) \quad , \quad t = 0 \quad .$$

Thus,

$$S = F_3(0) \tag{4.45b}$$
$$T = \dot{F}_3(0)/\omega \tag{4.45c}$$

Now, integrating the second equation

$$\dot{F}_2(t) = -2\omega F_1(t) + \dot{F}_2(0) + 2\omega F_1(0)$$

and substituting into the first allows the following solution to be developed:

$$\ddot{F}_1 - 3\omega^2 F_1 - 2\omega\left(-2\omega F_1 + \dot{F}_2(0) + 2\omega F_1(0)\right) = 0$$
$$\ddot{F}_1 + \omega^2 F_1 = 2\omega \dot{F}_2(0) + 4\omega^2 F_1(0) \tag{4.46a}$$
$$F_1(t) = S' \cos\omega t + T' \sin\omega t + \frac{2}{\omega}\dot{F}_2(0) + 4F_1(0) \quad ,$$

where

$$F_1(t) = F_1(0) \quad , \quad t = 0$$
$$\dot{F}_1(t) = \dot{F}_1(0) \quad , \quad t = 0 \quad .$$

Thus,

$$S' = -\frac{2}{\omega}\dot{F}_2(0) - 3F_1(0) \tag{4.46b}$$
$$T' = \frac{\dot{F}_1(0)}{\omega} \tag{4.46c}$$

and

$$F_1(t) = F_1(0)\left[4 - 3\cos\omega t\right] + \dot{F}_2(0)\left[\frac{2}{\omega} - \frac{2}{\omega}\cos\omega t\right] + \dot{F}_1(0)\left[\frac{1}{\omega}\sin\omega t\right] . \tag{4.47}$$

Substituting this latter equation into the first integral of $\ddot{F}_2$ yields

$$\dot{F}_2(t) = -2\omega\left[F_1(0)(4-3\cos\omega t) + 2\frac{\dot{F}_2(0)}{\omega}(1-\cos\omega t) + \frac{\dot{F}_1(0)}{\omega}\sin\omega t\right] + \dot{F}_2(0) + 2\omega F_1(0)\,\Delta\dot{F}$$

or

$$\dot{F}_2(t) = F_1(0)\left[-6\omega(1-\cos\omega t)\right] + \dot{F}_2(0)\left[-(3-4\cos\omega t)\right] - 2\dot{F}_1(0)\left[\sin\omega t\right] . \tag{4.48}$$

Thus,

$$\begin{aligned} F_2(t) &= -F_1(0)\left[6\omega t - 6\sin\omega t\right] - \dot{F}_2(0)\left[3t - \frac{4}{\omega}\sin\omega t\right] + \frac{2\dot{F}_1(0)}{\omega}\cos\omega t + C \\ &= F_1(0)\left[6(\sin\omega t - \omega t)\right] + \dot{F}_2(0)\left[\frac{1}{\omega}(4\sin\omega t - 3\omega t)\right] + \dot{F}_1(0)\left[\frac{2}{3}\cos\omega t\right] + C . \end{aligned} \tag{4.49}$$

Finally, employing the boundary conditions

$$F_2(t) = F_2(0) \quad , \quad t = 0$$

yields

$$\dot{F}_1(0)\frac{2}{\omega} + C = F_2(0)$$

or

$$C = F_2(0) - \frac{2\dot{F}_1(0)}{\omega} .$$

Thus,

$$\begin{aligned} F_2(t) = {} & F_1(0)\left[6(\sin\omega t - \omega t)\right] \\ & + \dot{F}_2(0)\left[\frac{1}{\omega}(4\sin\omega t - 3\omega t)\right] \\ & + \dot{F}_1(0)\left[\frac{2}{\omega}(\cos\omega t - 1)\right] + F_2(0) \quad , \end{aligned} \tag{4.50}$$

The extension of this analysis to the case in which inertial rather than rotating coordinates are utilized is obtained simply by noting that the first derivatives of $\vec{\delta}$ can be expressed as

$$\Delta\dot{\vec{F}} = A\hat{\chi} + B\hat{Y} + C\hat{Z} \tag{4.51}$$

where

$$\left.\begin{aligned} A &= \dot{F}_1 \\ B &= \dot{F}_2 \\ C &= \dot{F}_3 \\ \hat{\chi} &= \hat{F}_0(t) \\ \hat{Z} &= \hat{F}_0(0) \times \dot{\hat{F}}_0(0) \\ \hat{Y} &= \hat{Z} \times \hat{\chi} \end{aligned}\right\} \text{for rotating coordinates}$$

or where

$$A = \dot{F}_1 - \omega F_2$$

$$B = \dot{F}_2 + \omega F_1$$

$$C = \dot{F}_3$$

$$\hat{X}, \hat{Y}, \hat{Z} = \hat{r}_0(0), \hat{Z} \times \hat{X}, \hat{r}_0(0) \times \dot{\hat{r}}(0).$$

} inertial coordinates

In this latter case, of course, the initial conditions for the functions F_1 . . . F_3 become

$$F_1(0) = \Delta\vec{r}(0) \cdot \hat{X}(0) \tag{4.52a}$$

$$F_2(0) = \Delta\vec{r}(0) \cdot \hat{Y}(0) \tag{4.52b}$$

$$F_3(0) = \Delta\vec{r}(0) \cdot \hat{Z}(0) \tag{4.52c}$$

$$\dot{F}_1(0) = \Delta\dot{\vec{r}}(0) \cdot \Delta\hat{X}(0) + \omega F_2(0) \tag{4.52d}$$

$$\dot{F}_2(0) = \Delta\dot{\vec{r}}(0) \cdot \hat{Y}(0) - \omega F_1(0) \tag{4.52e}$$

$$\dot{F}_3(0) = \Delta\dot{\vec{r}}(0) \cdot \hat{Z}(0) \tag{4.52f}$$

The results of this process are presented in Tables 2.4.1 and 2.4.2 for the cases of rotating and inertial coordinates, respectively.

THE STATE TRANSITION MATRIX FOR CIRCULAR MOTION - ROTATING COORDINATES

Table 2.4.1

$$\begin{Bmatrix} dr \\ dy \\ dz \\ dv \\ d\dot{r} \\ d\dot{z} \end{Bmatrix}_2 = \begin{bmatrix} 4-3\cos\omega T & 0 & 0 & \frac{2}{\omega}(1-\cos\omega T) & \frac{1}{\omega}\sin\omega T & 0 \\ 6(\sin\omega T-\omega T) & 1 & 0 & \frac{1}{\omega}(4\sin\omega T-3\omega T) & \frac{2}{\omega}(\cos\omega T-1) & 0 \\ 0 & 0 & \cos\omega T & 0 & 0 & \frac{1}{\omega}\sin\omega T \\ 6n(\cos\omega T-1) & 0 & 0 & 4\cos\omega T-3 & -2\sin\omega T & 0 \\ 3\omega\sin\omega T & 0 & & 2\sin\omega T & \cos\omega T & 0 \\ 0 & 0 & -\omega\sin\omega T & 0 & 0 & \cos\omega T \end{bmatrix} \begin{Bmatrix} dr \\ dy \\ dz \\ dv \\ d\dot{r} \\ d\dot{z} \end{Bmatrix}$$

$$T = t - t_0$$

$$\omega = \frac{2\pi}{\tau} = \frac{\mu}{r_0^3}$$

THE STATE TRANSITION MATRIX FOR CIRCULAR MOTION - INERTIAL COORDINATES

Table 2.4.2

$$\begin{Bmatrix} dx \\ dy \\ dz \\ d\dot{x} \\ d\dot{y} \\ d\dot{z} \end{Bmatrix} = \begin{bmatrix} 2-\cos\omega t & \sin\omega t & 0 & \frac{\sin\omega t}{\omega} & \frac{2}{\omega}(1-\cos\omega t) & 0 \\ 2\sin\omega t - 3\omega t & 2\cos\omega t - 1 & 0 & \frac{2\cos\omega t - 2}{\omega} & \frac{1}{\omega}(4\sin\omega t - 3\omega t) & 0 \\ 0 & 0 & \cos\omega t & 0 & 0 & \frac{\sin\omega t}{\omega} \\ \omega(3\omega t - \sin\omega t) & \omega(1-\cos\omega t) & 0 & 2-\cos\omega t & 3\omega t - 2\sin\omega t & 0 \\ \omega(\cos\omega t - 1) & -\omega(\sin\omega t) & 0 & -\sin\omega t & 2\cos\omega t - 1 & 0 \\ 0 & 0 & -\omega\sin\omega t & 0 & 0 & \cos\omega t \end{bmatrix} \begin{Bmatrix} dx \\ dy \\ dz \\ d\dot{x} \\ d\dot{y} \\ d\dot{z} \end{Bmatrix}$$

$$\hat{x} = \hat{r}_o(0)$$

$$\hat{y} = \hat{r}_o(0) \times \dot{\hat{r}}_o(0) \times \hat{r}_o(0)$$

$$\hat{z} = \hat{r}_o(0) \times \dot{\hat{r}}_o(0)$$

2.4.2.2 Elliptic and Hyperbolic Motion (Rotating Coordinates)

2.4.2.2.1 Elliptic Motion. The complete description of the motion in an elliptic orbit is obtained from the following set of equations:

r	$= a(1 - e\cos E)$	$=$	radial distance
a	$= ru(2u - rv^2)^{-1}$	$=$	semimajor axis
p	$= (rv\cos\gamma)^2 u^{-1}$	$=$	semiparameter
e	$= (1 - \frac{p}{a})^{1/2}$	$=$	eccentricity
$(t-t_p)$	$= (E - e\sin E)n^{-1}$	$=$	time from periapse
n	$= u^{1/2}a^{-3/2}$	$=$	mean motion
$(\varphi - \omega)$	$= \cos^{-1}\left(\frac{p-r}{re}\right)$	$=$	true anomaly
φ	$=$ angle in the plane of motion measured from the ascending node		
ω	$=$ argument of periapse		
i	$= \cos^{-1}(\cos\angle\cos\beta)$	$=$	inclination
ν	$= \cos^{-1}\left(\frac{\cos\beta}{\sin i}\right)$	$=$	longitude relative to the ascending node
β	$=$ azimuth		

Thus, by straightforward differentiation, the column vector representing the change in the elements can be constructed for a given set of errors occurring in the initial position and velocity vectors. Now, since the modified constants of the motion (a + Δa, etc) are known, the implied variations in the position and velocity vectors at all subsequent times can be obtained by solving the six simultaneous equations for these quantities in terms of the known errors in the elements. This process, like the first, is direct though slightly involved.

The transition matrix, i.e., the matrix of partials relating errors in the state vector at any epoch to errors in the state at a given epoch, is readily constructed from these two matrices by direct substitution; that is,

$$\delta\vec{C} = \Theta_1(t_1, T)\,\vec{\delta}_1$$

$$\vec{\delta}_2 = \Theta_2(t_2, T)\,\delta\vec{C} \tag{4.53}$$

$$= \Theta_2\,\Theta_1\,\delta\vec{x}_1$$

$$\equiv \Phi(t_2, t_1)\,\vec{\delta}_1 \,. \tag{4.54}$$

Thus, the analysis will turn to the development of the two matrices denoted $[\Theta]_2$ and $[\Theta]_1$.

2.4.2.2.1.1 $\Delta\vec{r}, \Delta\vec{v} = f(\Delta a, \Delta e \ldots \Delta\Omega)$. Expansion of those equations required to define the components of $\Delta\vec{r}$ and $\Delta\vec{v}$ in terms of variations in the elements will be presented on the following pages. These expansions and the associated substitutions are presented in detail to facilitate a rapid understanding of the process, to demonstrate the manner in which computational difficulties can arise, and to establish the means of inverting the process to define the changes in the elements as a function of the changes in the position and velocity.

Consider first the radial distance

$$r = a(1 - e\cos E) \quad . \tag{4.55}$$

Thus,

$$dr = \frac{r}{a}da + ae\sin E\, dE - a\cos E\, de$$

But, t rather than E is the independent variable of the analysis. Thus Kepler's equation must be differentiated.

$$t_p = t - \frac{1}{n}(E - e\sin E)$$

$$dt_p = -\frac{1}{n}\frac{r}{a}dE + \frac{1}{n}\sin E\, de - \frac{3}{2}\frac{(t-t_p)}{a}da \quad .$$

Now, substitution yields

$$\begin{aligned} dr = &\left[\frac{r}{a} - \frac{3}{2}\frac{ea}{r}M\sin E\right]da \\ &+ \left[-a\cos E + \frac{a^2e}{r}\sin^2 E\right]de \\ &+ \left[-\frac{na^2e}{r}\sin E\right]dt_p \quad . \end{aligned} \tag{4.56}$$

Now, consider the displacement in the circumferential direction (i.e., perpendicular to r but in the plane of motion).

$$\varphi = \theta + \omega$$

$$rd\varphi = rd\theta + rd\omega \tag{4.57}$$

where

$$\cos\theta = \frac{a(1-e^2)-r}{re}$$

or

$$-\sin\theta\, d\theta = \frac{1}{re}\left[(1-e^2)\,da - 2ae\,de - dr\right] - \frac{p-r}{r^2e^2}\left[r\,de + e\,dr\right] \quad .$$

Substituting for dr in terms of dE reduces this equation to

$$r\,d\theta = \frac{1}{\sin\theta}\left\{-\left[\frac{(r-a)^2-a^2e^2}{re^2}\,de + \frac{pa\sin E}{r}\,dE\right]\right\} \quad .$$

Now, substituting for dE yields

$$r\,d\theta = \left[-\frac{3}{2}\,\frac{p}{r}\,\frac{M}{\sqrt{1-e^2}}\right]da + \left[-\frac{(r-a)^2-a^2e^2}{\sin\theta\,(re^2)} + \frac{pa}{r}\,\frac{\sin E}{\sqrt{1-e^2}}\right]de + \left[-\frac{pan}{r\sqrt{1-e^2}}\right]dt_p \qquad (4.58)$$

The next step in the analysis is the construction of the errors in the planar components of the velocity vector. The rotating axes for this application will be selected along, and perpendicular to, the direction of the velocity itself. Consider the energy equation

$$V^2 = \frac{2\mu}{r} - \frac{\mu}{a} \qquad (4.59)$$

$$dV = \frac{1}{2V}\left[-\frac{2\mu}{r^2}\,dr + \frac{\mu}{a^2}\,da\right]$$

substituting for dr

$$2V dV = \left[-\frac{2\mu}{ra} + \frac{\mu}{a^2} + \frac{3}{2} \frac{eaM}{r} \sin E \left(\frac{2\mu}{r^2} \right) \right] da$$

$$+ \left[+ \frac{2\mu}{r^2} a \cos E - \frac{2\mu}{r^2} \frac{a^2 e}{r} \sin^2 E \right] de$$

$$+ \left[\frac{2\mu}{r^2} \frac{na^2 e}{r} \sin E \right] dt_p$$

or

$$dV = \left[-\frac{V}{2a} + \frac{3}{2} \frac{\mu}{r^2 V} \frac{a}{r} eM \sin E \right] da$$

$$+ \left[\left(\frac{\mu a}{r^2 V} \right) \left(\cos E - \frac{a}{r} e \sin^2 E \right) \right] de \qquad (4.60)$$

$$+ \left[\left(\frac{\mu a}{r^2 V} \right) \frac{a}{r} en \sin E \right] dt_p .$$

Finally, consideration of the angular momentum will provide the information necessary to construct the variation in the velocity perpendicular to the nominal direction

$$r^2 V^2 \cos^2 \gamma = \mu a (1 - e^2)$$

$$2 \frac{\mu p}{r} dr + \frac{2\mu p}{V} dV - 2\mu p \tan \gamma \, d\gamma \qquad (4.61)$$

$$= 2\mu p \left[\frac{(1 - e^2)}{2p} da - \frac{ae}{p} de \right]$$

so that

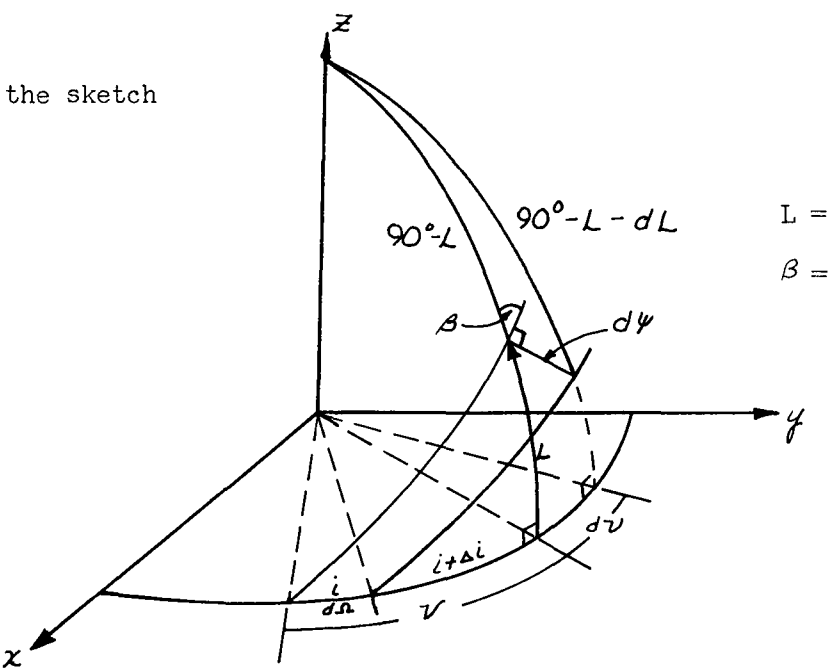
$$V d\gamma = \frac{V}{\tan \gamma} \left[\frac{dr}{r} + \frac{dV}{V} + \frac{ae}{p} de - \frac{da}{2a} \right] .$$

Substitution of previously derived expressions and the identity $\sin E/\tan\gamma = \sqrt{1-e^2}/e$ now reduces this final equation to

$$
\begin{aligned}
V d\gamma = &\left[-\frac{3}{2}\frac{VaM}{r^2}\sqrt{1-e^2}\left(1-\frac{\mu}{rV^2}\right)\right] da \\
&+\left[-\frac{a}{r}\frac{V\cos\theta}{\tan\gamma}\left(1-\frac{\mu}{rV^2}\right)+\frac{Ve}{\tan\gamma(1-e^2)}\right] de \\
&+\left[-Vn\left(\frac{a}{r}\right)^2\sqrt{1-e^2}\left(1-\frac{\mu}{rV^2}\right)\right] dt_p \quad .
\end{aligned}
\tag{4.62}
$$

The development of the variation in the out-of-plane angle and the azimuth (or the corresponding position and velocity perpendicular to the plane of motion) will be accomplished in a slightly different manner (Ref. 4.1). Since the dimensionality of this subset of variations is two, and since these variations are independent of the in-plane variations, this set and the inverse set [i.e., Δi, $\Delta\Omega = f(\Delta\beta, \Delta\psi)$] will be derived simultaneously in the following paragraphs.

Consider the sketch



L = Latitude

β = Azimuth

Now, employing the law of cosines for the spherical triangle involving the colatitudes and $d\psi$

$$
\begin{aligned}
\cos(90-L-dL) = &\cos(90-L)\cos d\psi \\
&+\sin(90-L)\sin d\psi\cos(90-\beta)
\end{aligned}
$$

which is approximately

$$\sin L + \cos L \, dL = \sin L - \cos L \sin \beta \, d\psi$$

or

$$dL = -\sin \beta \, d\psi. \tag{4.63}$$

But, from spherical trigonometry

$$\cos i = \cos L \sin \beta$$

so that

$$\begin{aligned} di &= -\frac{\cos L \cos \beta}{\sin i} d\beta + \frac{\sin L \sin \beta}{\sin i} dL \\ &= -\frac{\cos L \cos \beta}{\sin i} d\beta - \frac{\sin L \sin^2 \beta}{\sin i} d\psi. \end{aligned} \tag{4.64}$$

The error in the node ($d\Omega$) is obtained by considering the projection of the erroneous trajectory on the reference plane, i.e.

$$\cos(\nu + d\nu - d\Omega) = \frac{\cos(\beta + d\beta)}{\sin(i + di)} \tag{4.65}$$

which is approximately

$$\begin{aligned} \cos \nu - \sin \nu (d\nu - d\Omega) &= \frac{\cos \beta - \sin \beta \, d\beta}{\sin i + \cos i \, di} \\ &\approx \frac{(\cos \beta - \sin \beta \, d\beta)}{\sin i} (1 - \cot i \, di) \end{aligned}$$

or

$$\sin \nu \, (d\nu - d\Omega) = \frac{\sin \beta}{\sin i} d\beta + \frac{(\cos \beta \cot i)}{\sin i} di$$

But

$$\sin(d\nu) = \frac{\sin(90 + \beta)}{\sin(90 - L - \Delta L)} \sin d\psi$$

so

$$d\nu \approx \frac{\cos\beta}{\cos L}\, d\psi . \tag{4.66}$$

Thus,

$$d\Omega = -\frac{\sin\beta}{\sin\nu \sin i}\, d\beta - \frac{\cos\beta \cot i}{\sin i \sin\nu}\, di + \frac{\cos\beta}{\cos L}\, d\psi$$

or

$$\begin{aligned} d\Omega = &\left[-\frac{\tan\beta}{\tan\nu} + \frac{\cos\beta \cot i \cos L}{\sin i \tan\nu}\right] d\beta \\ &+ \left[\frac{\cos\beta}{\cos L} + \frac{\cot i \sin L \sin^2\beta}{\tan\nu \sin i}\right] d\psi . \end{aligned} \tag{4.67}$$

These variations can be written in the form

$$\begin{bmatrix} di \\ d\Omega \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} d\beta \\ d\psi \end{bmatrix} \tag{4.68}$$

Thus, solving for the column vector ($d\beta$, $d\psi$) is possible by computing the inverse matrix

$$A^{-1} = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \frac{1}{|A|}$$

The results of this process are

$$\begin{aligned} d\beta = &\left[-\sin L \sin^2\beta\right] d\Omega \\ &+ \left[-\cos\beta\left(\frac{\sin i}{\cos L} + \frac{\cos i \sin\beta}{\sin L}\right)\right] di \end{aligned} \tag{4.69}$$

and

$$\begin{aligned} d\psi = &\left[\cos L \cos\beta\right] d\Omega \\ &+ \left[\frac{\sin i}{\sin L}\left(\cot^2 i \cot^2\beta - 1\right)\right] di \end{aligned} \tag{4.70}$$

These results are summarized in Table 2.4.3.

THE LINEAR RELATIONSHIPS BETWEEN ELEMENT ERRORS AND STATE ERRORS FOR ELLIPTIC MOTION - ROTATING COORDINATES

Table 2.4.3

$$
\begin{Bmatrix} dr \\ rd\phi \\ rd\psi \\ dv \\ vd\gamma \\ vd\beta \end{Bmatrix} =
\begin{bmatrix}
\frac{r}{a} - \frac{3aeM}{2r}\sin E & -a\cos E + \frac{a^2 e}{r}\sin^2 E & \frac{-a^2 en}{r}\sin E & 0 & 0 & 0 \\
\frac{-3pM}{2r\sqrt{1-e^2}} & \frac{a\sin E}{\sqrt{1-e^2}}\left(1+p/r\right) & \frac{-pan}{r\sqrt{1-e^2}} & r & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{r\sin i}{\sin L}\left(\operatorname{ctn}^2 i \operatorname{ctn}^2\beta - 1\right) & r\cos L\cos\beta \\
\frac{-V}{2a} + \frac{3\mu eaM}{2r^3 V}\sin E & \frac{\mu a\cos E}{r^2 V} - \frac{\mu a^2 e\sin^2 E}{r^3 V} & \frac{\mu a^2 en}{r^3 V}\sin E & 0 & 0 & 0 \\
\frac{-3vaM\sqrt{1-e^2}}{2r^2}\left(1-\mu/rv^2\right) & \frac{-av\cos\theta}{r\tan\gamma}\left(1-\mu/rv^2\right) + \frac{ve}{\tan\gamma(1-e^2)} & \frac{-va^2 n\sqrt{1-e^2}}{r^2}\left(1-\mu/rv^2\right) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -v\cos\beta\left(\frac{\sin i}{\cos L} + \operatorname{ctn} i \sin\beta\right) & -v\sin L\sin^2\beta
\end{bmatrix}
\begin{Bmatrix} da \\ de \\ dtp \\ d\omega \\ di \\ d\Omega \end{Bmatrix}
$$

$M = n(t - t_p)$

THE LINEAR RELATIONSHIPS BETWEEN STATE ERRORS AND ELEMENT ERRORS FOR ELLIPTIC MOTION - ROTATING COORDINATES

Table 2.4.4

$$
\begin{Bmatrix} da \\ de \\ dt_p \\ d\omega \\ di \\ d\Omega \end{Bmatrix} =
\begin{bmatrix}
\frac{2a^2}{r^2} & 0 & 0 & \frac{2va^2}{\mu} & 0 & 0 \\
\cos P/r^2 & 0 & 0 & 2\cos E\, P/rV & r\sin\theta/Va & 0 \\
-\frac{3a}{r}(t-t_p) + \frac{\sin E}{nr}\left(P/r\cos E + \frac{1}{e}\right) & 0 & 0 & -\frac{3Va}{\mu}(t-t_p) + \frac{2\sin E}{nV}\left(P/r\cos E + \frac{1}{e}\right) & \frac{p\sin^2 E}{Vna\sqrt{1-e^2}} - \frac{r\sqrt{1-e^2}\cos E}{Vnae} & 0 \\
\frac{\sin\theta}{er} & \frac{1}{r} & 0 & \frac{2\sin\theta}{eV} & -\frac{1}{V} - \frac{\cos E}{Ve} & 0 \\
0 & 0 & -\frac{\sin L\sin^2\beta}{r\sin i} & 0 & 0 & -\frac{\cos L\cos\beta}{V\sin i} \\
0 & 0 & \left(\frac{\cos\beta}{\cos L} + \frac{\cos i\cos\beta\sin\beta}{\sin^2 i}\right) & 0 & 0 & \frac{ctn^2 i\, ctn^2\beta - 1}{V\sin L}
\end{bmatrix}
\begin{Bmatrix} dr \\ rd\varphi \\ rd\psi \\ dV \\ Vd\gamma \\ Vd\beta \end{Bmatrix}
$$

2.4.2.2.1.2 (Δa, Δe . . . Δe) = ($\Delta\vec{r}$, $\Delta\vec{v}$). The derivatives required to construct the linear equations representing the inverse of those presented in Table 2.4.3 were, for the most part, presented in the preceding section. Thus, the development of this section will be much simplified.

The first variation of interest will be in the semimajor axis. This equation was presented intact as:

$$da = \frac{2Va^2}{\mu} dV + \frac{2a^2}{r^2} dr \tag{4.71}$$

The variation in eccentricity is from the equation defining the variation in the flight path angle, simply by substituting to eliminate da from the equation. That is,

$$de = \frac{p}{ae}\left[\frac{da}{2a} + TAN\,\gamma\, d\gamma - \frac{dV}{V} - \frac{dr}{r}\right]$$

or, upon substituting for da

$$\begin{aligned} de &= \frac{p}{ae}\left[\left(1-\frac{a}{r}\right)\frac{dr}{r} + \left(1-\frac{aV^2}{\mu}\right)\frac{dV}{V} - TAN\,\gamma\, d\gamma\right] \\ &= \frac{p}{r^2} COS\,E\, dr + \frac{2p}{rV} COS\,E\, dV + \frac{r}{Va} SIN\,\theta\,(V d\gamma)\ . \end{aligned} \tag{4.72}$$

The development of the change in the epoch of periapse passage initiates with Kepler's equation in the form (presented previously)

$$dt_p = -\frac{3}{2}\frac{M}{n_a} da + \frac{SIN\,E}{n} de - \frac{r}{n_a} dE\ .$$

This time, however, it is necessary to substitute for dE in terms of dr, dv . . . . Thus,

$$r = a(1 - e\,COS\,E)$$

$$dE = \frac{1}{ae\,SIN\,E}\left[\frac{dr}{a} - r\,da + a\,COS\,E\,de\right]$$

or

$$dt_p = \left[-\frac{3}{2}\frac{M}{na} + \frac{r^2}{na^3 e \sin E} \right] da + \left[\frac{\sin E}{na} - \frac{r \cos E}{nae \sin E} \right] de$$

$$+ \left[-\frac{r}{na^2 e \sin E} \right] dr .$$

Now, eliminating da and de from the equation by substitution, and combining terms yields

$$dt_p = \left[-\frac{3Ma}{nr^2} + \frac{\sin E}{nr}\left(\frac{p}{r}\cos E + \frac{1}{e}\right) \right] dr$$

$$+ \left[-\frac{3M}{n\mu} Va + \frac{2 \sin E}{nV}\left(\frac{p}{r}\cos E + \frac{1}{e}\right) \right] dV \qquad (4.73)$$

$$+ \left[\frac{p \sin^2 E}{nVa\sqrt{1-e^2}} - \frac{r\sqrt{1-e^2}}{nVae} \cos E \right] (V d\gamma) .$$

The final equation required is obtained by noting that

$$\phi = \theta + \omega$$

so that

$$\cos(\phi - \omega) = a\left(\frac{1-e^2}{er}\right) - r$$

or

$$d\omega = \frac{1}{\sin\theta}\left[\frac{p}{ear} da + \left(\frac{-2a}{r} - \frac{p-r}{e^2 r} \right) de \right.$$

$$\left. + \left(\frac{-1}{er} - \frac{p-r}{er^2} \right) dr \right] + d\phi \qquad (4.74)$$

Now, once again substituting for da and de yields

$$d\omega = \frac{1}{SIN\,\theta}\left\{\left[\frac{p}{aer}\left(\frac{2a^2}{r^2}\right) - \frac{p}{er^2} - \left(\frac{2a}{r} + \frac{p-r}{e^2 r}\right)\frac{(a-r)p}{aer^2}\right] dr \right.$$

$$+ \left[\frac{p}{aer}\left(\frac{2Va^2}{\mu}\right) - \left(\frac{2a}{r} + \frac{p-r}{e^2 r}\right)\frac{2(a-r)p}{rVae}\right] dV$$

$$\left. + \left[-\left(\frac{2a}{r} + \frac{p-r}{e^2 r}\right)\left(\frac{p}{ea}\,TAN\,\gamma\right)\right] d\gamma\right\} + d\phi$$

which reduces to the following

$$d\omega = \frac{SIN\,\theta}{er}\,dr + \frac{2\,SIN\,\theta}{eV}\,dV - \left[\frac{1}{V} + \frac{COS\,E}{Ve}\right](V d\gamma) + \frac{1}{r}(r d\phi) \quad . \tag{4.75}$$

The results of this analysis are summarized in Table 2.4.4. Note that in **contrast** to all previous equations, the equations for dt_p and dw contain terms of with factors (1/e). This fact displays the physical problem associated with the indeterminancy of the line of apsides. This approach, nonetheless, completely describes the propagation of the differentials around the orbit, as can be seen by examining the limits of the terms contained in the product matrix (matrix of Table 2.4.3 times the matrix of Table 2.4.4) and comparing the results with the terms of the matrix presented in Table 2.4.1.

2.4.2.2.2 Hyperbolic Motion. The equations presented in Tables 2.4.3 and 2.4.4 can be applied to the case of hyperbolic motion when the following substitutions are made.

$$E = -iF$$

$$\sin E = i \sinh F$$

$$\cos E = \cosh F$$

$$n = \sqrt{\frac{\mu}{a^3}} = i\sqrt{\frac{\mu}{a_h}} = i n_h$$

$$a_h = -a$$

$$\sqrt{1-e^2} = i\sqrt{e^2-1}$$

The results of this set of substitutions are presented in Tables 2.4.5 and 2.4.6.

THE LINEAR RELATIONSHIPS BETWEEN ELEMENT ERRORS AND STATE ERRORS FOR HYPERBOLIC MOTION - ROTATING COORDINATES

Table 2.4.5

$$
\begin{Bmatrix} dr \\ rd\varphi \\ rd\psi \\ dv \\ vd\gamma \\ vd\beta \end{Bmatrix} =
\begin{bmatrix}
\frac{r}{a} + \frac{3aeM}{2r}\sinh F & a\cosh F - \frac{a^2 e}{r}\sinh^2 F & \frac{a^2 ne}{r}\sinh F & 0 & 0 & 0 \\
\frac{3PM}{2r\sqrt{e^2-1}} & -\frac{a\sinh F}{\sqrt{e^2-1}}\left(1 + P/r\right) & \frac{pan}{r\sqrt{e^2-1}} & r & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{r\sin L}{\sin i}\left(\operatorname{ctn}^2 i \operatorname{ctn}^2\beta - 1\right) & r\cos L\cos\beta \\
-\frac{V}{2a} - \frac{3\mu aeM}{2r^3 V}\sinh F & -\frac{\mu a\cosh F}{r^2 V} + \frac{\mu a^2 e}{r^3 V}\sinh^2 F & -\frac{\mu a^2 en}{r^3 V}\sinh F & 0 & 0 & 0 \\
\frac{3VaM\sqrt{e^2-1}}{2r^2}\left(1 - \mu/rV^2\right) & \frac{qv\cos\theta\left(1 - \mu/rV^2\right)}{r\tan\gamma} - \frac{Ve}{\tan\gamma\left(e^2-1\right)} & \frac{va^2 n\sqrt{e^2-1}}{r^2}\left(1 - \mu/rV^2\right) & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -V\cos\beta\left(\frac{\sin i}{\cos L} + \operatorname{ctn} i \sin\beta\right) & -V\sin L\sin^2\beta
\end{bmatrix}
\begin{Bmatrix} da \\ de \\ dt_p \\ d\omega \\ di \\ d\Omega \end{Bmatrix}
$$

THE LINEAR RELATIONSHIPS BETWEEN STATE ERRORS AND ELEMENT ERRORS FOR HYPERBOLIC MOTION - ROTATING COORDINATES

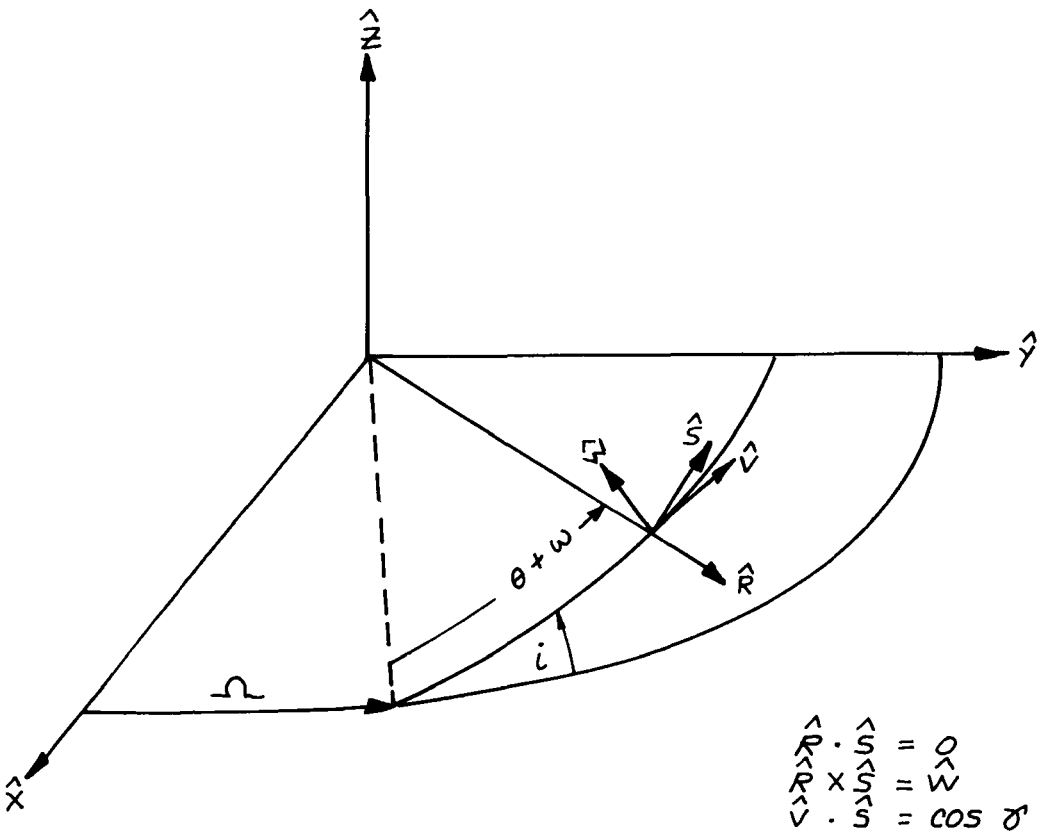
Table 2.4.6

$$
\begin{Bmatrix} da \\ de \\ dt_p \\ dw \\ di \\ d\Omega \end{Bmatrix} =
\begin{bmatrix}
-\frac{2a^2}{r^2} & 0 & 0 & -\frac{2va^2}{\mu} & 0 & 0 \\
\cosh F\,\frac{p}{r^2} & 0 & 0 & 2\cosh F\,\frac{p}{rv} & -\frac{r\sin\theta}{va} & 0 \\
\frac{3a}{r^2}(t-t_p) + \frac{\sinh F}{nr}\left(\frac{p}{r}\cosh F + \frac{1}{e}\right) & 0 & 0 & \frac{3va}{\mu}(t-t_p) + \frac{\sinh F}{vn}\left(\frac{p}{r}\cosh F + \frac{1}{e}\right) & -\frac{p\sinh^2 F}{vna\sqrt{e^2-1}} + \frac{r\sqrt{e^2-1}}{vnae}\cosh F & 0 \\
\frac{\sin\theta}{er} & \frac{1}{r} & 0 & \frac{2\sin\theta}{ev} & -\frac{1}{v} - \frac{\cosh F}{ve} & 0 \\
0 & 0 & -\frac{\sin L\sin^2\beta}{r\sin i} & 0 & 0 & -\frac{\cos L\cos\beta}{v\sin i} \\
0 & 0 & \left(\frac{\cos\beta}{\cos L} + \frac{\cos i\cos\beta\sin\beta}{\sin^2 i}\right)\frac{1}{r} & 0 & 0 & \frac{ctn^2 i\, ctn^2\beta - 1}{v\sin L}
\end{bmatrix}
\begin{Bmatrix} dr \\ rd\varphi \\ rd\psi \\ dv \\ vd\gamma \\ vd\beta \end{Bmatrix}
$$

2.4.2.2.3 Elliptic and Hyperbolic Motion (Inertial Coordinates)

2.4.2.2.3.1 Transition Matrices and Coordinate Transformation. The propagation equations developed previously (with the exception of the case of circular motion) were built around the transition matrix relating errors occurring in a rotating local coordinate system. However, there are sound reasons for studying this propagation in other frames of reference (e. g., in an inertial frame as in the case with an IMU, in systems which exhibit the principal values of the error ellipsoids of position and velocity on the coordinate axis, and for the purposes of propagation studies in patched conic trajectories). For this reason, several systems of primary interest will be introduced, and the mechanism through which these effects can be introduced will be presented.

For the purposes of generality, a composite transformation matrix will be established which will transform the errors represented in the local system into those expressed in an arbitrary inertial system. In the process, the transformation necessary to establish the relationship between the rotating system and any other frame of reference will be apparent.



Referring to the sketch, if the following shorthand notation is adopted

$T_X(\alpha)$ = CCW ROTATION ABOUT X

$T_Y(\alpha)$ = CW ROTATION ABOUT Y

$T_Z(\alpha)$ = CCW ROTATION ABOUT Z

the vectors X, Y, Z, $\dot{X}$, $\dot{Y}$, and $\dot{Z}$ can be readily expressed in terms of the known vectors R, S, W, v and $v\Delta\delta$.

$$\begin{Bmatrix} X \\ Y \\ Z \end{Bmatrix} = T_Z(-\Omega)\, T_X(-i)\, T_Z\left[-(\theta+\omega)\right] \begin{Bmatrix} R \\ S \\ W \end{Bmatrix} \equiv T \begin{Bmatrix} R \\ S \\ W \end{Bmatrix}$$

$$\begin{Bmatrix} R \\ S \\ W \end{Bmatrix} = T_Z(-90+\delta) \begin{Bmatrix} v \\ v\Delta\delta \\ w \end{Bmatrix}$$

$$\begin{Bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{Bmatrix} = T_Z(-\Omega)\, T_X(-i)\, T_Z[-90-(\theta+\omega)+\delta] \begin{Bmatrix} v \\ v\Delta\delta \\ w \end{Bmatrix}$$

$$\equiv T^* \begin{Bmatrix} v \\ v\Delta\delta \\ w \end{Bmatrix}$$

Now, combining the transformations and utilizing subscripts I and R to denote inertial and rotating, respectively, the transformation becomes:

$$\begin{Bmatrix} \delta r \\ \delta \dot{r} \end{Bmatrix}_I = \begin{bmatrix} T & 0 \\ 0 & T^* \end{bmatrix} \begin{Bmatrix} \delta r \\ \delta \dot{r} \end{Bmatrix}_R$$

$$\equiv S \begin{Bmatrix} \delta r \\ \delta \dot{r} \end{Bmatrix}_R$$

Similarly, if other frames of reference are desired, the transformation can be readily constructed by performing a series of rotations (in an inverse manner) from the inertial frame to any other frame.

2.4.2.2.3.2 Transition Matrices from Universal Variables (Inertial Coordinates). The equations of conic motion in terms of the so-called universal variables of Dr. S. Herrick will be utilized to develop another means of defining the partial derivatives of the components of position and velocity at any given time (on the conic section) relative to the components of position and velocity at some other arbitrary epoch. This task will be performed to avoid limiting cases and to simplify the development when interest is centered with inertial coordinates. The development will utilize a formulation valid for a non-rotating coordinate system and will be based on the development presented in the discussion of the two-body problem (Ref. 1.3). The required expressions for this analysis are:

$$\vec{r} = f\vec{r}_o + g\vec{s}_o \tag{4.76a}$$

$$\vec{s} = \dot{f}\vec{r}_o + \dot{g}\vec{s}_o \tag{4.76b}$$

where

$$f = 1 - \hat{c}/r_o$$

$$g = \tau - U$$

$$\dot{f} = -\hat{s}/r r_o$$

$$\dot{g} = 1 - \hat{c}/r$$

$\vec{r}$ = inertial position vector = $X\hat{X}^* + Y\hat{Y} + Z\hat{Z}$

$\vec{s}$ = normalized velocity vector = $\vec{v}/\sqrt{\mu} = (\dot{X}\hat{X}^* + \dot{Y}\hat{Y} + \dot{Z}\hat{Z})\sqrt{\mu}$

$\hat{X}^*$ is the X unit vector. (This notation is adopted to avoid confusion with a variable to be defined subsequently)

	ELLIPTIC MOTION	HYPERBOLIC MOTION
$\hat{c}$	$= a(1-\cos X)$	$= a(1-\cosh X)$
$\hat{U}$	$= a^{3/2}(X-\sin X)$	$= (-a)^{3/2}(X-\sinh X)$
$\hat{s}$	$= a^{1/2}(\sin X)$	$= (-a)^{1/2}(\sinh X)$
$\hat{x}$	$= a^{1/2} X$	$= (-a)^{1/2} X$

$$a = -1/\alpha$$

$$D_o = \vec{r}_o \cdot \vec{s}_o$$

$$C_o = 1 + r_o \alpha$$

$$X = E - E_o \quad \text{(ELIPTIC MOTION)} \qquad = F - F_o \quad \text{(HYPERBOLIC MOTION)}$$

The first step in obtaining the desired partial derivatives involves differentiation of the equations for $\vec{r}$ and $\vec{s}$ with respect to the components of $\vec{r}_o$ and $\vec{s}_o$. This task will be drastically simplified if full advantage is taken of the similar form of these derivatives at the outset. Thus, a shorthand notation will be adopted in that u and v ($\dot{u}$ and $\dot{v}$) can assume the values of s, y, and z ($\dot{x}$, $\dot{y}$, and $\dot{z}$).

$$\frac{\partial u}{\partial v_o} = f\delta_{uv} + u_o \frac{\partial f}{\partial v_o} + \dot{u}_o \frac{\partial g}{\partial v_o} \tag{4.77a}$$

$$\frac{\partial u}{\partial \dot{v}} = g\,\delta_{uv} + u_o \frac{\partial f}{\partial \dot{v}_o} + \dot{u} \frac{\partial g}{\partial \dot{v}_o} \tag{4.77b}$$

$$\frac{\partial \dot{u}}{\partial v_o} = \dot{f}\delta_{uv} + u_o \frac{\partial \dot{f}}{\partial v_o} + \dot{u}_o \frac{\partial \dot{g}}{\partial v_o} \tag{4.77c}$$

$$\frac{\partial \dot{u}}{\partial \dot{v}_o} = \dot{g}\,\delta_{uv} + \frac{u_o \partial \dot{f}}{\partial \dot{v}_o} + \dot{u}_o \frac{\partial \dot{g}}{\partial \dot{v}_o} \tag{4.77d}$$

where

$$\delta_{uv} = 0 \qquad u \neq v$$
$$= 1 \qquad u = v$$

Thus, the problem has reduced itself to one of obtaining the derivatives of f, g, $\dot{f}$, and $\dot{g}$ with respect to $\vec{r}$ and $\vec{s}$. This task will in turn be simplified if a set of intermediate parameters is selected, since the x . . . $\dot{z}$ do not appear explicitly in the equations for f . . . $\dot{g}$. The set to be utilized is suggested by the equation for the magnitude of r in this set of variables.

$$r = r_o + D_o \hat{s} + (1 + r_o\alpha)\hat{c}$$
$$= F(r_o, D_o, \alpha)$$

Having selected the intermediate variables, the next task is the differentiation of f, g, $\dot{f}$, and $\dot{g}$.

For f

$$\frac{\partial f}{\partial r_o} = \frac{-r_o \frac{\partial \hat{c}}{\partial r_o} r\hat{c}}{r_o^2} \tag{4.78a}$$

$$\frac{\partial f}{\partial D_o} = -\frac{1}{r_o} \frac{\partial \hat{c}}{\partial D_o} \tag{4.78b}$$

$$\frac{\partial f}{\partial \alpha} = -\frac{1}{r_o} \frac{\partial \hat{c}}{\partial \alpha} \tag{4.78c}$$

For g

$$\frac{\partial g}{\partial r_o} = \frac{\partial \tau}{\partial r_o} - \frac{\partial \hat{U}}{\partial r_o} \tag{4.79a}$$

$$\frac{\partial g}{\partial D_o} = \frac{\partial \tau}{\partial D_o} - \frac{\partial \hat{U}}{\partial D_o} \tag{4.79b}$$

$$\frac{\partial g}{\partial \alpha} = \frac{\partial \tau}{\partial \alpha} - \frac{\partial \tilde{U}}{\partial \alpha} \tag{4.79c}$$

For $\dot{f}$

$$\frac{\partial \dot{f}}{\partial r_o} = \frac{-(r r_o)\frac{\partial \hat{S}}{\partial r_o} + S(r + r_o \frac{\partial r}{\partial r_o})}{(r r_o)^2} \tag{4.80a}$$

$$\frac{\partial \dot{f}}{\partial D_o} = \frac{-(r r_o)\frac{\partial \hat{S}}{\partial D_o} + \hat{S}(r_o \frac{\partial r}{\partial D_o})}{(r r_o)^2} \tag{4.80b}$$

$$\frac{\partial \dot{f}}{\partial \alpha} = \frac{-(r r_o)\frac{\partial \hat{S}}{\partial \alpha} + \hat{S}(r_o \frac{\partial r}{\partial \alpha})}{(r r_o)^2} \tag{4.80c}$$

and for $\dot{g}$

$$\frac{\partial \dot{g}}{\partial r_o} = \frac{-r \frac{\partial \hat{C}}{\partial r_o} + \hat{C} \frac{\partial r}{\partial r_o}}{r^2} \tag{4.81a}$$

$$\frac{\partial \dot{g}}{\partial D_o} = \frac{-r \frac{\partial \hat{C}}{\partial D_o} + \hat{C} \frac{\partial r}{\partial D_o}}{r^2} \tag{4.81b}$$

$$\frac{\partial \dot{g}}{\partial \alpha} = \frac{-r \frac{\partial \hat{C}}{\partial \alpha} + \hat{C} \frac{\partial r}{\partial \alpha}}{r^2} \tag{4.81c}$$

Now, attention turns to the derivatives of $\hat{C}$, $\hat{S}$, $\hat{U}$, τ, etc., with respect to r_o, D_o, α.

For $\hat{C}$

$$\frac{\partial \hat{c}}{\partial r_o} = a \sin\chi \frac{\partial \chi}{\partial r_o} = \hat{S} \frac{\partial \hat{\chi}}{\partial r_o} \tag{4.82a}$$

$$\frac{\partial \hat{c}}{\partial D_o} = \hat{S} \frac{\partial \hat{\chi}}{\partial D_o} \tag{4.82b}$$

$$\frac{\partial \hat{c}}{\partial \alpha} = \hat{S} \frac{\partial \hat{\chi}}{\partial \alpha} + (1 - \cos\chi)\frac{\partial a}{\partial \alpha} = \hat{S} \frac{\partial \hat{\chi}}{\partial \alpha} + \hat{C} a \tag{4.82c}$$

For $\hat{U}$

$$\frac{\partial \hat{U}}{\partial r_o} = a^{3/2}(1 - \cos\chi)\frac{\partial \chi}{\partial r_o} = \hat{C} \frac{\partial \hat{\chi}}{\partial r_o} \tag{4.83a}$$

$$\frac{\partial \hat{U}}{\partial D_o} = \hat{C} \frac{\partial \hat{\chi}}{\partial D_o} \tag{4.83b}$$

$$\frac{\partial \hat{U}}{\partial \alpha} = \hat{C} \frac{\partial \hat{\chi}}{\partial \alpha} + \frac{3}{2}(\chi - \sin\chi) a^{1/2} \frac{\partial a}{\partial \alpha} = \hat{C} \frac{\partial \hat{\chi}}{\partial \alpha} + \frac{3}{2} \hat{U} a \tag{4.83c}$$

For $\hat{S}$

$$\frac{\partial \hat{S}}{\partial r_0} = \cos\chi \frac{\partial \hat{\chi}}{\partial r_0} \tag{4.84a}$$

$$\frac{\partial \hat{S}}{\partial D_0} = \cos\chi \frac{\partial \hat{\chi}}{\partial D_0} \tag{4.84b}$$

$$\frac{\partial \hat{S}}{\partial \alpha} = \cos\chi \frac{\partial \hat{\chi}}{\partial \alpha} + \frac{1}{2}\sin\chi\, a^{3/2} \tag{4.84c}$$

For τ

$$\frac{\partial \tau}{\partial r_0} = 0 \tag{4.85a}$$

$$\frac{\partial \tau}{\partial D_0} = 0 \tag{4.85b}$$

$$\frac{\partial \tau}{\partial \alpha} = \frac{1}{2} a^{3/2} t = \frac{1}{2} a \tau \tag{4.85c}$$

and for r

$$\frac{\partial r}{\partial r_0} = 1 + D_0 \frac{\partial \hat{S}}{\partial r_0} + C_0 \frac{\partial \hat{C}}{\partial r_0} + \alpha \hat{C} \tag{4.86a}$$

$$\frac{\partial r}{\partial D_0} = \hat{S} + D_0 \frac{\partial \hat{S}}{\partial D_0} + C_0 \frac{\partial \hat{C}}{\partial D_0} \tag{4.86b}$$

$$\frac{\partial r}{\partial \alpha} = D_0 \frac{\partial \hat{S}}{\partial \alpha} + C_0 \frac{\partial \hat{C}}{\partial \alpha} + r_0 \hat{C} \equiv r_\alpha \tag{4.86c}$$

The final set of partials required is now recognized to be that of $\hat{X}$ with respect to r_0, D_0 and α. This set requires the equation for time (analogous to Kepler's equation) be differentiated as follows:

$$\sqrt{a}\, t \equiv \tau = r_0 \hat{\chi} + D_0 \hat{C} + (1 + \alpha r_0)\hat{U}$$

for $\frac{\partial \hat{\chi}}{\partial r_0}$

$$0 = \hat{\chi} + r_0 \frac{\partial \hat{\chi}}{\partial r_0} + D_0 \frac{\partial \hat{C}}{\partial r_0} + \alpha \hat{U} + (1 + \alpha r_0) \frac{\partial \hat{U}}{\partial r_0}$$

$$= \hat{\chi} + \alpha \hat{U} + r \frac{\partial \hat{\chi}}{\partial r_0} \tag{4.87a}$$

$$\frac{\partial \hat{\chi}}{\partial r_0} = -\frac{\hat{S}}{r}$$

for $\frac{\partial \hat{\chi}}{\partial D_o}$ $\quad 0 = r_o \frac{\partial \hat{\chi}}{\partial D_o} + \hat{C} + D_o \frac{\partial \hat{c}}{\partial D_o} + C_o \frac{\partial \hat{U}}{\partial D_o}$

$$\frac{\partial \hat{\chi}}{\partial D_o} = -\frac{\hat{C}}{r} \tag{4.87b}$$

for $\frac{\partial \hat{\chi}}{\partial \alpha}$

$$\begin{aligned} \tfrac{1}{2} a\tau &= r_o \frac{\partial \hat{\chi}}{\partial \alpha} + D_o \frac{\partial \hat{c}}{\partial \alpha} + C_o \frac{\partial \hat{U}}{\partial \alpha} + \hat{U} r_o \\ &= \hat{U}\left(r_o + \tfrac{3}{2} C_o a\right) + D_o(\hat{C}_o) + \frac{\partial \hat{\chi}}{\partial \alpha}(r) \end{aligned} \tag{4.87c}$$

$$\frac{\partial \hat{\chi}}{\partial \alpha} = -\frac{1}{r}\left[\hat{U}(C_o a + r_o) + \frac{a}{2}(D_o \hat{C} - r_o \hat{\chi})\right] \equiv -\frac{\chi_\alpha}{r}$$

Now, substituting back into the previous expressions and collecting terms, the derivatives required to compute the partials of f, g, $\dot{f}$ and $\dot{g}$ with respect to the intermediate parameters are:

For $\hat{C}$

$$\frac{\partial \hat{C}}{\partial r_o} = -\frac{\hat{S}^2}{r} \tag{4.88a}$$

$$\frac{\partial \hat{C}}{\partial D_o} = -\frac{\hat{S}\hat{C}}{r} \tag{4.88b}$$

$$\frac{\partial \hat{C}}{\partial \alpha} = -\frac{\hat{S}}{r}\chi_\alpha + \hat{C}a \equiv C_\alpha \tag{4.88c}$$

For $\hat{U}$

$$\frac{\partial \hat{U}}{\partial r_o} = -\frac{\hat{C}\hat{S}}{r} \tag{4.89a}$$

$$\frac{\partial \hat{U}}{\partial D_o} = -\frac{\hat{C}^2}{r} \tag{4.89b}$$

$$\frac{\partial \hat{U}}{\partial \alpha} = -\frac{\hat{C}\chi}{r} + \frac{3}{2}\hat{U}a \equiv U_\alpha \tag{4.89c}$$

For $\hat{S}$

$$\frac{\partial \hat{S}}{\partial r_o} = -\cos\chi\,\frac{\hat{S}}{r} \tag{4.90a}$$

$$\frac{\partial \hat{S}}{\partial D_o} = -\cos\chi\,\frac{\hat{C}}{r} \tag{4.90b}$$

$$\frac{\partial \hat{S}}{\partial \alpha} = \cos\chi\,\chi_\alpha + \tfrac{1}{2}\sin\chi\, a^{3/2} \equiv S_\alpha \tag{4.90c}$$

For r

$$\frac{\partial r}{\partial r_o} = 1 + D_o\left(-\cos\chi\,\frac{\hat{S}}{r}\right) + C_o\left(-\frac{\hat{S}^2}{r}\right) + \alpha\hat{C} = \frac{r_o}{r} f \tag{4.91a}$$

$$\frac{\partial r}{\partial D_o} = \hat{S} + D_o\left(-\cos\chi\,\frac{\hat{C}}{r}\right) + C_o\left(-\frac{\hat{S}\hat{C}}{r}\right) = \frac{1}{r} g \tag{4.91b}$$

$$\frac{\partial r}{\partial D_o} = D_o S_\alpha + C_o C_\alpha + r_o \hat{C} \equiv r_\alpha \tag{4.91c}$$

Finally, the required partials of f, g, $\dot{f}$, and $\dot{g}$ with respect to r_o, D_o, and α are:

For f

$$\frac{\partial f}{\partial r_o} = \frac{\hat{C}}{r_o^2} + \frac{\hat{S}^2}{r r_o} \tag{4.92a}$$

$$\frac{\partial f}{\partial D_o} = \frac{3\hat{C}}{r r_o} \tag{4.92b}$$

$$\frac{\partial f}{\partial \alpha} = -\frac{1}{r_o} C_\alpha \tag{4.92c}$$

For g

$$\frac{\partial g}{\partial r_o} = \frac{\hat{S}\hat{C}}{r} \tag{4.93a}$$

$$\frac{\partial g}{\partial D_o} = \frac{\hat{C}^2}{r} \tag{4.93b}$$

$$\frac{\partial g}{\partial \alpha} = \frac{1}{2} a\left[\tau - 3\hat{U}\right] + \frac{\hat{C}}{r}\chi_\alpha \tag{4.93c}$$

For f

$$\frac{\partial \dot{f}}{\partial r_o} = \frac{\hat{S}}{r^3 r_o^2}\left[r^2 + r_o(r\cos\chi + r_o f)\right] \tag{4.94a}$$

$$\frac{\partial \dot{f}}{\partial D_o} = \frac{1}{r^3 r_o}\left[\hat{S} g + \cos\chi\, r\hat{C}\right] \tag{4.94b}$$

$$\frac{\partial \dot{f}}{\partial \alpha} = \frac{1}{r r_o} S_\alpha + \frac{\hat{S}}{r^2 r_o} r_\alpha \tag{4.94c}$$

For g

$$\frac{\partial \dot{g}}{\partial r_o} = \frac{1}{r^3}\left[r\hat{S}^2 + \hat{C} r_o f\right] \tag{4.95a}$$

$$\frac{\partial \dot{g}}{\partial D_o} = \frac{\hat{C}}{r^3}\left[r\hat{S} + g\right] \tag{4.95b}$$

$$\frac{\partial \dot{g}}{\partial \alpha} = -\frac{r C_\alpha + \hat{C} r_\alpha}{r^2} \tag{4.95c}$$

Thus, the only remaining steps at this point are to provide the derivatives of the intermediate set of parameters with respect to the components of $\vec{r}_o$ and $\vec{s}_o$, to construct the derivatives of f, g, $\dot{f}$, and $\dot{g}$ with respect to $\vec{r}_o$ and $\vec{s}_o$, and to relate the results to the parameters $\vec{r}$ and $\vec{v}$. The first step is accomplished by referring to the definitions of r_o, D_o, and α.

$$r_o^2 = \sum_{i=1}^{3} x_i^2 \qquad s_o^2 = \sum_{i=1}^{3} s_i^2$$

$$\frac{\partial r_o}{\partial x_i} = \frac{x_i}{r_o} \qquad \frac{\partial r_o}{\partial s_i} = 0$$

$$D_o = \vec{r}_o \cdot \vec{s}_o$$

$$\frac{\partial D_o}{\partial x_i} = s_i \qquad \frac{\partial D_o}{\partial s_i} = x_i$$

$$\alpha = \vec{s} \cdot \vec{s} - \frac{2}{r_o}$$

$$\frac{\partial \alpha}{\partial x_i} = \frac{2x_i}{r_o^3} \qquad \frac{\partial \alpha}{\partial s_i} = s_i$$

The second step is accomplished through the medium of the chain rule, i.e.,

$$\frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial r_o}\frac{\partial r_o}{\partial x_i} + \frac{\partial f}{\partial D_o}\frac{\partial D_o}{\partial x_i} + \frac{\partial f}{\partial \alpha}\frac{\partial \alpha}{\partial x_i}$$

etc., and the third and final step is employed to remove the normalization factor applied to the components of the velocity. Since

$$\vec{s} = \frac{\vec{v}}{\sqrt{\mu}}$$

$$d\vec{s} = \frac{1}{\sqrt{\mu}}\, d\vec{v}$$

the desired matrix is

$$\begin{Bmatrix} d\vec{r} \\ d\vec{v} \end{Bmatrix} = \left[\begin{array}{c|c} \dfrac{\partial \vec{r}}{\partial \vec{r}_o} & \dfrac{1}{\sqrt{\mu}}\dfrac{\partial \vec{r}}{\partial \vec{s}_o} \\ \hline \sqrt{\mu}\dfrac{\partial \vec{s}}{\partial \vec{r}_o} & \dfrac{\partial \vec{s}}{\partial \vec{s}_o} \end{array}\right] \begin{Bmatrix} d\vec{r}_o \\ d\vec{v}_o \end{Bmatrix} \qquad (4.96)$$

2.4.2.3 Approximate Method of Including the Effects of Trajectory Perturbations

If the trajectory itself is being generated with precision (i.e., all perturbative forces are being integrated in such a manner that their net effect is accurately known), two simple means exist for compensating for the deletion of the effects of the variations in the perturbing accelerations in the computation of the transition matrix along the true trajectory. The first and most accurate of the two schemes is obtained by considering the true trajectory to be composed of a series of conic arcs, one terminal (i.e., $\vec{r}$, $\vec{v}$) of which matches the true trajectory at the epoch corresponding to the terminal selected. Under this assumption, the total transition matrix relating the present epoch and some arbitrary initial epoch is obtained by forming the product of the most recent transition matrix [relating t_n and t_{n-1}; $\varphi(t_o, t_o) = I$], and that formed using the present values of $\vec{r}_o$ and $\vec{v}$ and a propagation time equal to the step size of the numerical integration. That is,

$$\phi(t_n, t_o) = \phi(t_n, t_{n-1})\,\phi(t_{n-1}, t_o)$$

This method does not involve integration but is, nonetheless, subject to error accumulation due to the fact that loss of significance can result in the series of required matrix multiplications. Thus, improved accuracy may result if larger steps are taken for the purpose of incrementing $\varphi(t, t_o)$ [i.e., increment $\varphi(t, t_o)$ each nth integration step]. It is important to note, however, that the limiting assumption inherent in the equations of motion will never admit great precision in this approach. Thus, extremely involved logic to determine the proper propagation time appears unwarranted.

A second, less accurate, approximation is obtained by considering only the two terminals $\vec{r}_o$, $\vec{v}_o$, t_o and $\vec{r}$, $\vec{v}$, T. For this case, two estimates of the transition matrix can be generated. The first is obtained simply by evaluating one of the conic matrices for the condition $\vec{r}_o$, $\vec{v}_o$, $T-t_o$, i.e.,

$$\vec{\delta}_1(T) = \phi_1(T, t_o)\,\vec{\delta}_1(t_o)$$

The second estimate is obtained by considering the trajectory from $\vec{r}_T$, $\vec{v}_T$, T in negative time [i.e., $-(T - t_o)$]

$$\vec{\delta}_2(t_o) = \phi_2(t_o, T)\,\vec{\sigma}_2(T)$$

Thus,

$$\vec{\delta}_2(T) = \phi_2^{-1}(t_o, T)\,\vec{\delta}_2(t_o)$$

These two estimates can now be combined by weighing them in some fashion. For the case of equal weights

$$\phi(T, t_o) \approx .5 \left[\phi_1(T, t_o) + \phi_2^{-1}(t_o, T)\right]$$

This second procedure is easily mechanized if the transition matrices relate inertial errors, since the analytic inverse theorem can be applied to avoid numerical problems in the process. In fact, this theorem can be utilized even in those cases where rotating coordinates are employed when it is noted that the rotating and inertial matrices are related as follows:

$$\phi_R(t, t_o) = S(t)\, \phi_I(t, t_o)$$

where S (t) is an orthogonal transformation. Thus,

$$\phi_R^{-1}(t, t_o) = \phi_I^{-1}(t, t_o)\, S^T(t)$$

The accuracy of the second technique is somewhat unknown. From physical reasoning of the case in which T-t is less than one period, it can be argued that the inclusion of the second estimate will provide an improvement in the accuracy. However, a measure of the improvement is difficult to construct since the magnitude of the perturbative displacement (secular plus periodic; resulting from the non-central forces of the problem) from the conic trajectory varies from point to point along the trajectory. For the case where $T-t_o$ is greater than one period, the same effects exist. In addition, errors are introduced due to the fact that pure conic reference is being assumed for increasingly long intervals of time.

Another family of approximations can, of course, be generated by combining the first two types.

2.5 DATA WEIGHTING

2.5.1 General Theory

The implementation of the various estimators that are introduced in this monograph requires a knowledge of the covariance matrix of the observables. This covariance matrix can be determined from the covariance matrices of instrument errors and navigation model errors. This section will present the general expression that relates these covariance matrices. The following section (2.5.2) will give the detailed expressions for each type of navigation measurement that will explicitly define this covariance relationship for any combination of navigation measurement.

When a navigation observation is made, it will differ from a predetermined nominal measurement because of several reasons. First, the vehicle will be off the nominal trajectory by some amount in both position and velocity. Second, the constants of the navigation model that were used to calculate the nominal measurement are themselves in error, because of uncertainties in the physical dimensions and uncertainties in celestial body positions. Finally, the uncertainty of the instrument that is used for the measurement will constribute to the difference in the measurement.

Since all of the sources of the deviations are small and the nominal measurement is assumed to be known, the measurement may be expanded in a Taylor Series about the nominal value by including derivatives with respect to all of the variables that influence the measurement. Furthermore, since the contribution of each deviation source is very small, the higher order terms may be neglected in the Taylor Series. Letting q_i be the measurement of interest, the Taylor Series can now be expressed as

$$q_i = q_{i_0} + \frac{\partial q_i}{\partial \underline{X}} d\underline{X} + \frac{\partial q_i}{\partial \underline{M}} d\underline{M} + \frac{\partial q_i}{\partial \underline{\epsilon}_i} d\underline{\epsilon}_i \qquad (5.1)$$

where q_{j_0} is the nominal value of the measurement

$d\underline{X}$ is the true state vector deviation
$d\underline{M}$ is a vector composed of the model uncertainties
$d\underline{\epsilon}_i$ is the instrument measurement error associated with q_i
$\underline{X}$ is the state vector
$\underline{M}$ is the vector composed of the navigation model parameters
$\underline{\epsilon}_i$ is the vector of measurement parameters

Equation (5.1) can be recognized as a more general form of the position deviation equation that was presented in a previous monograph, Reference 1.1. It is recalled that a deviation in some observation from a nominal value could be related to the position deviation as follows:

$$\delta \underline{q} = H \delta R \qquad (5.2)$$

where $\underline{\delta q}$ is an observation vector composed of the δq_i's
$\delta q_i = q_i - q_{i_o}$
$\underline{\delta R}$ = position deviation from nominal

A comparison of Equations (5.1) and (5.2) reveals that Equation (5.2) corresponds to the second term of Equation (5.1), where the state vector deviation is considered to be position deviation only. The significance of all of terms in Equation (5.1) should now be apparent. The first term is merely a predetermined quantity that gives the value of the quantity under perfect conditions, i.e., that value that would be measured if the vehicle were exactly on the nominal trajectory, all positions and dimensions of objects used for the observation were known exactly, and the instrument used for the measurement were error free. The second term corresponds to the deviations in the measurement that are due to perturbations of the state vector about the nominal trajectory. The third term represents all measurement deviations that are introduced by uncertainties in the navigation model. The number of elements in this vector depends on the number of uncertainties in the model that have a direct first order effect on the measurement. Finally, the last term in Equation (5.1) represents the error due to the uncertainties in the instrument that is used for the ith measurement. In cases where the navigation observable and the measured quantity are the same, the vector $\frac{\partial \underline{q}}{\partial \underline{\epsilon}}$ will merely contain a "one" in the appropriate location. However, if the measured quantity and the navigation observable differ, as in radar measurements where phase and frequency measurements are indirect range and range-rate measurements, then the vector $\frac{\partial \underline{q}}{\partial \underline{\epsilon}}$ must contain the appropriate partial derivative to account for the uncertainty in the sensor.

If a series of navigation measurements is made. Equation (5.1) can be extended to a vector equation.

$$\underline{q} = \underline{q}_o + \frac{\partial \underline{q}}{\partial \underline{X}} \underline{dX} + \frac{\partial \underline{q}}{\partial \underline{M}} \underline{dM} + \frac{\partial \underline{q}}{\partial \underline{\epsilon}} \underline{d\epsilon} \qquad (5.3)$$

It should be noted that the model uncertainty term is now shown as the product of a matrix, $\frac{\partial \underline{q}}{\partial \underline{M}}$, and a vector $\underline{dM}$. This notation can be adopted because model uncertainties, even though they all may not be used for a particular observation, may be entered into a column vector. If a particular model uncertainty is not used for an observation, then the $\frac{\partial \underline{q}}{\partial \underline{M}}$ matrix will contain zeroes in the appropriate locations.

The covariance matrix for $\underline{q}$ can be found by employing the expected value theorem. (See Appendix A.) The result is

$$cov(\underline{q}) = E\left\{[\underline{q} - E(\underline{q})][\underline{q} - E(\underline{q})]^T\right\} \qquad (5.4)$$

Equation (5.4) can be evaluated once the expected value of q is determined. This step can be accomplished by taking the expected value of Equation (5.3), i.e.,

$$E(\underline{q}) = E(\underline{q}_0) + E\left[\frac{\partial \underline{q}}{\partial \underline{X}} d\underline{X}\right] + E\left[\frac{\partial \underline{q}}{\partial \underline{M}} d\underline{M}\right] + E\left[\frac{\partial \underline{q}}{\partial \underline{\epsilon}} d\underline{\epsilon}\right] \qquad (5.5)$$

Since q_0 is a calculated nominal vector, its expected value is q_0. Further, the expected value of $\underline{dX}$ is $\underline{dX}$ because $\underline{dX}$ is the true deviation. The last two terms in Equation (5.5) are zero, since $\underline{dM}$ and $\underline{d\epsilon}$ are random variables with zero mean (by choice of the nominal values of the parameters). It should be noted that the result obtained from Equation (5.5) after the expected value is taken, is consistent with Equation (5.2) which assumes the true values of the measurements and position deviation are known, i.e.,

$$E(\underline{q}) = \bar{\underline{q}} = \underline{q}_0 + \frac{\partial \underline{q}}{\partial \underline{X}} d\underline{X} \qquad (5.6)$$

or

$$\delta \underline{q} = \bar{\underline{q}} - \underline{q}_0 = \frac{\partial \underline{q}}{\partial \underline{X}} d\underline{X}$$

Equation (5.4) can now be evaluated if the quantity q - E (q) is expressed as

$$\begin{aligned} \underline{q} - E(\underline{q}) &= \underline{q}_0 + \frac{\partial \underline{q}}{\partial \underline{X}} d\underline{X} + \frac{\partial \underline{q}}{\partial \underline{M}} d\underline{M} + \frac{\partial \underline{q}}{\partial \underline{\epsilon}} d\underline{\epsilon} - \underline{q}_0 - \frac{\partial \underline{q}}{\partial \underline{X}} d\underline{X} \\ &= \frac{\partial \underline{q}}{\partial \underline{M}} d\underline{M} + \frac{\partial \underline{q}}{\partial \underline{\epsilon}} d\underline{\epsilon} \end{aligned} \qquad (5.7)$$

The result is

$$cov(\underline{q}) = E\left[B\, d\underline{M}\, d\underline{M}^T B^T + B\, d\underline{M}\, d\underline{\epsilon}^T C^T + C\, d\underline{\epsilon}\, d\underline{M}^T B^T + C\, d\underline{\epsilon}\, d\underline{\epsilon}^T C^T\right] \qquad (5.8)$$

where $B \equiv \frac{\partial \underline{q}}{\partial \underline{M}}$

$C \equiv \frac{\partial \underline{q}}{\partial \underline{\epsilon}}$

Before proceeding, it should be noted that cross correlations between the vectors $\underline{dM}$ and $\underline{d\epsilon}$ usually do not exist, since components of these two vectors are not affected by one another. This statement becomes evident from an examination of the following chart that shows the nature of any correlation that would exist if the vectors were not uncorrelated.

Influencing Vector \ Affected Vector	$\underline{dM}$	$\underline{d\epsilon}$
$\underline{dM}$	X	Influences of uncertainties in the navigation model on the accuracy of the instruments used in the measurement.
$\underline{d\epsilon}$	Influence of instrument uncertainties on the uncertainties of the navigation model.	X

Since there is no cross correlation between $\underline{dM}$ and $\underline{d\epsilon}$ the expected value of any cross product term in Equation (5.8) will be the null matrix. Hence, the covariance of q can be written as

$$\begin{aligned} cov(q) &= E[B\,\underline{dM}\,\underline{dM}^T B^T] + E[B\,\underline{dM}\,\underline{d\epsilon}^T C^T] \\ &\quad + E[C\,\underline{d\epsilon}\,\underline{dM}^T B^T] + E[C\,\underline{d\epsilon}\,\underline{d\epsilon}^T C^T] \\ &= E[B\,\underline{dM}\,\underline{dM}^T B^T] + E[C\,\underline{d\epsilon}\,\underline{d\epsilon}^T C^T] \end{aligned} \tag{5.9}$$

Now, the covariance matrix of the observables can be related to the covariance matrices of the navigation model uncertainties and sensor uncertainties as follows:

$$cov(q) = B[cov(\underline{dM})]B^T + C[cov(\underline{d\epsilon})]C^T \tag{5.10}$$

The matrices B and C have been defined to be matrices of partial derivatives. In the following section, 2.5.2, a detailed analysis that will define

the elements of the B and C matrices will be presented for each type of navigation measurement and sensor developed in Reference 1.1. The covariance matrices for the navigation model and sensors can be constructed from the various uncertainties that are encountered in the particular measurements and sensors being used in the observations. Following the derivation of the elements of the B and C matrices for a particular measurement, a sample problem will be presented in Section 2.5.4 to demonstrate how the complete B and C matrices are constructed from several measurements and how the covariance matrices of the model and sensor uncertainties are determined.

2.5.2 Navigation Measurement Uncertainties

This section will develop the detailed expressions that represent the B anc C matrices introduced earlier. The expressions for the navigation model uncertainty will be presented in Section 2.5.2.1 and those for the sensor uncertainties in Section 2.5.2.2. It should be noted that each measurement (i.e., observable) defines but a row of these matrices; thus simultaneous measurement of several quantities requires that successive rows be added to both matrices.

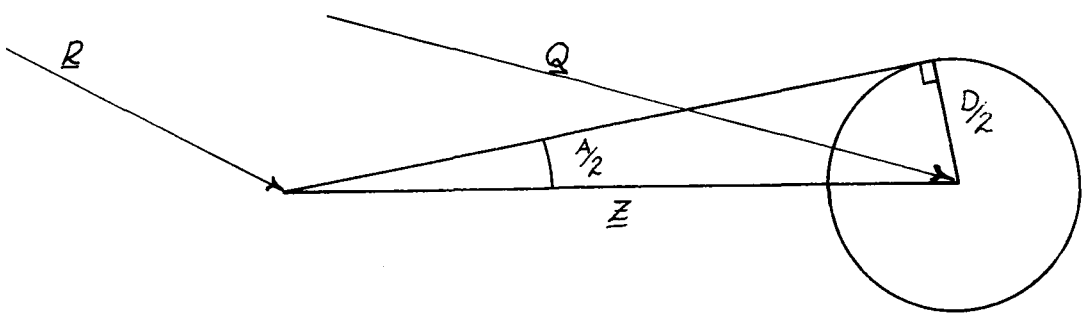
2.5.2.1 Navigation Model Uncertainties

The following section will present the expressions for the navigation observable uncertainties as related to the navigation model uncertainties. The terms developed herein will be elements of the B matrix, i.e., they are the partial derivatives of the navigation observables with respect to the model uncertainties. The terminology was introduced in Reference 1.1.

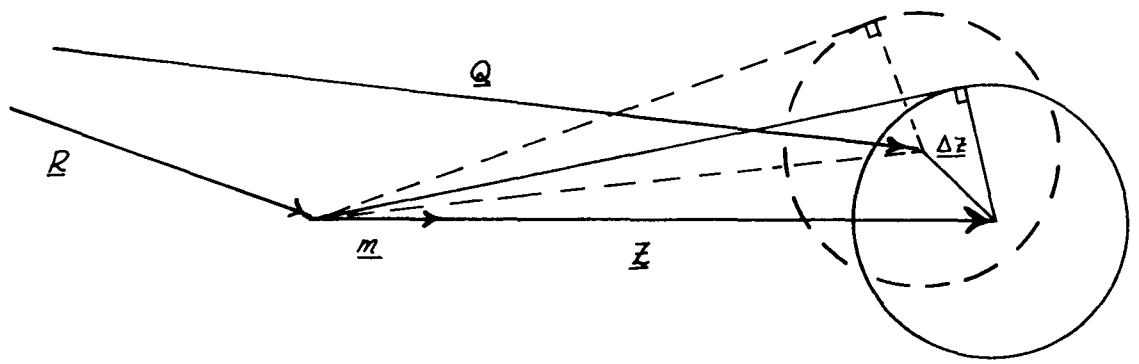
2.5.2.1.1 Planet Diameter Measurement

The planet diameter measurement as discussed in a previous monograph (Reference 1.1), assumed that the physical dimensions of the model were known exactly, so the deviations from the nominal trajectory could be found by taking first order variations of the angle measurement with respect to position deviation in the radial direction. If consideration is to be given to uncertainties in the model, the partial derivatives of the angle measured with respect to planet position and planet diameter must be found and be incorporated into the B matrix.

Consider the following planet diameter measurement:



If the planet were in a slightly different position than anticipated and all else were exactly correct, then a first order change in the angle measurement due to this plane position uncertainty could be shown as follows:



Since the original relationship that describes the exact measurement is

$$\sin\frac{A}{2} = \frac{D/2}{Z} \quad , \tag{5.11}$$

the variation of A with respect to Z can be determined by straightforward differentiation. The same technique can be used to find the first order variation of A with respect to the uncertainty in the planet diameter. The differentiation is shown below:

$$\frac{1}{2}\cos\left(\frac{A}{2}\right)dA = \frac{D}{2}(-1)\,Z^{-2}dZ + \frac{1}{2Z}dD \tag{5.12}$$

Hence, the total model uncertainty in the measured angle is

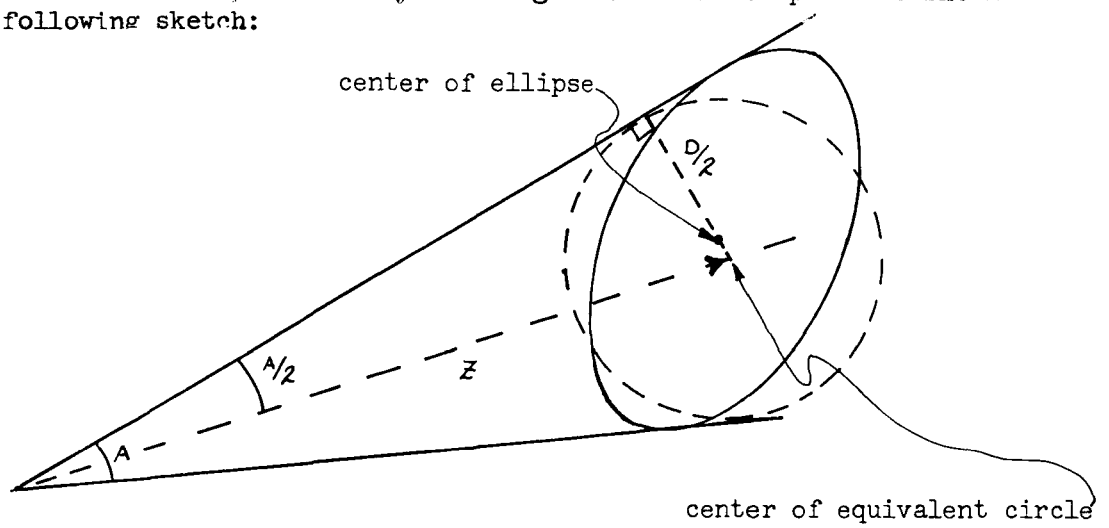
$$dA = \frac{-D\,dZ}{Z^2\cos\left(\frac{A}{2}\right)} + \frac{dD}{Z\cos\left(\frac{A}{2}\right)} \tag{5.13}$$

The quantity dZ is the projection of the $\underline{\Delta Z}$ vector in the radial direction. So dZ can be written as $\underline{m}\cdot\underline{\Delta Z}$, and now

$$dA = \frac{-D(\underline{m}\cdot\Delta\underline{z})}{z^2\cos\left(\frac{A}{2}\right)} + \frac{dD}{z\cos\left(\frac{A}{2}\right)} \qquad (5.14)$$

Equation (5.14) is the final expression for the measurement uncertainty due to the navigation model uncertainty in the planet diameter measurement.

The previous analysis also assumed a perfectly spherical planet. However, significant difficulty is encountered if an analysis that considers planet flattening is attempted. The difficulty stems from the need to express in an analytical manner the general angle subtended by an ellipsoid from a point in space. An approximate analysis of the flattened planet measurement can be performed by defining an equivalent sphere as shown in the following sketch:



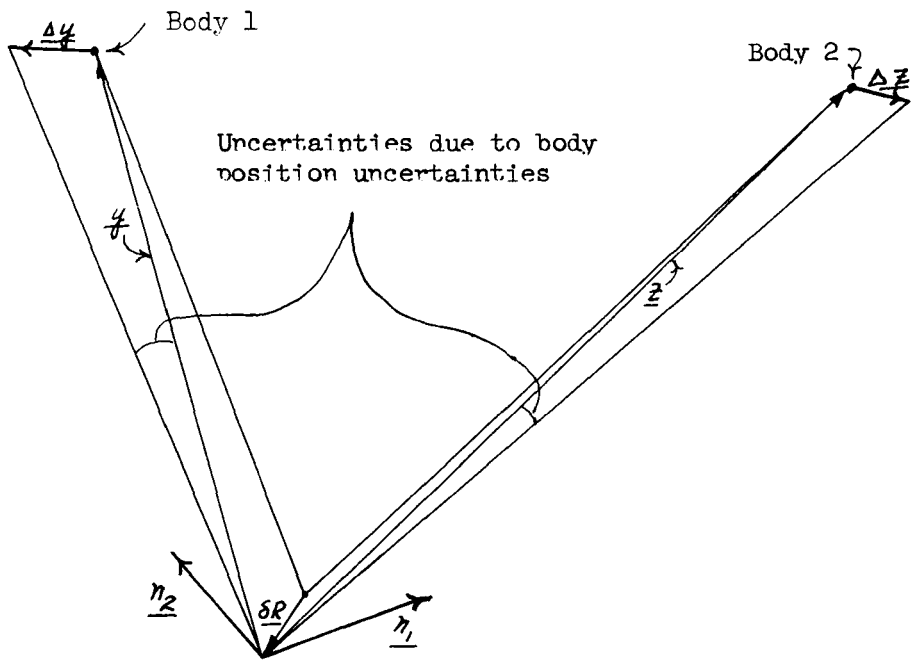
In this sketch, a circle is defined in the subtended angle that is formed by the flattened planet. In order for this circle to be as close to the flattened planet in its diameter measurement characteristics, it should have its center as close as possible to the center of the planet. In general, the two center points will not coincide. Therefore, the circle centermust be defined to be at a point such that a radius that is perpendicular to a line of sight passes through the center point of the planet. An uncertainty in the relative attitude between the observation point and the planet would introduce an uncertainty in the diameter of the equivalent circle to be used. Furthermore, the flattening will cause a different measurement for every different plane that is defined by the lines of sight of the measurement of the ellipsoid. Such an approximate analysis would be acceptable if the equivalent planet diameter uncertainties were known for some predetermined nominal model. Usually, however, a continuous mathematical

relationship between the observable and the uncertainty is desired so that a general case can be treated. The derivation of this relationship is quite lengthy and has been placed in Appendix D for completeness.

2.5.2.1.2 The Angle Between the "Close" Bodies

The model and sensor uncertainties that influence the measurement of the angle between two close bodies will now be considered. (There are uncertainties due to the fact that the positions of the bodies are not known exactly, and due to the fact that the sensor has a limited accuracy). First, the uncertainties in the angle due to the uncertainties in the body positions will be considered. If each body position deviation vector is resolved into components in the $\underline{n}_1$ and $\underline{n}_2$ directions, and each is related to a change in the angle measurement, it is evident that the components that are perpendicular to the lines-of-sight have a contribution to the angle measurement. (Reference 1.1 defined $\underline{n}_1$ ($\underline{n}_2$) to be a unit vector perpendicular to $\underline{y}$ ($\underline{z}$) and in the plane determined by the angle measurement.) Those that are parallel to the line-of-sight have no contribution. From the following sketch, the body position uncertainties, Δy and Δz, can be seen to have an influence dA on A, where

$$dA = \frac{\underline{n}_2 \cdot \Delta \underline{z}}{z} + \frac{\underline{n}_1 \cdot \Delta \underline{y}}{y} \qquad (5.15)$$



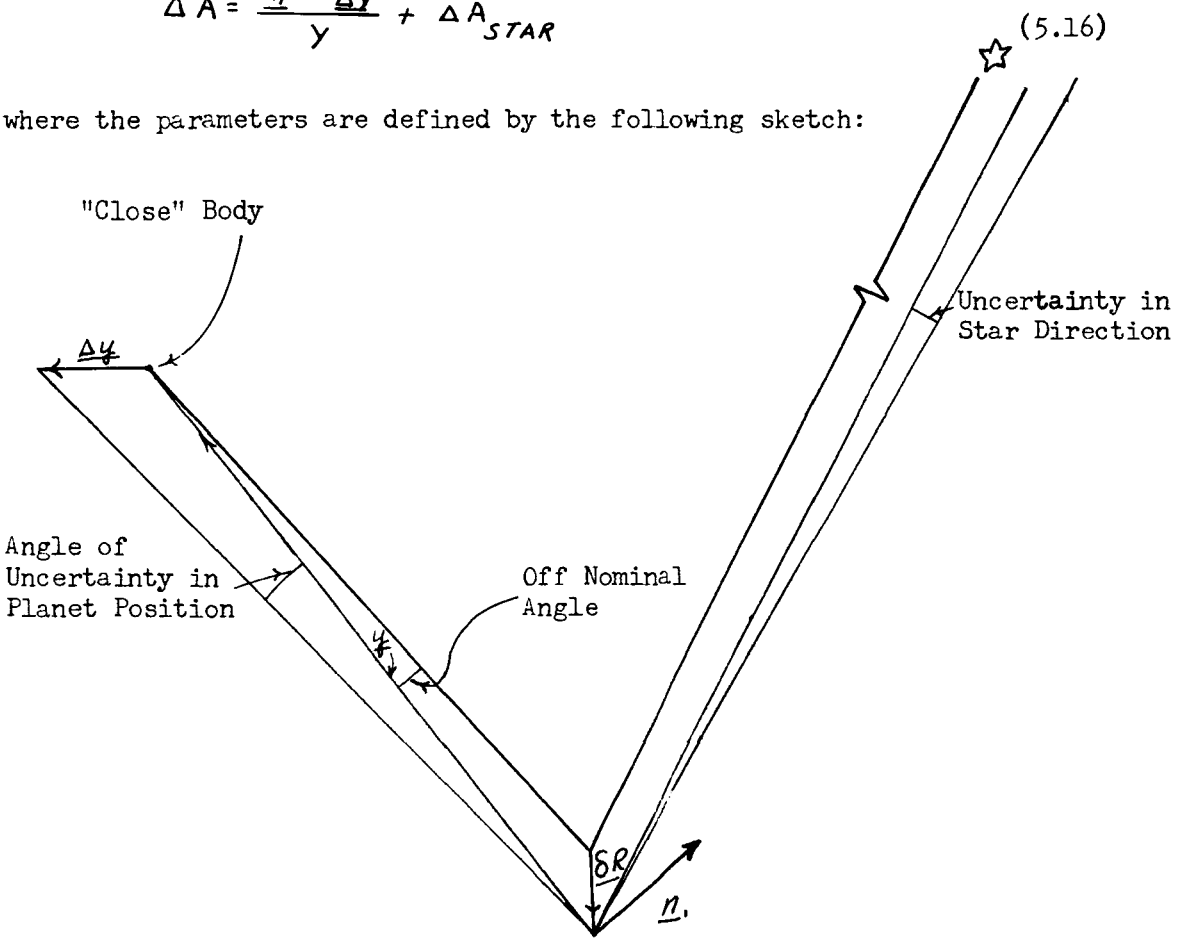
Since the sensor device measures the angle directly, the uncertainty in the angle due to the sensor uncertainty is just the amount in the sensor.

2.5.2.1.3 The Angle Between a "Close" Body and a Star

The inaccuracies associated with the measurement of the angle between a close celestial body and a star are the uncertainty in the position of the celestial body, the uncertainty in the direction of the star, and the uncertainty in the instrument making the measurement. However, only the component of the body position uncertainty vector that is perpendicular to line-of-sight to the body has a direct influence of the angle measurement. Therefore, the total model uncertainty in this measurement is

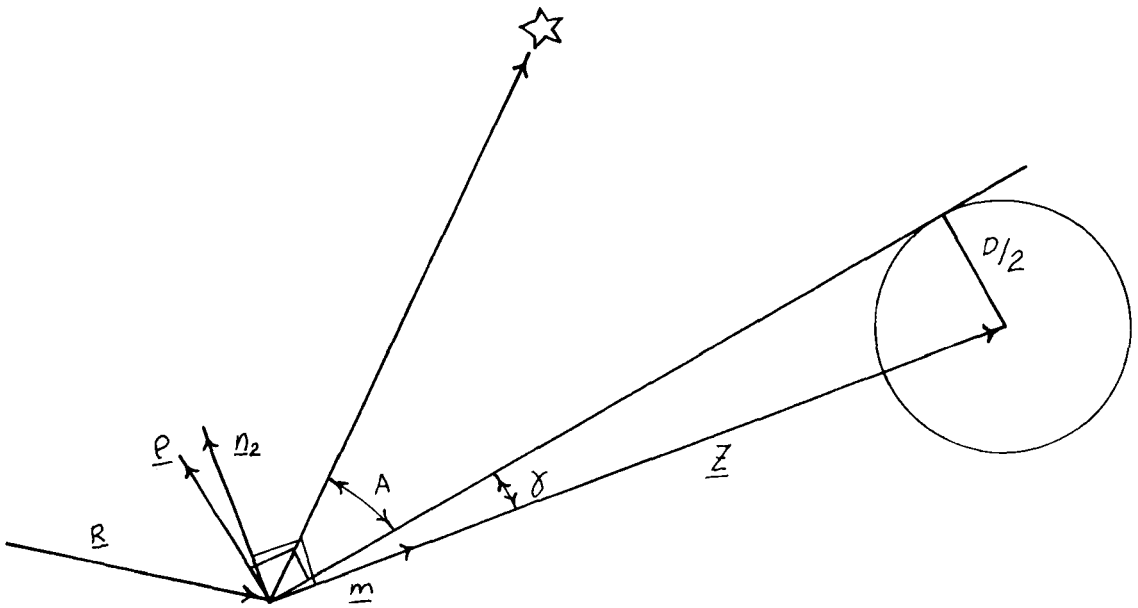
$$\Delta A = \frac{n_1 \cdot \Delta Y}{y} + \Delta A_{STAR} \qquad (5.16)$$

where the parameters are defined by the following sketch:

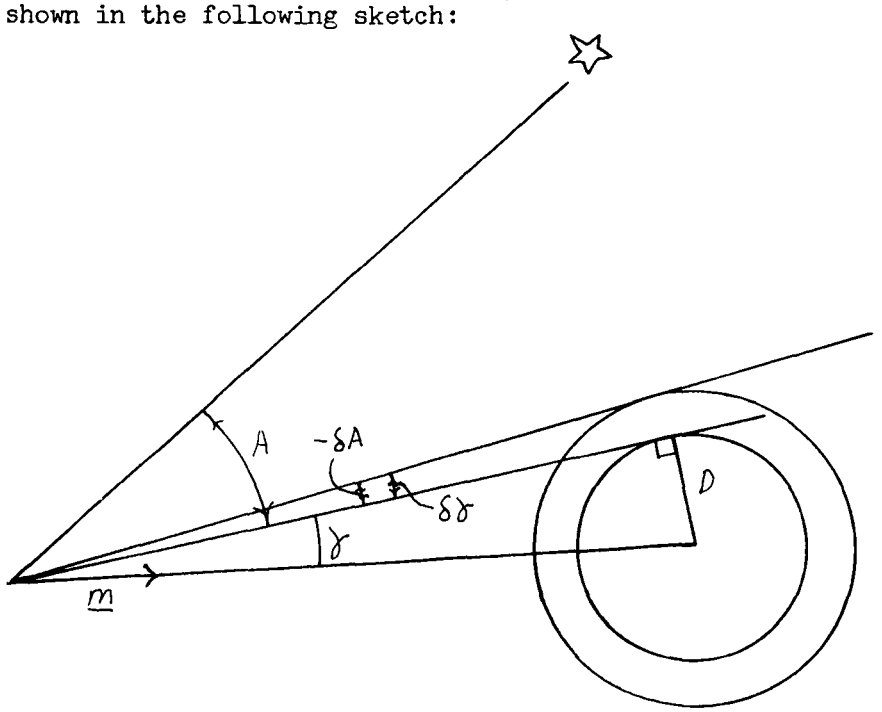


2.5.2.1.4 Star Elevation Measurement

The inaccuracies associated with the star elevation measurement are the planet diameter, planet position, and star direction. From a previous monograph, Reference 1.1, the geometry of the star elevation measurement is recalled to be:



Thus, if an error in the planet diameter is considered, the corresponding error in the elevation angle can be found by keeping all of the other parameters of the measurement fixed. The geometry for the diameter uncertainty analysis is shown in the following sketch:



Since changes in D do not change the magnitude of the angle $(A + \gamma)$, then it can be said that $\Delta A = -\Delta\gamma$. Hence, the uncertainty in the angle A is the same as the uncertainty in γ . The relationship between the uncertainty in γ and D can be found from the following relationship:

$$\sin\gamma = \frac{D}{z} \tag{5.17}$$

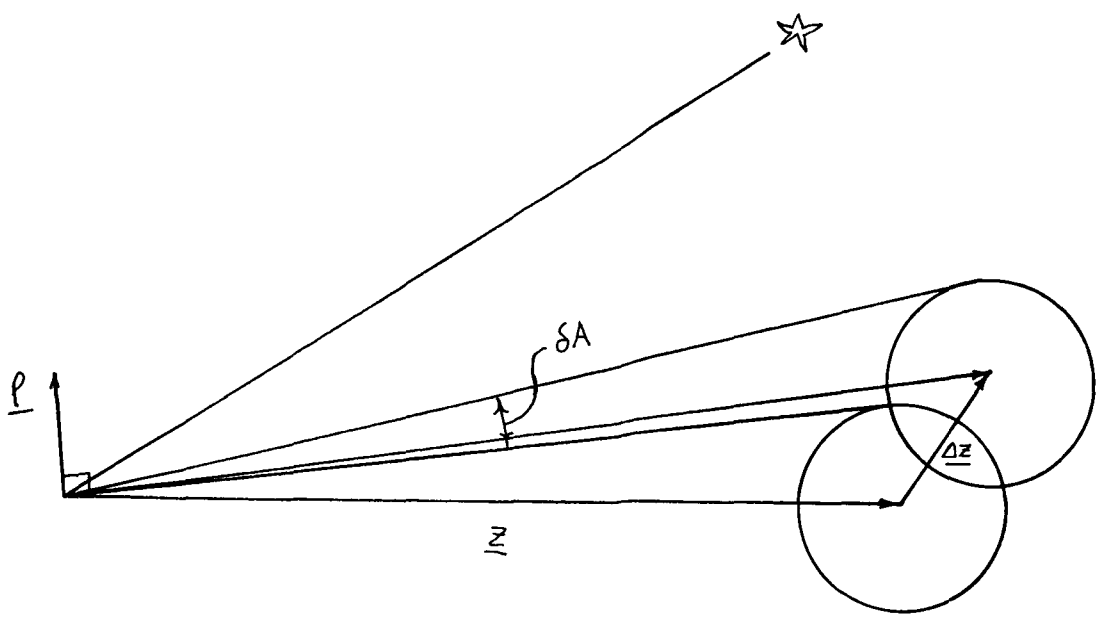
Now,

$$\frac{\partial\gamma}{\partial D} = \frac{1}{z\cos\gamma} \tag{5.18}$$

The uncertainty in the elevation angle due to the planet diameter uncertainty can now be expressed as

$$(\Delta A)_D = \frac{\Delta D}{z\cos\gamma} \tag{5.19}$$

Another uncertainty that will influence the elevation angle measurement is the uncertainty in the planet position. The following sketch shows the geometry of this condition with $\underline{\Delta z}$ being the planet position uncertainty vector and all other parameters fixed to their nominal values.



Here, the vector $\Delta \underline{z}$ can be resolved into components in the $\underline{z}$ and $\underline{\rho}$ directions. The component in the $\underline{\rho}$ direction will influence angle A, but the component in the $\underline{z}$ direction will not. Hence, the uncertainty in A due to the uncertainty in the planet position is

$$(\Delta A)_{\underline{z}} = \frac{\underline{\rho} \cdot \Delta \underline{z}}{z} \tag{5.20}$$

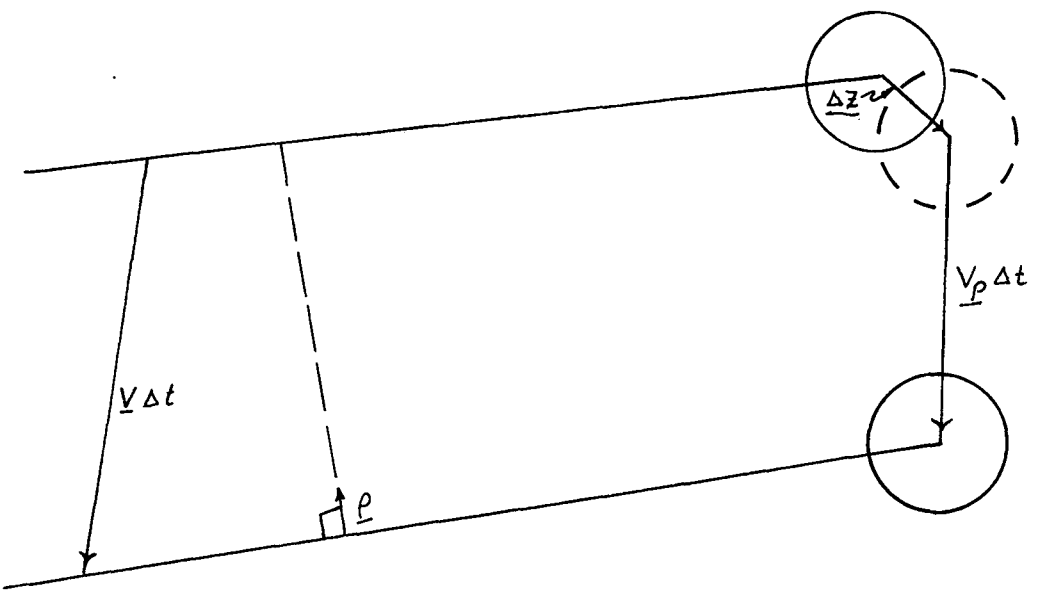
Any uncertainties in the star direction or the instrument have a direct influence on the measurement and require no further analysis. The total model uncertainty in the star elevation measurement is thus:

$$\begin{aligned} \Delta A &= (\Delta A)_D + (\Delta A)_z + (\Delta A)_{STAR} \\ &= \frac{\Delta D}{z \cos\delta} + \frac{\underline{\rho} \cdot \Delta \underline{z}}{z} + (\Delta A)_{STAR} \end{aligned} \tag{5.21}$$

2.5.2.1.5 Star Occulation Measurement

The model uncertainties that influence the star occultation measurement include the planet position, the planet velocity, the vehicle velocity, the point of tangency during occultation, and the star direction. The effect of each uncertainty will be investigated independently in order that the effective partial differential change can be found.

First, consider a change in the planet position with all other parameters being exact. The following sketch shows the geometry:



During the change in time between the expected and actual occultation, the vehicle moved a distance $\underline{V}\,\Delta t$. The planet, being slightly off the expected position, moved a distance $\underline{V}_p\,\Delta t$. Now, an equation relating the components in the $\underline{\rho}$ direction can be written as

$$\underline{\rho}\cdot\underline{V}\,\Delta t = \underline{\Delta z}\cdot\underline{\rho} + \underline{\rho}\cdot\underline{V_p}\cdot\Delta t \tag{5.22}$$

or

$$\underline{\rho}\cdot(\underline{V}-\underline{V_p})\Delta t = \underline{\Delta z}\cdot\underline{\rho} \tag{5.23}$$

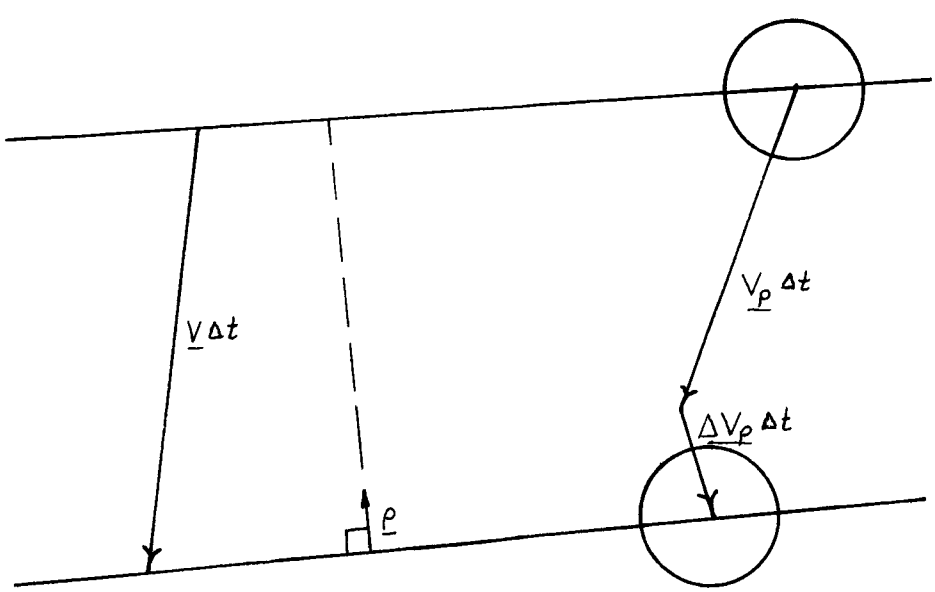
Thus, the uncertainty in the occultation time that is due to the uncertainty in the planet position can be written as

$$(\Delta t)_z = \frac{\underline{\Delta z}\cdot\underline{\rho}}{\underline{\rho}\cdot(\underline{V}-\underline{V_p})} = \frac{\Delta z\cdot\rho}{(\underline{\rho}\cdot\underline{V_R})} \tag{5.24}$$

where

$$\underline{V_R} = \underline{V} - \underline{V_p}$$

The following sketch shows the effect of an uncertainty in the planet velocity on the occultation time:



A completely analogous derivation can be performed for this case. Here $\underline{\Delta V_p}\,\Delta t$ is the additional distance that the planet travels during the time difference between the expected and the actual occultation.

The equation for the components in the $\underline{\rho}$ direction becomes:

$$\underline{\rho}\cdot\underline{V}\Delta t = \underline{\rho}\cdot\underline{V_p}\,\Delta t + \underline{\rho}\cdot\underline{\Delta V_p}\,\Delta t \tag{5.25}$$

In order to determine the differential that $\underline{\Delta V_p}$ introduces in Δt, the occultation time deviation, it is necessary to find the deviation without $\underline{\Delta V_p}$. This requirement introduces a problem since there would be no time deviation if all parameters were exactly correct during the occultation. Yet, the partial differential change is desired, requiring that all parameters other than those being considered be assumed known. The paradox can be circumvented if a fictitious δt is assumed to exist. δt is the result of some parameter being off nominal. This parameter will not be specified since the parameter being considered, $\underline{\Delta V_p}$, has already been specified. With $\underline{\Delta V_p} = \underline{0}$, equation (5.25) becomes

$$\underline{\rho}\cdot\underline{V}\,\delta t = \underline{\rho}\cdot\underline{V_p}\,\delta t \tag{5.26}$$

where δt is the fictitious time deviation. The influence of ΔV_p on Δt can now be determined by subtracting Equation (5.26) from Equation (5.25). The result is

$$\rho\cdot V(\Delta t - \delta t) = \rho\cdot V_p(\Delta t - \delta t) + \underline{\rho}\cdot\underline{\Delta V_p}\,\Delta t \tag{5.27}$$

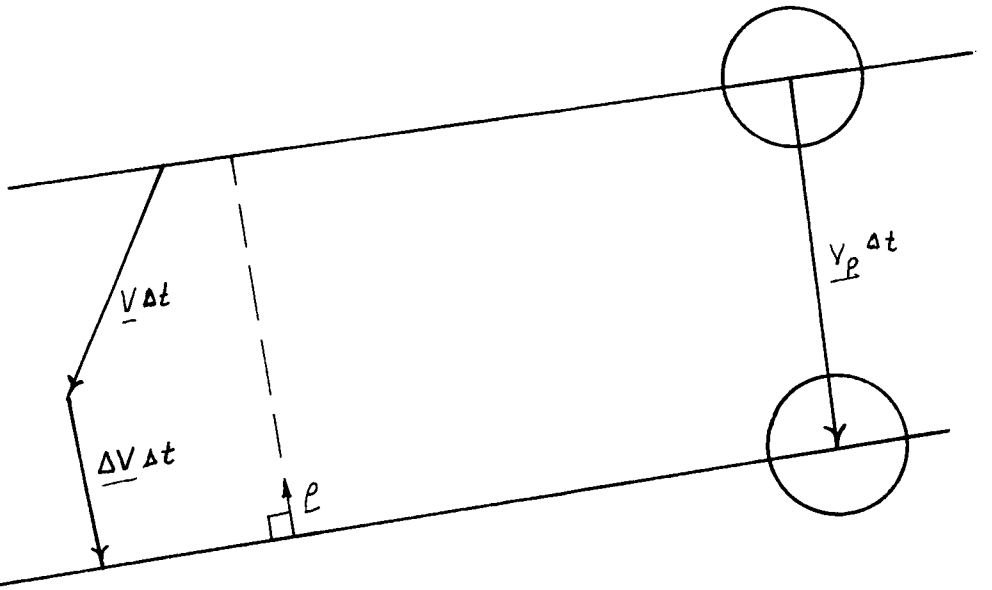
The quantity $(\Delta t - \delta t)$ can be considered to be the uncertainty in the occultation time deviation $(\Delta t)_{V_p}$ due to the uncertainty in the planet velocity.

$$(\Delta t)_{V_p} = \frac{\underline{\rho}\cdot\underline{\Delta V_p}\,\Delta t}{\underline{\rho}\cdot\underline{V_R}} \tag{5.28}$$

Where

$$\underline{V_R} = \underline{V} - \underline{V_p}$$

The uncertainty in the vehicle velocity can be handled in a manner similar to that of the planet velocity. The following sketch shows the geometry:



The equation for the components in the $\underline{\rho}$ direction becomes

$$\underline{\rho} - \underline{V}\Delta t + \underline{\rho}\cdot\Delta\underline{V}\Delta t = \underline{\rho}\cdot V_{\underline{\rho}}\,\Delta t \tag{5.29}$$

For no inaccuracy in $\underline{V}$, the previous expression becomes

$$\underline{\rho}\cdot\underline{V}\,\delta t = \underline{\rho}\cdot V_{\underline{\rho}}\,\delta t \tag{5.30}$$

where δt is again the fictitious time deviation discussed earlier. The uncertainty in the time deviation due to the vehicle velocity uncertainty can now be found as before by the difference between Equations (5.30) and (5.29). Solving for the time difference ($\Delta t - \delta t$), the result is

$$(\Delta t)_{V} = (\Delta t - \delta t) = \frac{\underline{\rho}\cdot\Delta\underline{V}\,\Delta t}{\underline{\rho}\cdot\underline{V}_{\underline{\rho}}}. \tag{5.31}$$

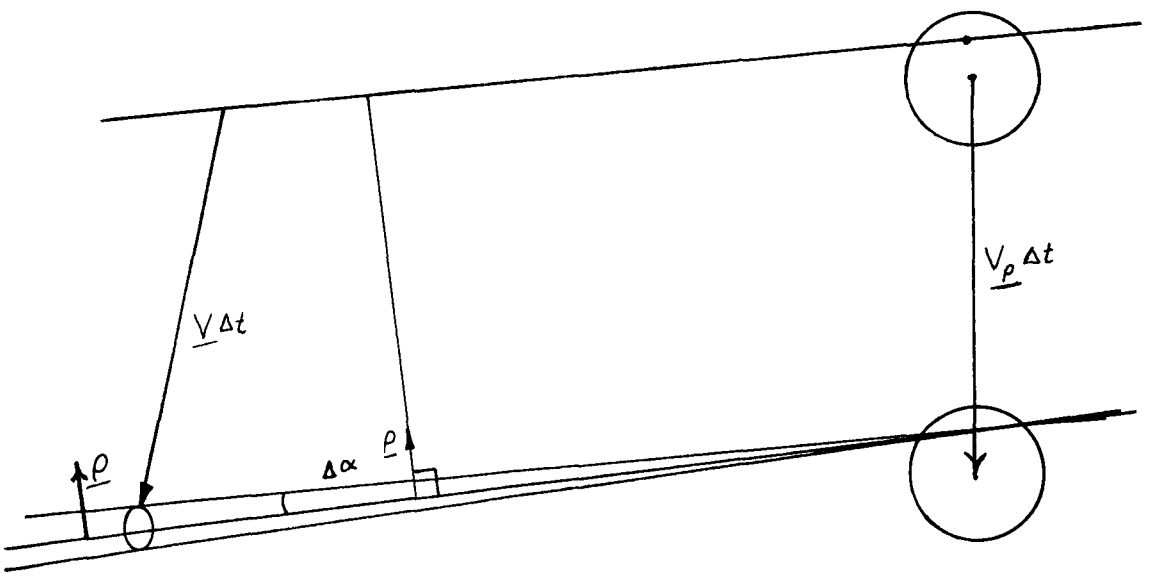
If the actual point of tangency differs from the nominal point of tangency during the occultation time, a correction term for this effect could be included in the time deviation equation. This is pointed out in Reference 1.1. Since the point of tangency cannot be determined exactly, there is an uncertainty in the time deviation due to this tangent point uncertainty. The component of the time deviation, δt , due to the change of the point of tangency is shown in Reference 1.1 to be

$$(\delta t)_E = \frac{\underline{\rho} \cdot (\underline{E}_2 - \underline{E}_1)}{\underline{\rho} \cdot \underline{V}_\rho} \tag{5.32}$$

If there are uncertainties involved with an exact determination of $\underline{E}_2$ and $\underline{E}_1$, then the difference vector, $(\underline{E}_2 - \underline{E}_1)$, has a corresponding uncertainty, $\underline{\Delta E}$. Associated with $\underline{\Delta E}$ is a time uncertainty $(\Delta t)_E$. The final expression for the deviation time uncertainty due to the tangent point uncertainty can be written immediately from Equation (5.32) as

$$(\Delta t)_E = \frac{\underline{\rho} \cdot \underline{\Delta E}}{\underline{\rho} \cdot \underline{V}_\rho} \tag{5.33}$$

The last model error to be considered for the star occultation measurement is the uncertainty in the direction of the star used for the occultation. The following sketch illustrates how such an uncertainty can influence the occultation time deviation uncertainty:



Once again the fictitious time deviation (δt) used earlier is employed so that the partial differential change can be evaluated. Note should be made of the fact that the vector between the two points of tangency is $\underline{V}_\rho \Delta t$, because all parameters other than star direction are assumed to be known exactly. If the components of the vectors in the $\underline{\rho}$ direction are written, the following expression is obtained.

$$\underline{\rho}\cdot\underline{V}\,\delta t + |\underline{N}_2|\Delta\alpha = \underline{\rho}\cdot\underline{V}_p\,\Delta t \qquad (5.34)$$

where $\underline{N}_2$ is a vector from the observer to the point of tangency. As before, the expression that is obtained with no star direction uncertainty is

$$\underline{\rho}\cdot\underline{V}\,\delta t = \underline{\rho}\cdot\underline{V}_p\,\delta t \qquad (5.35)$$

The difference in Equations (5.35) and (5.34) gives the time uncertainty due to the star direction uncertainty.

$$\underline{\rho}\cdot\underline{V}\,(\Delta t - \delta t) + |\underline{N}_2|\Delta\alpha = \underline{\rho}\cdot\underline{V}_p\,(\Delta t - \delta t) \qquad (5.36)$$

or $$\rho\cdot(V - V_p)(\delta t - \Delta t) = |\underline{N}_2|\Delta\alpha$$

Hence,

$$(\Delta t)_\alpha = \frac{|\underline{N}_2|\Delta\alpha}{\underline{\rho}\cdot\underline{V}_R} \qquad (5.37)$$

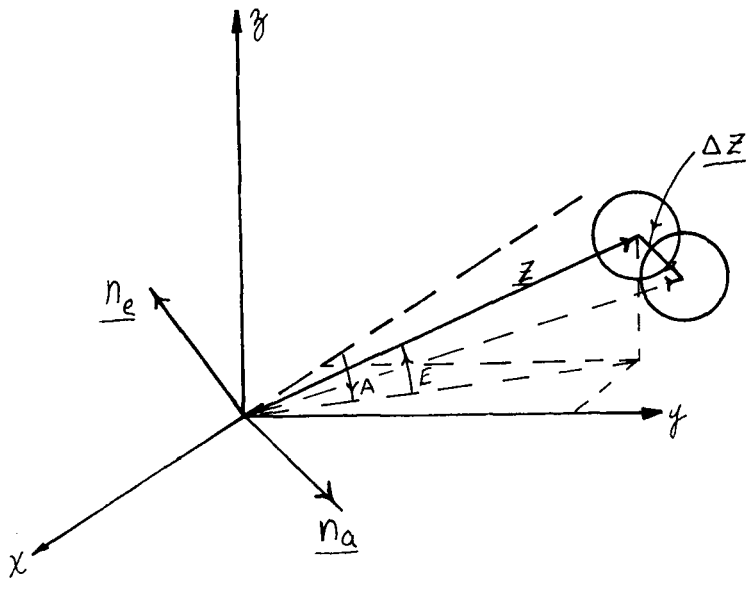
Finally, the total uncertainty for the star occultation measurement is:

$$\Delta t = (\Delta t)_z + (\Delta t)_{V_p} + (\Delta t)_V + (\Delta t)_E + (\Delta t)_\alpha$$

$$= \frac{\Delta\underline{Z}\cdot\underline{\rho}}{(\underline{\rho}\cdot\underline{V}_R)} + \frac{\underline{\rho}\cdot\Delta\underline{V}_p\,\Delta t}{(\underline{\rho}\cdot\underline{V}_R)} + \frac{\underline{\rho}\cdot\Delta\underline{V}\,\Delta t}{\underline{\rho}\cdot\underline{V}_R} + \frac{\underline{\rho}\cdot\Delta\underline{E}}{\underline{\rho}\cdot\underline{V}_R} + \frac{|\underline{N}_2|\,\Delta\alpha}{\underline{\rho}\cdot\underline{V}_R}$$

2.5.2.1.6 Elevation-Azimuth Angle Measurement

The model inaccuracies associated with the elevation and azimuth angle measurement are the inaccuracies in the planet position, landmark position, and the reference attitude of the platform from which the measurement is made. The following sketch illustrates the geometry for the case where the planet position is uncertain by an amount $\Delta\underline{Z}$ (η_a is a unit vector in the X-Y plane and is perpendicular to the projection of $\underline{Z}$ on the X-Y plane and η_e is a unit vector in the elevation angle plane and is perpendicular to $\underline{Z}$).



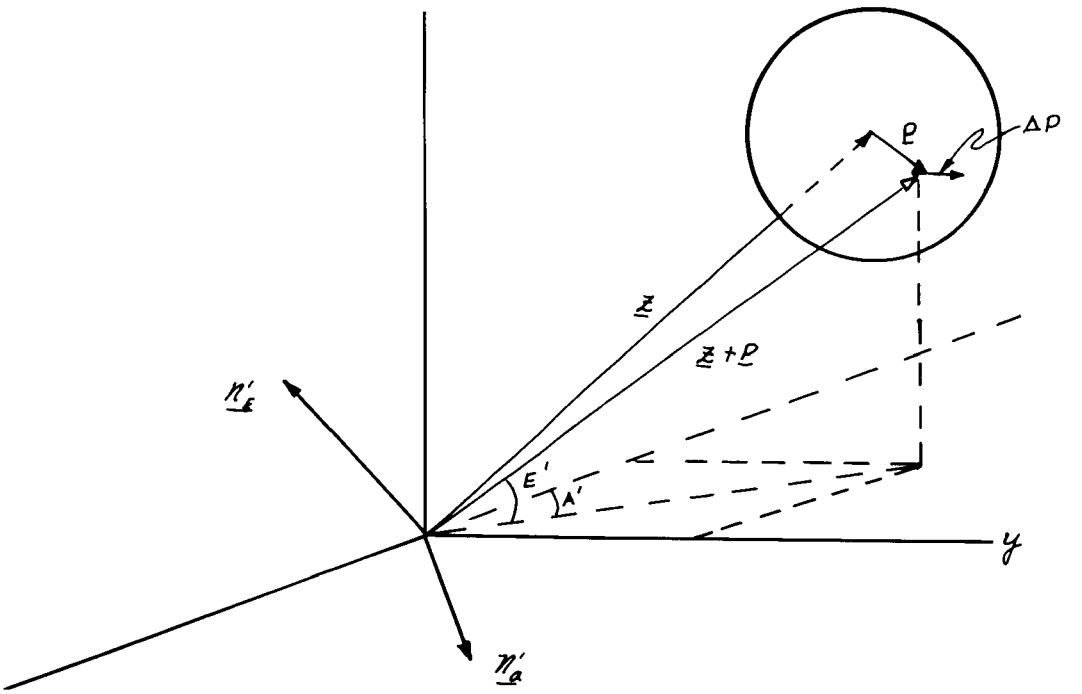
This $\Delta\underline{Z}$ can be resolved into components in the $\underline{n_a}$, $\underline{n_e}$, and $\underline{z}$ directions ($\underline{Z}$ is a vector from the vehicle to the planet center). The component in the $\underline{z}$ direction does not influence either of the two measurements, because its change is not seen from the origin. The component in the $\underline{n_a}$ direction does have an influence on the measurement of A by an amount

$$(\Delta A)_z = \frac{\underline{n_a} \cdot \Delta \underline{Z}}{|\underline{Z}| \cos E} \tag{5.38}$$

Similarly, the component in the $\underline{n_e}$ direction contributes to the measurement of E.

$$(\Delta E)_z = \frac{\underline{n_e} \cdot \Delta \underline{Z}}{|\underline{Z}|} \tag{5.39}$$

A completely analogous derivation can be performed for the uncertainty in the landmark position, ΔP .



The results are

$$(\Delta E)_P = \frac{\eta_e' \cdot \Delta P}{|\underline{z} + \underline{P}|} \qquad (\Delta A)_P = \frac{\eta_a' \cdot \Delta P}{|\underline{z} + \underline{P}| \cos E'} \qquad (5.40)$$

where E', η_e', and η_a' are defined with respect to $\underline{z} + \underline{P}$ instead of $\underline{z}$.

The uncertainty in the attitude of the platform is a function of the hardware limitations and the accuracy of the star directions with which the platform is aligned. These uncertainties can be expressed directly in terms of A and E and will not be discussed any further here. In summary,

$$\Delta A = (\Delta A)_z + (\Delta A)_p + \Delta A_{STAR}$$

$$= \frac{\eta_a \cdot \Delta \underline{z}}{|z| \cos E} + \frac{\eta_a' \cdot \Delta P}{|\underline{z} + \underline{P}| \cos E'} + \Delta A_{STAR}$$

$$\Delta E = (\Delta E)_z + (\Delta E)_p + \Delta E_{STAR}$$

$$= \frac{\eta_e \cdot \Delta z}{|z|} + \frac{\eta_e' \cdot \Delta P}{|\underline{z} + \underline{P}|} + \Delta E_{STAR}$$

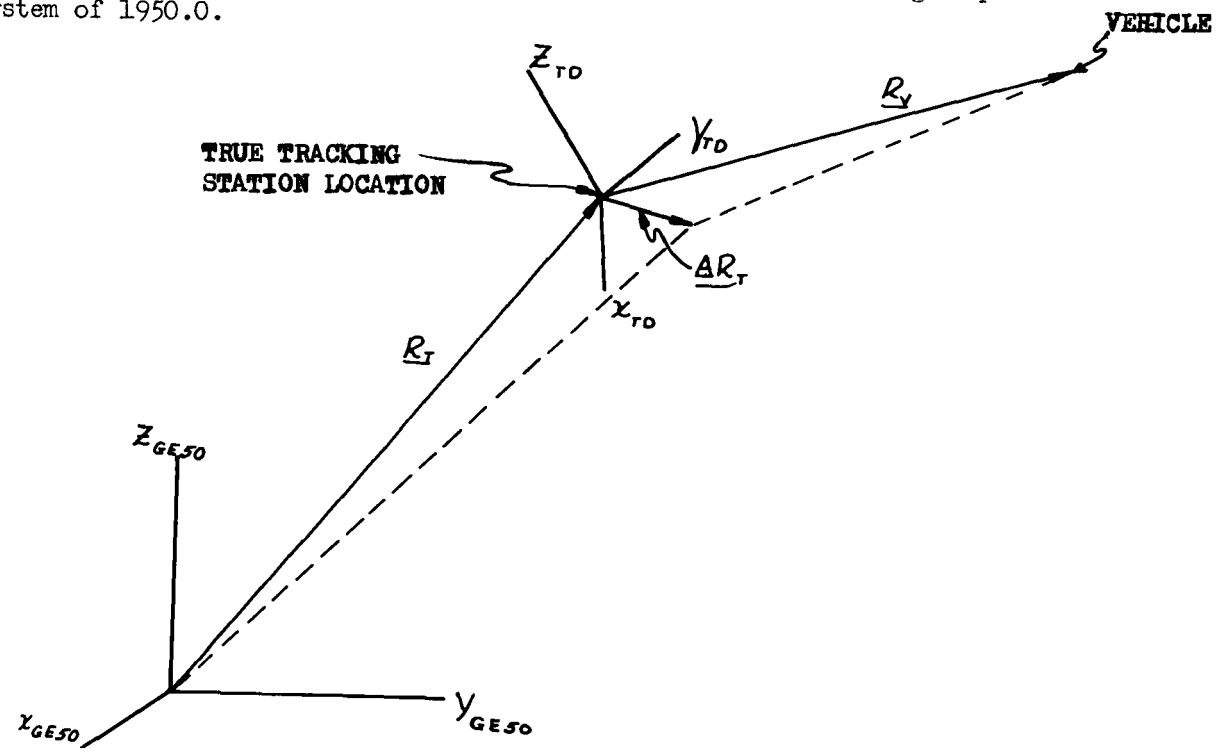
2.5.2.1.7 Tracking Station Measurement

An earth based tracking station has a position uncertainty that introduces corresponding uncertainties in both the position and velocity observables for the object that is being tracked. Since the station location is usually expressed in terms of altitude, latitude, and longitude, the uncertainty can best be described for present purposes in terms of the topodetic axis system described in Reference 5.1. In this system, the uncertainty in the altitude can be expressed in the $\hat{z}_{TD}$, or outward normal to spheroid model direction. The uncertainty in latitude and longitude can similarly be expressed in the $\hat{x}_{TD}$ or south and $\hat{y}_{TD}$ or east directions respectively. Using the geocentric spherical coordinates to specify the station location, the uncertainty in the location can be written as

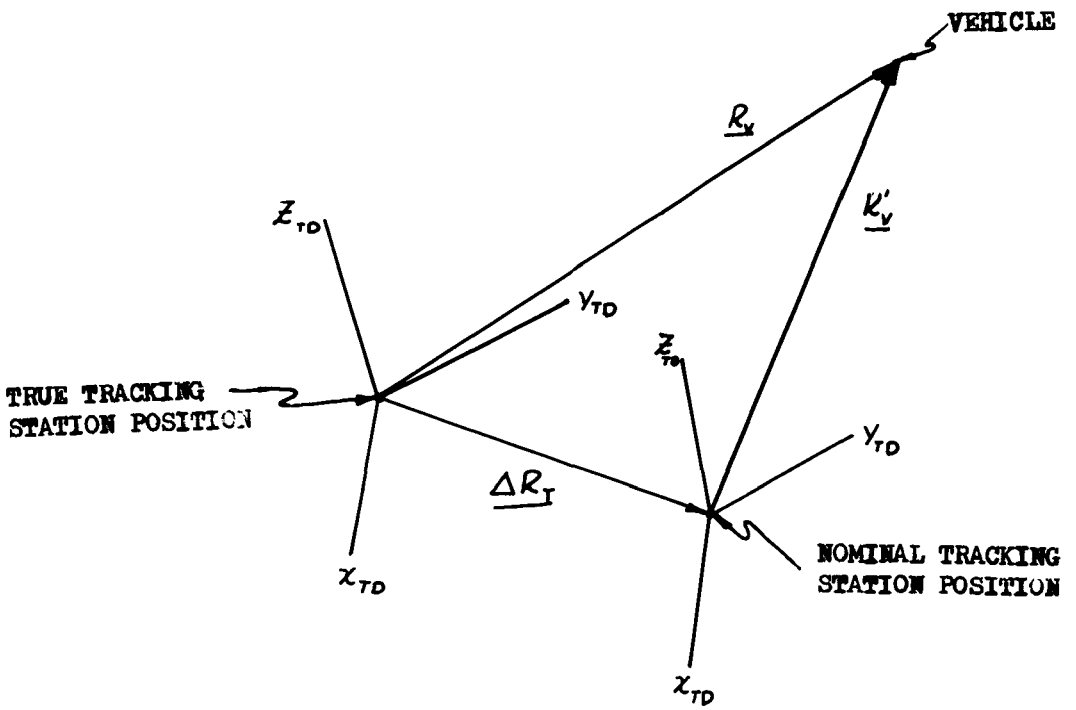
$$\underline{\Delta R}_T = R_G \, \Delta \phi'_{GC} \, \hat{x}_{TD} + R_G \cos \phi'_{GC} \, \Delta \lambda_{GG} \, \hat{y}_{TD} + \Delta H \hat{z}_{TD} \qquad (5.41)$$

where R_G = geocentric radial distance to station
ϕ'_{GC} = geocentric latitude of station
λ_{GG} = geocentric longitude of station

Now, the effect of this station location error on the observation of some object will be investigated. The position of the tracking station and the vehicle are shown below in an inertial system such as the geoequatorial system of 1950.0.

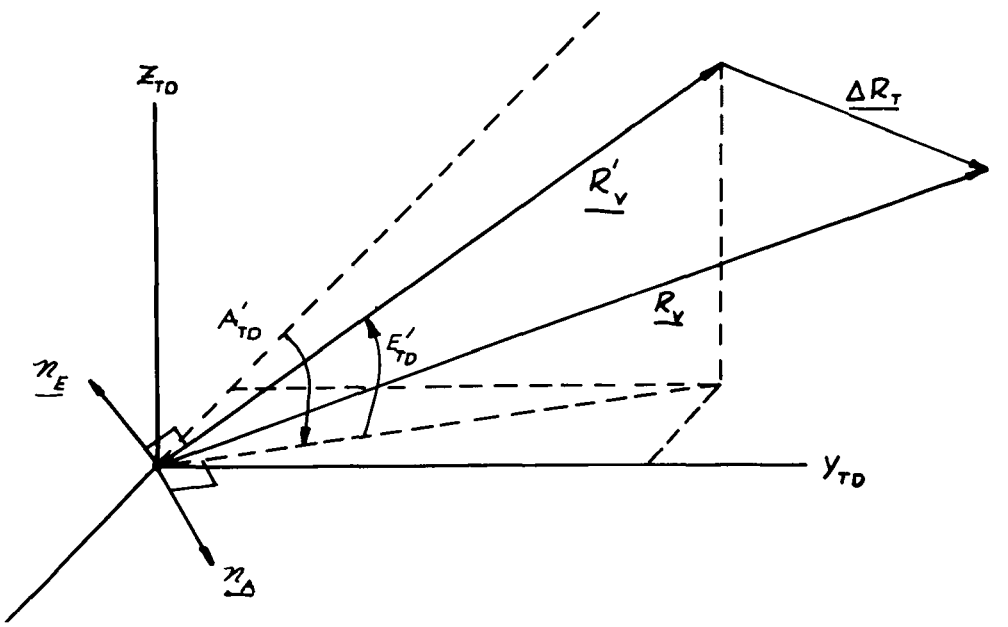


The position of the vehicle can be expressed in terms of the topodetic system by specifying the range, azimuth and elevation, (r_{TD} , A_{TD} , E_{TD}). It is desired to find the partial differential changes in r_{TD} , A_{TD} and E_{TD} that are caused by ΔR_T . This step can be accomplished by expressing the vehicle position in the true tracking station coordinate system and the nominal tracking station coordinate system.



Although the topodetic systems that could be constructed at these two distinct points would not align with each other, they are considered to be aligned in this analysis because the partial differential changes in range, azimuth, and elevation are desired as functions of the original displacement.

The vectors of interest can now be expressed in one coordinate system and the influence of ΔR_T can now be found. This is done by defining the unit vector η_A and η_E as shown in the following sketch, and resolving ΔR_T into components in the directions of R'_V , η_A , and η_E .



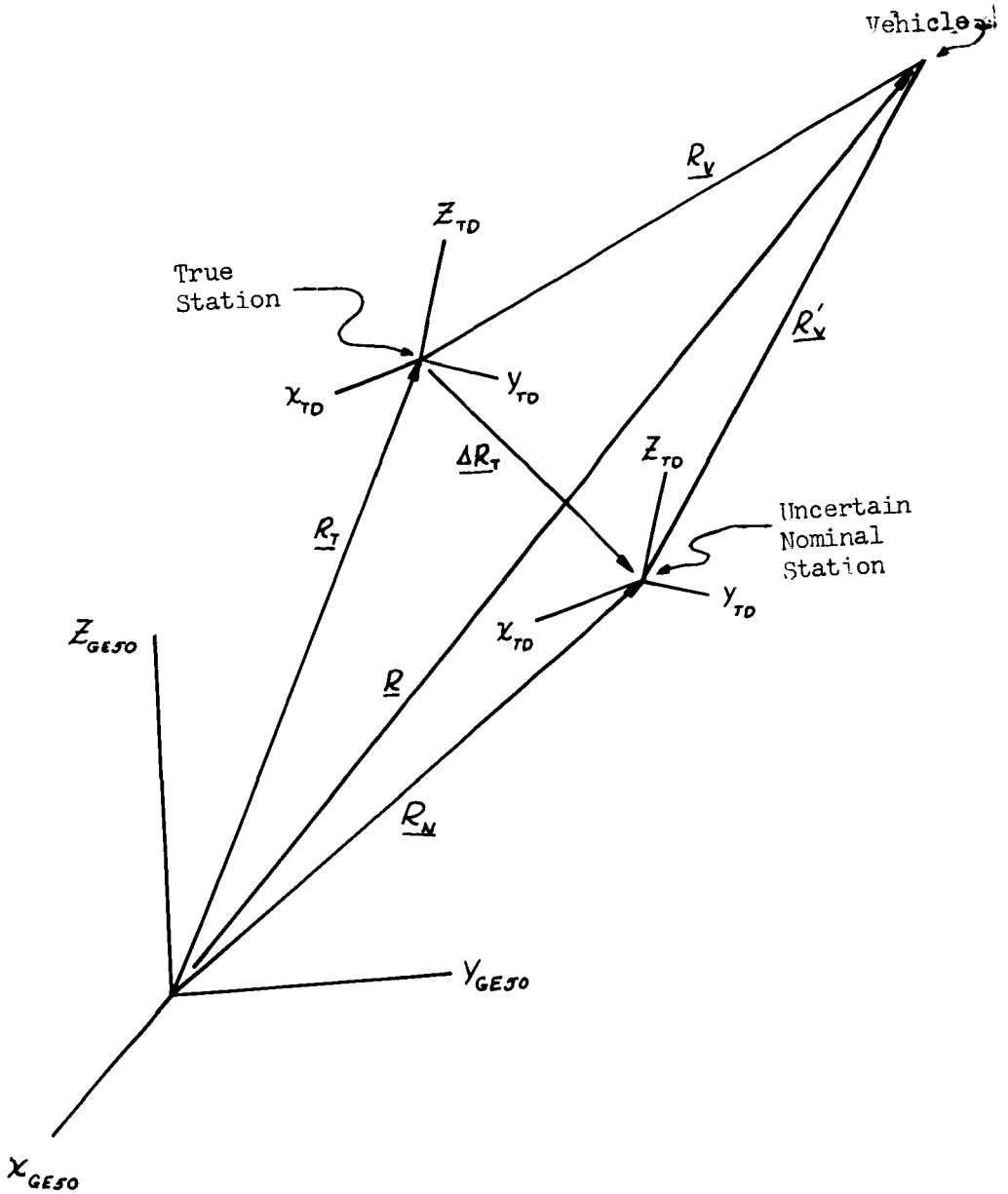
η_Δ and η_E are defined in a manner that is similar to the azimuth and elevation measurement. η_Δ is in the X-Y plane and is perpendicular to the projection of R'_v on the X-Y plane. η_E is also perpendicular to R'_v, but in the plane of the elevation angle measurement. When ΔR_T is resolved into components previously mentioned, the differential influences can be found as follows:

$$\Delta A_{TD} = \frac{\eta_a \cdot \Delta R_T}{|R'_v| \cos E_{TD}} \tag{5.42a}$$

$$\Delta E_{TD} = \frac{\eta_E \cdot \Delta R_T}{|R'_v|} \tag{5.42b}$$

$$\Delta R = \frac{\Delta R_T \cdot R'_v}{|R'_v|} \tag{5.42c}$$

The uncertainty in the velocity observables, due to the tracking station uncertainty, will now be considered. The geometry of the problem is shown in the following sketch.



As before, there are two aligned coordinate systems for the true and uncertain tracking station. The notation specifies these systems as being topodetic systems. Hence, they are fixed to the surface of the earth and experience the corresponding motion associated with the surface. The inertial reference is shown to be the geoequatorial system of 1950.0.

The object of the analysis is to determine any uncertainties in the velocity observables that are introduced by the tracking station location uncertainty. The observables that are necessary for a tracking station to determine velocity are the range, range-rate, elevation, elevation-rate, azimuth, and azimuth-rate. From the previous analysis on observable angle deviations, the true and uncertain observables can be related as follows:

$$\begin{aligned} A'_{TO} &= A_{TO} + \Delta A_{TO} \\ E'_{TO} &= E_{TO} + \Delta E_{TO} \\ R'_{V} &= R_{V} + \Delta R_{V} \end{aligned} \tag{5.43}$$

Thus, the rates of these observables may be written as

$$\begin{aligned} \dot{A}'_{TO} &= \dot{A}_{TO} + \Delta \dot{A}_{TO} \\ \dot{E}'_{TO} &= \dot{E}_{TO} + \Delta \dot{E}_{TO} \\ \dot{R}'_{V} &= \dot{R}_{V} + \Delta \dot{R}_{V} \end{aligned} \tag{5.44}$$

The quantities $\Delta \dot{A}_{TO}$, $\Delta \dot{E}_{TO}$,and $\Delta \dot{R}_{V}$ may be considered the velocity observable deviations that are introduced by the station location uncertainty and may be determined from the time derivatives of Equations (5.42). Before the time derivatives can be taken, however, the time derivatives of the vectors $\underline{\eta_a}$, $\underline{\eta_E}$, $\underline{\Delta R_T}$,and $\underline{R'_V}$ must be calculated.

Since these vectors are expressed in a moving coordinate system, the total derivatives are

$$\begin{aligned} \dot{\underline{\eta}}_a &= \frac{\partial \underline{\eta}_a}{\partial t} + \underline{\omega} \times \underline{\eta}_a \\ \dot{\underline{\eta}}_E &= \frac{\partial \underline{\eta}_E}{\partial t} + \underline{\omega} \times \underline{\eta}_E \\ \dot{\underline{\Delta R}}_T &= \frac{\partial \underline{\Delta R}_T}{\partial t} + \underline{\omega} \times \underline{\Delta R}_T \\ \dot{\underline{R}}'_V &= \frac{\partial \underline{\Delta R}_T}{\partial T} + \underline{\omega} \times \underline{R}'_V \end{aligned} \tag{5.45}$$

where $\frac{\partial}{\partial t}$ = the apparent rate of change seen by the moving observer in the topodetic system.

$\underline{\omega}$ = the rotation of the moving coordinate system which is the rotation of the earth in this case.

But, the apparent rate of change of $\underline{\eta_a}$ and $\underline{\eta_E}$ can be easily found as

$$\frac{\partial \underline{\eta_a}}{\partial t} = \underline{\dot{A}} \times \underline{\eta_a} \tag{5.46}$$

$$\frac{\partial \underline{\eta_E}}{\partial t} = \underline{\Omega} \times \underline{\eta_E}$$

where

$$\underline{\dot{A}} = -\dot{A}_{TO}\,\hat{\beta}$$

$$\underline{\Omega} = \underline{\dot{A}} + \dot{E}_{TO}\,\underline{\eta_\Delta}$$

Further, the vector $\underline{\Delta R_T}$ is always fixed in the topodetic axis system. This corresponds to the fact that the true and uncertain stations are fixed in position with respect to each other in terms of their own coordinate system. Hence,

$$\frac{\partial \underline{\Delta R_T}}{\partial t} = 0 \tag{5.47}$$

Finally, $\underline{\Delta R'_V}$ can be expressed in orthogonal components as follows

$$R'_V = \begin{bmatrix} -R'_V \cos E'_{TO} \cos A'_{TO} \\ R'_V \cos E'_{TO} \sin A'_{TO} \\ R'_V \sin E'_{TO} \end{bmatrix} = \begin{bmatrix} R'_{V_x} \\ R'_{V_y} \\ R'_{V_z} \end{bmatrix} \tag{5.48}$$

so that the vector $\frac{\partial \underline{R'_V}}{\partial t}$ can be found as

$$\frac{\partial \underline{R'_V}}{\partial t} = \begin{bmatrix} \dot{R}'_{V_x} \\ \dot{R}'_{V_y} \\ \dot{R}'_{V_z} \end{bmatrix} \tag{5.49}$$

where

$$\dot{R}'_{V_x} = f_1(R'_V, E'_{TO}, A'_{TO}, R'_V, \dot{E}'_{TO}, \dot{A}'_{TO})$$

$$\dot{R}'_{V_y} = f_2(R'_V, E'_{TO}, A'_{TO}, R'_V, \dot{E}'_{TO}, \dot{A}'_{TO})$$

$$\dot{R}'_{V_z} = f_3(R'_V, E'_{TO}, \dot{R}'_V, \dot{E}'_{TO})$$

Now that the terms in the total derivative expressions of Equation (5.45) have been defined, the differentiation of Equations (5.43) may be performed. First, the uncertainty in the azimuth rate will be derived. The expression of the azimuth deviation is

$$\Delta A_{TD} = \frac{\underline{n}_a \cdot \Delta \underline{R}_T}{R'_V \cos E_{TD}} \tag{5.50}$$

So

$$\Delta \dot{A} = \frac{\frac{d}{dt}[\underline{n}_a \cdot \Delta \underline{R}_T]}{R'_V \cos E} - \frac{[\underline{n}_a \cdot \Delta \underline{R}_T]\,[-R'_V \dot{E}(\sin E) + \dot{R}'_V \cos E]}{[R'_V \cos E]^2}$$

$$= \frac{(\dot{\underline{n}}_a \cdot \Delta \underline{R}_T) + (\underline{n}_a \cdot \Delta \dot{\underline{R}}_T)}{R'_V \cos E} - \frac{[\underline{n}_a \cdot \Delta \underline{R}_T]\,[-R'_V \dot{E} \sin E + \dot{R}'_V \cos E]}{[R'_V \cos E]^2}$$

and finally,

$$\Delta \dot{A} = \frac{[(\dot{\underline{A}} \times \underline{n}_a) + (\underline{\omega} \times \underline{n}_a)] \cdot \Delta \underline{R}_T + [\underline{n}_a \cdot (\underline{\omega} \times \Delta \underline{R}_T)]}{[R'_V \cos E]^2} \tag{5.51}$$

$$- \frac{[\underline{n}_a \cdot \Delta \underline{R}_T]\,[-R'_V \dot{E} \sin E + R'_V \cos E]}{[R'_V \cos E]}$$

In a similar manner, the expression for $\Delta \dot{E}$ can be found.

$$\Delta E = \frac{\underline{n}_E \cdot \Delta \underline{R}_T}{R'_V}$$

$$\Delta \dot{E} = \frac{\frac{d}{dt}[\underline{n}_E \cdot \Delta \underline{R}_T]}{R'_V} + \frac{-[\underline{n}_e \cdot \Delta \underline{R}_T]\,\dot{R}'_V}{[R'_V]^2} \tag{5.52}$$

$$= \frac{[\underline{n}_E \cdot \Delta \dot{\underline{R}}_T + \dot{\underline{n}}_e \cdot \Delta \underline{R}_T]}{R'_V} - \frac{[\underline{n}_E \cdot \Delta \underline{R}_T]\,\dot{R}'_V}{[R'_V]},$$

$$\dot{\underline{n}}_E = [\dot{\underline{A}} + \dot{E}\,\underline{n}_A] \times \underline{n}_E + \underline{\omega} \times \underline{n}_E$$

$$= [\dot{\underline{A}} + E\,\underline{n}_A + \underline{\omega}] \times \underline{n}_E$$

$$\Delta\dot{E} = \frac{\underline{n}_E \cdot [\underline{\omega} \times \underline{\Delta R}_T] + \left\{[\dot{\underline{A}} + \dot{E}\,\underline{n}_\Delta + \underline{\omega}] \times \underline{n}_E\right\} \cdot \underline{\Delta R}_T}{R'_V}$$

$$- \frac{[\underline{n}_E \cdot \underline{\Delta R}_T]\ \dot{R}'_V}{[R'_V]^2} \qquad (5.53)$$

The expression for $\underline{\Delta\dot{R}}$ can be obtained by differentiating

$$\Delta R = \frac{\underline{\Delta R}_T \cdot \underline{R}'_V}{R'_V}$$

or

$$\Delta\dot{R} = \frac{\frac{d}{dt}[\underline{\Delta R}_T \cdot \underline{R}'_V]}{R'_V} - \frac{[\underline{\Delta R}_T \cdot \underline{R}'_V]\ \dot{R}'_V}{[R'_V]^2} \qquad (5.54)$$

Using the expression for $\underline{\Delta\dot{R}}_T$ and $\dot{\underline{R}}'_V$ developed earlier, $\Delta\dot{R}$ becomes

$$\Delta\dot{R} = \frac{(\underline{\omega} \times \underline{\Delta R}_T) \cdot \underline{R}'_V + \underline{\Delta R}_T \cdot \left(\frac{\partial \underline{R}'_V}{\partial t} + \underline{\omega} + \underline{R}_V\right)}{R'_V} - \frac{[\underline{\Delta R}_T \cdot \underline{R}'_V]\ \dot{R}'_V}{[R'_V]^2} \qquad (5.55)$$

Equations (5.51, (5.53), and (5.55) are thus the final expressions for the uncertainty in the velocity observables due to the uncertainty in the tracking station location.

2.5.2.2 Sensor Uncertainties

In the previous section, 2.5.2.1, the measurement accuracies were related to the navigation model dimension uncertainties. This section will present the relationships that are necessary in order to convert sensor uncertainties that are not originally expressed in terms of the navigation observables to navigation observable uncertainties. Several measurements are accomplished by indirectly measuring the phase or time delay of wave forms electronically. The uncertainties of these measurements are expressed in terms of the measured quantity. If such information is to be useful in determining the covariance matrix of the observables, it must be converted to a measurement uncertainty in terms of the navigation observables. It is pointed out that many measurements are originally expressed in the corrected form, e.g., sextant angle measurements, and thus the conversion factor is 1.

The relationships derived in this section are those which are of use in determining the C matrix as discussed in the General Theory, Section 2.5.1.

2.5.2.2.1 Range Measurement Conversion

In a previous monograph, Reference 1.1, three basic methods of determining range were discussed. They were: (1) Pulse Time Dealy, (2) Frequency Modulated Continuous Wave Radar, and (3) Multiple Frequency Continuous Wave Radar. The final expressions relating the navigation observable, R, and the measured parameters are respectively:

$$R = \frac{c(t_2 - t_1)}{2} \qquad \text{(Pulse Time Delay)} \quad (5.56a)$$

$$R = \frac{f_r c}{4 f_m \Delta f} \qquad \text{(FM-CW)} \quad (5.56b)$$

$$R = \frac{c \Delta \phi}{4\pi (f_1 - f_2)} \qquad \text{(Multiple Freq. CW)} \quad (5.56c)$$

The three different techniques require three different types of measurement for the same navigation observable. In the first case, a time delay between the transmitted and reflected signal is measured. The accuracy with which the time of the occurance of the leading edge of a pulse can be measured is discussed in Reference 1.1. For convenience, the result is repeated here

$$\Delta t = \left[\frac{\tau}{2BE/N_o} \right]^{1/2} \qquad (5.57)$$

Hence, the accuracy of the range measurement can be related to the accuracy of the pulse measurement as

$$\Delta R = \frac{c}{2} \Delta t = \frac{c}{2} \left[\frac{\tau}{2BE/N_o} \right]^{1/2} \qquad (5.58)$$

If the known accuracy of the radar is given in terms of ΔR, then a conversion is not necessary. However, if the accuracy parameters are given in terms of signal-to-noise ratios, then this relationship must be used in order to convert the accuracy to navigation observable terms.

The second method of measuring range (FM-CW Radar) consists of measuring the beat frequency between the transmitted and reflected signal. The accuracy with which this frequency can be determined is discussed in Reference

1.1. The range accuracy can thus be determined as

$$\Delta R = \frac{c}{4 f_m \Delta f} \Delta f_r \tag{5.59}$$

where Δf_r is the uncertainty of the beat frequency measurement.

The third method of range measurement, Multiple Frequency C-W Radar, uses a phase measurement as the measured quantity. If the instrument uncertainty is known in terms of the phase measurement accuracy, Equation (5.56c) can be used to determine the range accuracy as

$$\Delta R = \frac{c}{4\pi(f_1 - f_2)} \Delta(\Delta\phi) \tag{5.60}$$

where $\Delta(\Delta\phi)$ is the uncertainty with which the instrument can measure phase.

2.5.2.2.2 Range-Rate Measurement Conversion

Reference 1.1 gives the range-rate measurement as

$$V_R = \pm \frac{c f_d}{2 f_o} \tag{5.61}$$

The accuracy of the range-rate measurement, therefore, depends on the ability to measure Doppler frequency. Since the accuracy for the Doppler frequency is given in Reference 1.1, the accuracy of the range-rate measurement can be found from Equation (5.61) to be

$$\Delta V_R = \frac{c}{2 f_o} \Delta f_d \tag{5.62}$$

where Δf_d is the uncertainty in the Doppler frequency measurement.

2.5.2.2.3 Angular Measurements

The uncertainties of most instruments that measure angles are usually expressed in terms of the navigation observables originally. In such cases the C matrix will contain 1's in the appropriate locations so that the uncertainty in the navigation observable due to the instrument is identically the uncertainty in the instrument. Reference 1.1 gives the accuracy to be expected from the measurement of azimuth and elevation angles by radar devices. Most other angle measuring devices have uncertainties that are intimately related to the features of the device.

2.5.3 Accuracy Data

In order to use the weighting theory that has previously been discussed, it is necessary to have a knowledge of the uncertainties that are involved in the navigation model and the sensors that are being used for the particular observation. The purpose of this section is to provide sources of data that enable the determination of these uncertainties. When feasible, the actual data will be presented in the text. Many quantities, however, require extensive tables and can only be referenced.

2.5.3.1 Planet Diameter Uncertainty

The diameters of the planets in our solar system have been measured by many individuals. For convenience, the values adopted by R.M.L. Baker, Jr. are included in this section. A compilation of other available sources of planet diameter measurements is presented in Reference 5.4 along with confidence levels.

Planet	Equatorial Diameter (km)	Polar Diameter (km)	1/f
Mercury	4,660 ± 30	?	?
Venus	12,200 ± 20	?	?
Earth	6,378.150 ± .050	6,356. ± ?	298.30 ± 0.05
Mars	6,830 ± 10	6,784 ± 24	150 ± 50
Jupiter	142,750 ± 100	133,358 ± 100	15.2 ± 0.1
Saturn	121,000 ± 100	109,138 ± 90	10.2 ± ?
Uranus	49,700 ± 100	?	?
Neptune	50,000 ± 500	49,146 ± 500	58.5 ± ?
Pluto	3,000 ± 1,000	?	?
Earth's Moon	1,738.57 ± 0.07 1,738.58 ± 0.07	1,738.21 ± 0.07	

Since the purpose of this section is to provide sources of data that estimate the uncertainty in the navigation model, the detailed information will not be discussed any further. Instead, the reader who is interested in more data as determined by other individuals is referenced to Reference 5.3,

5.5, 5.6, 5.7, 5.8, 5.9, 5.10, 5.11, 5.12, 5.13, 5.14, and 5.15.

2.5.3.2 Planet Position and Velocity Uncertainty

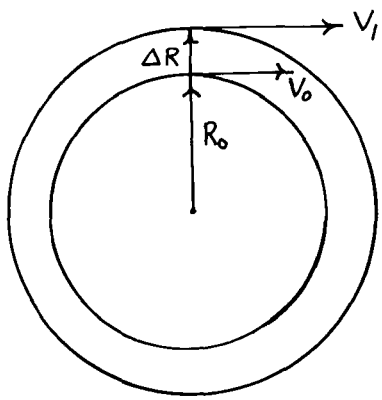
The uncertainties associated with planet positions and velocities stem from the fact that the astrodynamical constants of the solar system are not known accurately. Realizing that the accuracy of these constants will continually be improved, astronomers decided to "freeze" the values of the astrodynamical constants in order that an Ephemeris could be written. This enabled them to construct tables of the planets that matched previous observations and predicted future observations. The uncertainties in the constants prohibited the use of standard units in these catalogs, for the values would have to be modified each time a constant became known more accurately. Instead, a system of units that were independent of the uncertain constants was employed. This gave rise to the use of the astronomical unit. Although the astronomical unit is not known accurately in terms of conventional length measurements, it still provides an excellent parameter to describe the orbital behavior of the planets, because the relative motion of the bodies in our solar system can be defined very precisely in terms of this unit. Further, the Ephemeris is independent of the uncertainties of the astrodynamical constants. For these reasons, the Ephemeris can contain an extremely precise and self consistent set of data that describes the motion of the planets.

It should be noted that although the data in the Ephemeris is extremely precise, the accuracy with which it agrees with reality may not be as good. This observation results from the fact that the conversions to conventional measurements require a knowledge of the constants which are usually not known as accurately as the precision of the Ephemeris. In particular, the position of a planet may be known to an uncertainty in the eighth plance in terms of astronomical units, but the astronomical unit is only known to five places. The uncertainty in position can now be defined in terms of conventional units.

A convenient way to visualize the relationship between the uncertainty in the astronomical unit and the planet positions is to consider the Ephemeris as a catalog of the planets that differs from reality by some scale factor. The scale factor does not change the apparent angular motion of the planets, but does change their absolute positions and velocities. The uncertainty in the value of the astronomical unit can be thought of as a change in the scale of the solar system model. The accuracy of the model to reality depends on the accuracy of the astronomical unit.

The predicted position of a planet for a navigation observation can be obtained from the Ephemeris in terms of astronomical units. The uncertainty in the astronomical unit can then be used in order to find the position uncertainty. This estimate of the position uncertainty is about the best that can be achieved with the Ephemeris. If the astronomical unit is someday known much more accurately, the method of computing the uncertainty from the Ephemeris will be the same, although the uncertainty will be smaller.

The uncertainty in the velocity of a planet is needed in the star occultation measurement. A very good approximation can be made by assuming the orbit of the planet to be perfectly circular. Since the angular velocity is known, the uncertainty in the radius of the orbit can be used to calculate a very good estimate of the velocity uncertainty. Furthermore, the uncertainty in the direction of the velocity vector for circular orbits is not affected by variations in the radius. The reader interested in the exact expressions of elliptical orbit velocity sensitivities is referred to Table 3 in Section 2.4 of this monograph. It can be seen that the enumerated approximations are very good for most practical cases, if the appropriate values for the eccentricity are used. The following sketch illustrates the velocity uncertainty associated with the orbital radius uncertainty of a circular orbit.



If the known mean angular velocity of the planet about the sun is ω, the tangential velocity of the planet is

$$V_o = R_o \omega \tag{5.63}$$

Now a small change in V that would be introduced by the uncertainties in the angular velocity and the radial distance can be found as

$$\Delta V = R\Delta\omega + \omega\Delta R \tag{5.64}$$

where $\Delta\omega$ is the uncertainty in the angular velocity of the planet

ΔR is the uncertainty in the radial distance of the circular orbit

The uncertainty in ω will come from the limited precision that can be found in the Ephemeris, and the uncertainty in ΔR is the result of the limited accuracy of the astronomical unit. References 5.3, 5.5, 5.6, 5.7, 5.11, 5.12,

5.13, 5.16, 5.17, 5.18, 5.19, 5.20, 5.21, 5.22, 5.23 give some of the more recent determinations of the astronomical unit. At present, there is some discrepancy in the value for the astronomical unit calculated from radar reflections and dynamical theory. The consistency of the various radar measurements has been so good that astronomers interested in the best value of the astronomical unit have been impressed. Although some discrepancies have been explained, there are still differences between results of radar reflections and astronomical methods that cannot be explained. If successive results are obtained by the radar method under differing conditions and with planets other than Venus, then the radar method figure will probably be accepted as the standard figure.

2.5.3.3 Star Direction Uncertainty

The determination of the locations of stars in the celestial sphere is part of the field of Astrometry. It is beyond the scope of this monograph to discuss the error analysis in the star direction measurement. The reader interested in such analysis is referred to Reference 5.24. An order of magnitude estimate for the mean squared error in modern catalogs is

$$\pm 0.008 \text{ to } \pm 0.010 \text{ for right ascension (units - sec. of time)}$$

and $$\pm 0.15 \text{ to } \pm 0.20 \text{ for declination (units - sec. of arc)}$$

The unit of measurement for right ascension is the second of time and that of the declination is the second of arc. Since the earth rotates 15° in one hour, the second of time corresponds to about 15 seconds of arc. Hence, the accuracy of the right ascension is of the same order of magnitude as the declination measurement.

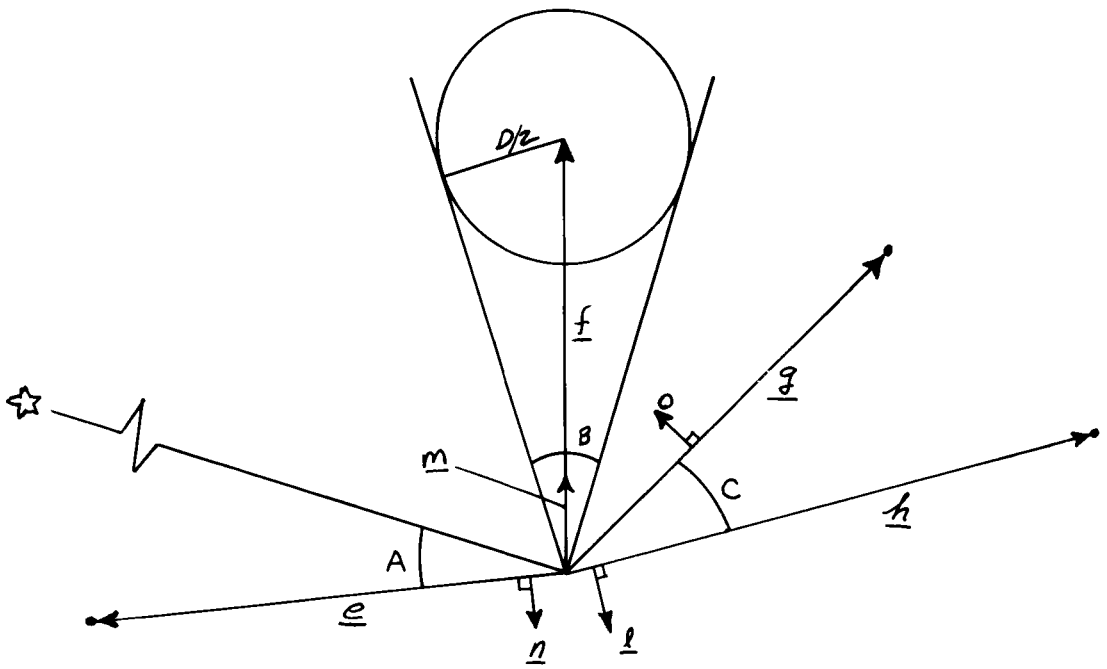
2.5.3.4 Tracking Station Location Uncertainty

The tracking station location uncertainties for various tracking station locations are given in Reference 5.2. There are several coordinate systems that are used in order to specify the station location error. Reference 5.2 uses the X-down range, Y-off range, and Z-verticle system. The system used in this monograph, however, is the topodetic coordinate system (Reference 5.1). This system was chosen because of its compatibility with the radar range, azimuth, elevation system and its compatibility with the uncertainties expressed in terms of latitude, longitude and altitude. A suitable transformation can be used to find the station uncertainty in terms of the topodetic system if not originally expressed in these coordinates.

2.5.4 Sample Problem

For convenience, a sample problem will be presented in order that the theory of Section 2.5 can be associated with a practical problem. The problem will be simple in nature for the sake of clarity. The final result, of course, is the covariance matrix of the observables.

Consider the measurement of a planet diameter, the angle between a planet and a star, and the angle between two planets.



Assuming that the flattening effect is negligible, the uncertainty in the angle of the diameter measurement, A, is

$$dB = \left[\frac{D}{f^2 \cos\left(\frac{B}{2}\right)} \quad \frac{1}{f\cos\left(\frac{B}{2}\right)}\right]\left[\begin{matrix}\underline{m}\cdot\underline{\Delta f} \\ \Delta D\end{matrix}\right] \qquad (5.65)$$

Rewriting this equation in terms of vector components yields

$$dB=\left[\frac{Dm_x}{f^2\cos\left(\frac{B}{2}\right)} \quad \frac{Dm_y}{f^2\cos\left(\frac{B}{2}\right)} \quad \frac{Dm_z}{f^2\cos\left(\frac{B}{2}\right)} \quad \frac{1}{f\cos\left(\frac{B}{2}\right)}\right]\left[\begin{matrix}\Delta f_x \\ \Delta f_y \\ \Delta f_z \\ \Delta D\end{matrix}\right] \qquad (5.66)$$

where m_x, m_y, m_z are direction cosines of unit vector m. Similarly, the uncertainties in the planet-star and planet-planet measurement are respectively:

$$dA = \begin{bmatrix} \frac{1}{e} & 1 \end{bmatrix} \begin{bmatrix} \underline{n} \cdot \underline{\Delta e} \\ \Delta A_{STAR} \end{bmatrix} \tag{5.67}$$

$$dC = \begin{bmatrix} \frac{1}{h} & \frac{1}{g} \end{bmatrix} \begin{bmatrix} \underline{\ell} \cdot \underline{\Delta h} \\ \underline{o} \cdot \underline{\Delta g} \end{bmatrix} \tag{5.68}$$

Expressed in terms of vector components, Equations (5.67) and (5.68) yield

$$dA = \begin{bmatrix} \frac{n_x}{e} & \frac{n_y}{e} & \frac{n_z}{e} & 1 \end{bmatrix} \begin{bmatrix} \Delta e_x \\ \Delta e_y \\ \Delta e_z \\ \Delta A_{STAR} \end{bmatrix} \tag{5.69}$$

$$dC = \begin{bmatrix} \frac{\ell_x}{h} & \frac{\ell_y}{h} & \frac{\ell_z}{h} & \frac{o_x}{g} & \frac{o_y}{g} & \frac{o_z}{g} \end{bmatrix} \begin{bmatrix} \Delta h_x \\ \Delta h_y \\ \Delta h_z \\ \Delta g_x \\ \Delta g_y \\ \Delta g_z \end{bmatrix} \tag{5.70}$$

Generally speaking, the measurement of angles A, B, and C will be at different epochs. But, if the covariance matrix of the observables is to be meaningful, it must be related to the covariance matrix of the model uncertainties when all components of the model uncertainty vector, $\underline{dM}$, are related to the same epoch. However, it is noted that each of the $\underline{dM}$ vectors utilized (Equations 5.66, 5.69, 5.70) must be expressed in components of the particular epoch of the measurement. (The same is true of the direction cosines of the various unit vectors, $\underline{\ell}$, $\underline{n}$, $\underline{m}$, and $\underline{o}$.) Since it is desired to have one model uncertainty vector and one model uncertainty covariance matrix, the components that change are related to an epoch at which error data is available by a transition matrix for the particular uncertainty being considered. For instance, if all uncertainties in the position of the planet used in the diameter measurement are to be related to epoch I and the planet diameter measurement is taken

at epoch II, then

$$\begin{bmatrix} \Delta f_x \\ \Delta f_y \\ \Delta f_z \\ \Delta \dot{f}_x \\ \Delta \dot{f}_y \\ \Delta \dot{f}_z \end{bmatrix}_{I} = {}^{f}\Phi_{I,II} \begin{bmatrix} \Delta f_x \\ \Delta f_y \\ \Delta f_z \\ \Delta \dot{f}_x \\ \Delta \dot{f}_y \\ \Delta \dot{f}_z \end{bmatrix}_{II} \tag{5.71}$$

where ${}^{f}\Phi_{I,II}$ is the transition matrix for Δf from Epoch II to epoch I. Although the velocity uncertainty is of no concern in this error analysis, it must be included for epoch transition calculations because of the coupling to subsequent position errors.

If the planet-star measurement is taken at epoch III and the planet-planet measurement at epoch IV, the following relations can be written:

$$\begin{bmatrix} \Delta e_x \\ \Delta e_y \\ \Delta e_z \\ \Delta \dot{e}_x \\ \Delta \dot{e}_y \\ \Delta \dot{e}_z \end{bmatrix}_{I} = {}^{e}\Phi_{I,III} \begin{bmatrix} \Delta e_x \\ \Delta e_y \\ \Delta e_z \\ \Delta \dot{e}_x \\ \Delta \dot{e}_y \\ \Delta \dot{e}_z \end{bmatrix}_{III} \tag{5.72}$$

$$\begin{bmatrix} \Delta h_x \\ \Delta h_y \\ \Delta h_z \\ \Delta \dot{h}_x \\ \Delta \dot{h}_y \\ \Delta \dot{h}_z \end{bmatrix}_{I} = {}^{h}\Phi_{I,IV} \begin{bmatrix} \Delta h_x \\ \Delta h_y \\ \Delta h_z \\ \Delta \dot{h}_x \\ \Delta \dot{h}_y \\ \Delta \dot{h}_z \end{bmatrix}_{IV} \tag{5.73}$$

$$\begin{bmatrix} \Delta g_x \\ \Delta g_y \\ \Delta g_z \\ \Delta \dot{g}_x \\ \Delta \dot{g}_y \\ \Delta \dot{g}_z \end{bmatrix}_{I} = {}^{g}\Phi_{I,IV} \begin{bmatrix} \Delta g_x \\ \Delta g_y \\ \Delta g_z \\ \Delta \dot{g}_x \\ \Delta \dot{g}_y \\ \Delta \dot{g}_z \end{bmatrix}_{IV} \tag{5.74}$$

Equations (5.66), (5.69), and (5.70) can now be written as:

$$dB = \begin{bmatrix} \frac{Dm_x}{f^2 \cos\left(\frac{\beta}{2}\right)} & \frac{Dm_y}{f^2 \cos\left(\frac{\beta}{2}\right)} & \frac{Dm_z}{f^2 \cos\left(\frac{\beta}{2}\right)} & 0 & 0 & 0 & \frac{1}{f \cos\left(\frac{\beta}{2}\right)} \end{bmatrix} \begin{bmatrix} \Delta f_x \\ \Delta f_y \\ \Delta f_z \\ \Delta \dot{f}_x \\ \Delta \dot{f}_y \\ \Delta \dot{f}_z \\ \Delta D \end{bmatrix}_{V} \tag{5.75}$$

$$dA = \begin{bmatrix} \frac{n_x}{e} & \frac{n_y}{e} & \frac{n_z}{e} & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta e_x \\ \Delta e_y \\ \Delta e_x \\ \Delta \dot{e}_x \\ \Delta \dot{e}_y \\ \Delta \dot{e}_z \\ \Delta A_{STAR} \end{bmatrix}_{III} \tag{5.76}$$

$$dC=\begin{bmatrix} \frac{\ell_x}{h} & \frac{\ell_y}{h} & \frac{\ell_z}{h} & 0\,0\,0 & \frac{o_x}{g} & \frac{o_y}{g} & \frac{o_z}{g} & 0\,0\,0 \end{bmatrix} \begin{bmatrix} \Delta h_x \\ \Delta h_y \\ \Delta h_z \\ \Delta \dot{h}_x \\ \Delta \dot{h}_y \\ \Delta \dot{h}_z \\ \Delta g_x \\ \Delta g_y \\ \Delta g_z \\ \Delta \dot{g}_x \\ \Delta \dot{g}_y \\ \Delta \dot{g}_z \end{bmatrix}_{\text{IV}} \tag{5.77}$$

Now combining the notations yields

$$dB=\left[\begin{array}{cccc:c} \frac{D m_x}{f^2\cos(\frac{\beta}{2})} & \frac{D m_y}{f^2\cos(\frac{\beta}{2})} & \frac{D m_z}{f^2\cos(\frac{\beta}{2})} & 0\,0\,0 & \frac{1}{f\cos(\frac{\beta}{2})} \end{array}\right] \left[\begin{array}{c:c} \Phi_{\text{II,I}} & 0 \\ \hdashline 0 & 1 \end{array}\right]_f \begin{bmatrix} \Delta f_x \\ \Delta f_y \\ \Delta f_z \\ \Delta \dot{f}_x \\ \Delta \dot{f}_y \\ \Delta \dot{f}_z \\ \hdashline \Delta D \end{bmatrix}_{\text{I}} \tag{5.78}$$

$$dA = \left[\begin{array}{ccc|ccc|c} \frac{n_x}{e} & \frac{n_y}{e} & \frac{n_z}{e} & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c|c} {}^{e}\Phi_{III,I} & 0 \\ \hline 0 & 1 \end{array}\right] \left[\begin{array}{c} \Delta e_x \\ \Delta e_y \\ \Delta e_z \\ \Delta \dot{e}_x \\ \Delta \dot{e}_y \\ \Delta \dot{e}_z \\ \hline \Delta A_{STAR} \end{array}\right]_I \quad (5.79)$$

$$dC = \left[\begin{array}{ccc|ccc|ccc|c} \frac{\ell_x}{h} & \frac{\ell_y}{h} & \frac{\ell_z}{h} & 0 & 0 & 0 & \frac{o_x}{g} & \frac{o_y}{g} & \frac{o_z}{g} & 0 \quad 0 \quad 0 \end{array}\right] \left[\begin{array}{c|c} {}^{h}\Phi_{III,I} & 0 \\ \hline 0 & {}^{g}\Phi_{III,I} \end{array}\right] \left[\begin{array}{c} \Delta h_x \\ \Delta h_y \\ \Delta h_z \\ \Delta \dot{h}_x \\ \Delta \dot{h}_y \\ \Delta \dot{h}_z \\ \hline \Delta g_x \\ \Delta g_y \\ \Delta g_z \\ \Delta \dot{g}_x \\ \Delta \dot{g}_y \\ \Delta \dot{g}_z \end{array}\right]_I \quad (5.80)$$

where

$${}^{f}\Phi_{II,I} = {}^{f}\Phi^{-1}_{I,II}$$

$${}^{e}\Phi_{III,I} = {}^{e}\Phi^{-1}_{I,III}$$

$${}^{h}\Phi_{IV,I} = {}^{h}\Phi^{-1}_{I,IV}$$

$${}^{g}\Phi_{IV,I} = {}^{g}\Phi^{-1}_{I,IV}$$

In order to obtain a form that is compatible with the theory presented in Section 2.5.1, these equations can be written symbolically as

$$\begin{bmatrix} dB \\ dA \\ dc \end{bmatrix} = \underline{\underline{B}}\,\Phi \begin{bmatrix} \\ \underline{dM} \\ \\ \end{bmatrix} \tag{5.81}$$

where

$$\underline{\underline{B}} = \left[\begin{array}{ccc|c|ccc|ccc|ccc|ccc} \frac{Dm_x}{f^2\cos\left(\frac{\beta}{2}\right)} & \frac{Dm_y}{f^2\cos\left(\frac{\beta}{2}\right)} & \frac{Dm_z}{f^2\cos\left(\frac{\beta}{2}\right)} & 0\;0\;0 & \frac{1}{f\cos\left(\frac{\beta}{2}\right)} & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 \\ 0\;0\;0 & 0\;0\;0 & 0 & \frac{n_x}{e}\;\frac{n_y}{e}\;\frac{n_z}{e} & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 \\ 0\;0\;0 & 0\;0\;0 & 0 & 0\;0\;0 & 0\;0\;0 & \frac{l_x}{h}\;\frac{l_y}{h}\;\frac{l_z}{h} & 0\;0\;0 & \frac{o_x}{g}\;\frac{o_y}{g}\;\frac{o_z}{g} & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 & 0\;0\;0 \end{array}\right] \tag{5.82}$$

$$\Phi = \begin{bmatrix} {}^{f}\Phi_{II,I} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & {}^{e}\Phi_{III,I} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & {}^{h}\Phi_{III,I} & 0 \\ 0 & 0 & 0 & 0 & 0 & {}^{g}\Phi_{IV,I} \end{bmatrix} \tag{5.83}$$

$$\underline{dM} = \begin{bmatrix} \Delta f_x \\ \Delta f_y \\ \Delta f_z \\ \Delta \dot{f}_x \\ \Delta \dot{f}_y \\ \Delta \dot{f}_z \\ \hline \Delta D \\ \hline \Delta e_x \\ \Delta e_y \\ \Delta e_z \\ \Delta \dot{e}_x \\ \Delta \dot{e}_y \\ \Delta \dot{e}_z \\ \hline \Delta A_{STAR} \\ \hline \Delta h_x \\ \Delta h_y \\ \Delta h_z \\ \Delta \dot{h}_x \\ \Delta \dot{h}_y \\ \Delta \dot{h}_z \\ \hline \Delta g_x \\ \Delta g_y \\ \Delta g_z \\ \Delta \dot{g}_x \\ \Delta \dot{g}_y \\ \Delta \dot{g}_z \end{bmatrix}_I \qquad (5.84)$$

It should be noted that many zeroes are introduced into the matrices in this process at the expense of space for purposes of illustration of the form of the problem. In an actual mechanization of this problem on a computer, it is more desirable to perform the computations in parts in order to avoid useless core storage of zeroes.

Now that the covariance matrix of the observables ($E(\underline{dM}\ \underline{dM}^T)$) has been "frozen" to a particular epoch, the results of Section 2.5.1 can be applied to determine the covariance matrix of observables to be

$$cov(q) = \underline{\underline{B}}\ \underline{\underline{\Phi}}\ cov(\underline{dM})\ \underline{\underline{\Phi}}^T \underline{\underline{B}}^T \qquad (5.85)$$

Note should be made that the previous results simplify if the three measurements are taken close enough in time to be considered simultaneous. If dM is known during the epoch of the three measurements, then

$$\underline{\underline{\Phi}} = \underline{\underline{I}} \text{ (the identity matrix)}$$

and then

$$cov(q) = \underline{\underline{B}}\ cov(\underline{dM})\ \underline{\underline{B}}^T \qquad (5.86)$$

where $\underline{\underline{B}}$ and $\underline{\underline{\Phi}}$ have been defined previously.

Since the measurement made in this sample problem are angles, it is not necessary to convert the measurements to navigation observables. Hence, the covariance matrix of the instrument inaccuracies can be added to the result of Equation (5.86) in order to obtain the total covariance due to the model and instrument uncertainties.

3.0 RECOMMENDED PROCEDURES

The discussions presented within the state-of-the-art review present several variations of the means which may be employed to obtain initial estimates of the orbit and subsequently improve these estimates through data smoothing. Thus, it is desirable to suggest specific applications for this material or conversely recommend combinations of these techniques for problems of common interest.

The initial fix on the trajectory as was shown can be accomplished in several distinctly different means depending on the data available. These methods are summarized in Table 3.1.

However, as was discussed, there is a distinct chance for non-negligible error in the initial orbit if raw data are employed in the process. Rather, it was shown in Section 2.1.4 that a series of preliminary operations should be performed to eliminate biases and random scatter in the data resulting from errors in the observation process, the mathematical model, and the instrument employed for the measurement.

The developments then turn to means of improving knowledge of the initial orbit and including the affects of perturbing accelerations. In particular, three types of estimators are developed:

least squares

weighted least squares

minimum variance (recursive and non-recursive)

are developed from simple concepts of "loss" or "optimality" of the fit. Subsequent developments then addressed themselves to the task of explaining the basic estimation problem, the mechanism in which the statistics of the errors, non-linear estimation techniques and/or non-linear dynamics might be introduced to the problem, and the equivalence of the estimators previously developed under certain assumptions regarding the statistics of the data. It was in this latter discussion where the concept of sufficient statistics was introduced and there it was shown that, in general, the Bayes estimator was superior to others which might be developed. It was also in this latter discussion that it was shown that the minimum variance biased estimator (MVB or Kalman) was a Bayes estimator where the statistics are Gaussian and where the loss function is simple.

Since the set of assumptions utilized in developing the MVB estimator closely corresponds to the nature of <u>most</u> orbit determination problems, it is recommended that the MVB estimator of the form developed in Sections 2.2.2.4 and 2.2.2.5 be applied. Care must be exercised, however, to confirm for any

TABLE 3.1 APPLICABLE INITIAL FIX METHODS

	Range	Range-Rate	Azimuth & Elevation
Range-Rate	6 epochs employ Eq. (1.25) without assuming $R(\dot{t}) = R(\dot{o})$	3 epochs employ Section 2.1.3 in the form of Equation (1.28) then differentially correct $\vec{r}_o$, $\vec{v}_o$ as in Eq. (1.11)	2 epochs employ Section 2.1.1 in conjunction with Lambert's Theorem (Ref. 1.3)
Range-Rate		6 epochs employ Eq. (1.26) to solve for estimate of $\vec{r}_o$, $\vec{v}_o$ then differentially correct /Eq. (1.1)/ to adjust for approximations in the formulation	2 epochs this combination of data was not investigated - but rather, it was assumed that 3 measurements would be taken and only azimuth and elevation data utilized
Azimuth & Elevation			employ Gauss's method /Eq.(1.16), (1.19), and 1.20)/ or Laplace's method /Eq. (1.6), (1.7), (1.8)/ then differentially correct for errors as in Eq. (1.11). Gauss's method is prefered for low eccentricity orbits.

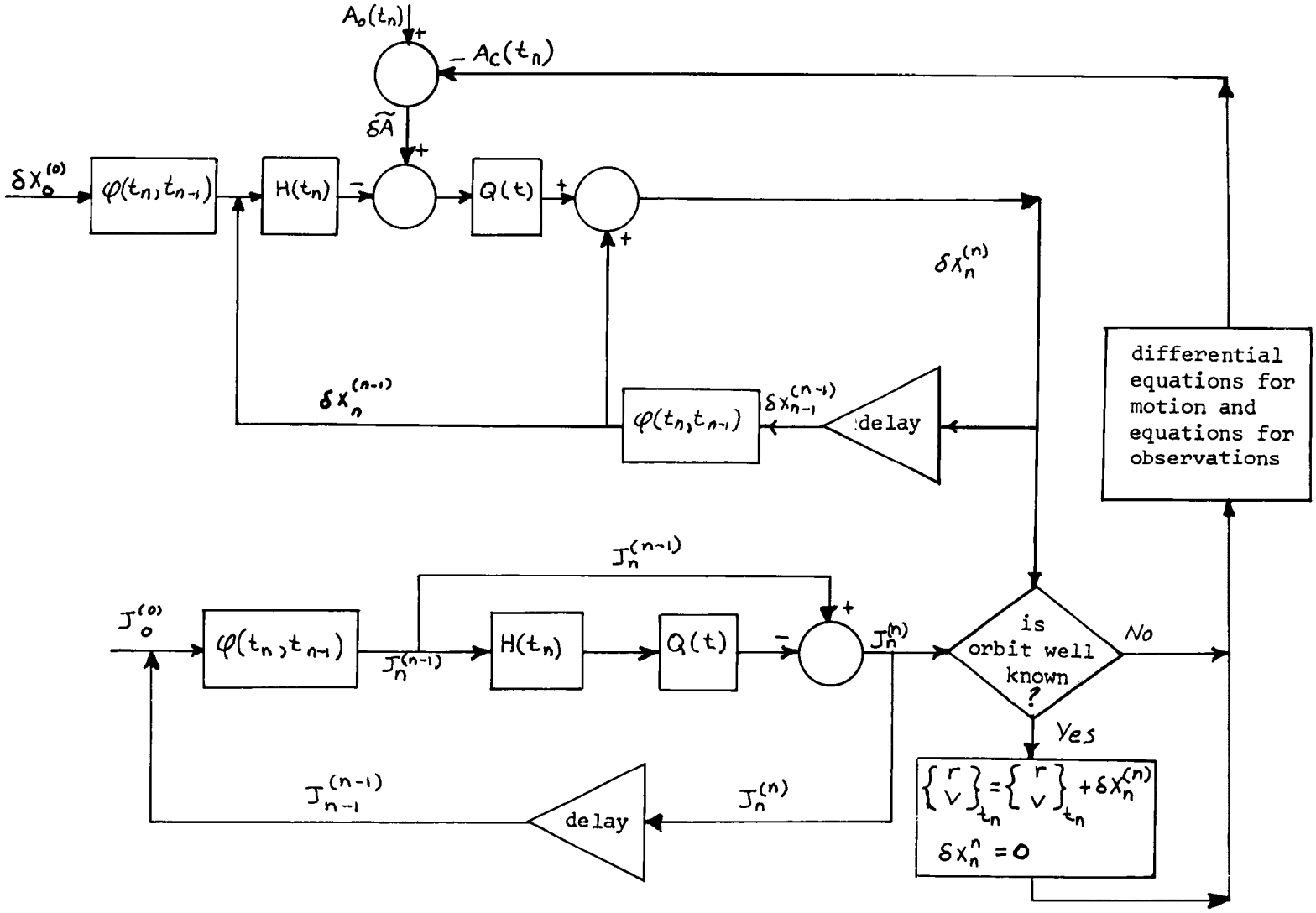


FIGURE 3.1 MECHANIZATION OF THE KALMAN ESTIMATOR

given application that the use of this estimator is justified. If not, an alogrithm must be developed from the Bayes formulation for that specific application. This test will not, in general, be easily performed; however, should the statistics of the problem be sufficiently different from those used in developing the MVB, the results could differ to a large degree. There is, however, one means of applying the MVB so that the results will be approximately correct if the number of error sources is large. This means exists due to the central limit theorem of statistics which states that the distribution of a function of random variables approaches Gaussian as the number of variables increases. Thus, a Gaussian model can be constructed which will be equivalent to the more precise process.

The mechanization of Kalman's form of the minimum variance estimator is shown in Figure 3.1. This procedure is preferred above other MVB estimators because of the fact that it provides for the utilization of initial data regarding the state of the system and because it is recursive (to minimize estimation problems and to provide a means of limiting the assumed duration for linear expansion about the nominal trajectory). One source of trouble may exist in that the procedure illustrated for updating the matrix J involves differencing. If J ever approaches the null matrix, it is thus possible for one of the eigen values of the updated matrix to be negative (due to roundoff, loss of numerical significance, etc.). Thus, for these cases, J should be updated in the manner specified in Equation (2.77).

Attention is drawn specifically to the decision function illustrated in the lower right-hand corner of the previous figure. A simple test is made to determine if the covariance matrix for the estimation errors is sufficiently small to allow the estimated state to be added to the reference trajectory so that future computations can employ a more precise reference. One such test consists of comparing the summation of the terms along the diagonal of J (or its diagonal equivalent, Appendix C) to a comparison function constructed to define an acceptable region (6 dimensions) for errors in the radius and velocity vectors. (For example,

$$F = \frac{\Delta\vec{r} \cdot \Delta\vec{r}}{\vec{r} \cdot \vec{r}} + \lambda \frac{\Delta\vec{v} \cdot \Delta\vec{v}}{\vec{v} \cdot \vec{v}} \cdot)$$

Attention then turns to the development of the state transition matrix for perturbed and conic motions. This material is intended to provide the user with a series of tools which can be applied to achieve a level of accuracy adequate for his needs. This objective is accomplished by approximating the true trajectory (for the purposes of constructing the transition matrix only) with a series of conic arcs, as described in Section 2.4.2.3, and by employing an analytic inverse property which is developed for this matrix.

The covariance matrix for the errors in the observables is constructed as the final major step in the presentation. This development relates the affects of navigation model uncertainties and sensor errors on the observables being processed for the purpose of providing the navigational filter with data required to weight the observed minus computed residuals. This step is accomplished by constructing the linear relationships between the errors in the observables and the errors in the model itself for each of the navigational

techniques discussed in Reference 1.1. These partial derivatives are then utilized in conjunction with error data for constants of the model (assumed to be normally distributed) to construct the desired covariance matrix. This procedure, while not always precise (due to non-Gaussion errors, etc.) is recommended for all cases in which more accurate data are not available.

The monograph concludes with a series of appendices which are designed to provide a background in the most normally applied statistical procedures. This material leads to the development of the concept of an error volume and the assignment of a probability of enclosure within the volume, thus allowing the covariance matrix for estimation errors (Sections 2.2 and 2.3) to be interpreted geometrically. This fact is particularly useful in discussing the results of a specific analysis.

4.0 REFERENCES

0.1 Kalman, R. E. "A New Approach to Linear Filtering and Prediction Problems," Journal of Basic Engineering. pp. 35-45 (March 1960).

0.2 Schmidt, S. F., et al. Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle. NASA, TR-R135 (1962).

0.3 Battin, R. H. "A Statistical Optimizing Procedure for Space Flight," ARS Journal. Vol. 32, No. 11, pp. 1681-1696 (November 1962).

0.4 Middleton, D. Statistical Communications Theory. New York: McGraw-Hill Book Company (1960).

0.5 Marshall, W. C. The Adjoint Method and Some Applications to Astronautics. Minneapolis-Honeywell MPG Document, U-RD-6171 (October 1961).

0.6 Danby, J. M. A. "Matrix Methods in the Calculation and Analysis of Orbits," AIAA Journal. Vol. 2, No. 1, pp. 13-19 (January 1964).

0.7 Kochi, K. C. "Exact Two-Body Error Sensitivity Coefficients," AIAA Journal. Vol. 2, No. 8, pp. 1502-1504 (August 1964).

0.8 DeVries, J. P. "Elliptic Elements in Terms of Small Increments of Position and Velocity Components," AIAA Journal. Vol. 1, No. 11, pp. 2626-2629 (November 1963).

0.9 Townsend, G. E. "The State Transition Matrix for Conic Motion," The Differential Corrections Geocentric Orbit Computation Program. NAA S&ID, SID 65-1203-1, pp. 288-306 (January 1966).

1.1 Miller, B. J. and Abbott, A. S. Observation Theory and Sensors. NAA S&ID, SID 65-1200-2 (December 1965).

1.2 Townsend, G. E. Differential Corrections Geocentric Orbit Computation Program. NAA S&ID, SID 65-1203-1, pp. 239-241 (January 1966).

1.3 Townsend, G. E. and Tamburro, M. B. The Two-Body Problem. NAA S&ID, SID 65-1200-3 (February 1966).

1.4 Danby, J. M. A. Fundamentals of Celestial Mechanics. The MacMillan Company: New York, pp. 169-172, pp. 146-148, pp. 177-182 (1962).

1.5 Herget, P. The Computations of Orbits. University of Cincinnati, published by the author (1948).

3.1 Mood, A. M. Introduction to the Theory of Statistics. 2nd Edition. New York: McGraw-Hill Book Company (1963).

3.2 Uspensky, J. V. Introduction to Mathematical Probability. New York: McGraw-Hill Book Company (1937).

4.1 Kraft, J. D. "Orbit Computation," The Orbital Flight Handbook. NASA, SP-33 (1962).

5.1 Tamburro, M. B., Abbott, A. S., and Townsend, G. E. Coordinate Systems and Time Measure. NAA S&ID, SID 65-1200-1 (October 1965).

5.2 Mann, H. P. The Accuracy of AMR Instrumentation. Air Force Missile Test Center, Patrick Air Force Base, Florida, Technical Documentary Report No. MTC-TDR-64-1 (December 1963).

5.3 Makemson, M. W., Baker, R. M. L., and Wistron, G. B. "Analysis and Standardization of Astrodynamic Constants," Astronautical Journal. Vol. VII, No. 1 (Spring 1961).

5.4 Townsend, G. E. and Kraft, J. D. Space Flight Handbooks. 3 Volumes. The Martin Company, Space Systems Division, Baltimore, Maryland, NASA, N63-21101 (1963).

5.5 MacDonald, G. J. F. "Radar Studies of the Planets," Reviews of Geophysics. Vol. 2 (November 1964).

5.6 Notes of Working Symposium on Solar System Constants. Rand Corporation, RM 3525, NASA Reproduction No. N63-18371 (May 1963).

5.7 Herrick, S., Baker, Jr., R. M. L. and Hilton, C. G. "Gravitational and Related Constants for Accurate Space Navigation," UCLA Astronomical Papers. Vol. 1, No. 14 (1958).

5.8 Kuiper, G. P. "The Diameter of Neptune," Astrophysics Journal. p. 10 (1949); "The Diameter of Pluto," Pub. A.S.P. p. 62 (1950).

5.9 Reuyl, D. "Photographic Determination of the Diameter of Mars," Astronomical Journal. p. 49 (1941).

5.10 Van de Kamp, P. "A Determination of the Diameter of Mars," Astronomical Journal. p. 38 (1928).

5.11 Allen, C. W. Astrophysical Quantities. University of London: The Athlone Press (1955).

5.12 Clemence, G. M. "On the System of Astronomical Constants," Astronomical Journal. Vol. 53, p. 178 (1947).

5.13 De Sitter, W. "On the System of Astronautical Constants," Bulletin of the Astronomical Institute of Netherlands VIII. pp. 213-231 (July 1938).

5.14 Victor, W. K., Stevens, R. and Golomb, S. W. Radar Explorations of Venus JPL, TR 32-131 (August 1961).

5.15 Pettingill, G. H., et al. "A Radar Investigation of Venus," Astronomical Journal. p. 67 (May 1962).

5.16 "Coordinates of Venus," Astronomical Papers. Vol. 15, Part 3, Nautical Almanac Office, U. S. Naval Observatory, GPO, Washington D.C. (1955).

5.17 Duncombe, R. L. "Motion of Venus," Astronomical Papers. Vol. 16, Part 1, Nautical Almanac Office, U. S. Naval Observatory, G.P.O., Washington, D.C. (1958).

5.18 Newcomb, S. "Tables of the Motion of the Earth on Its Axis and Around the Sun," Astronomical Papers. Vol. 6, Part 1, Nautical Almanac Office, U. S. Naval Observatory, G.P.O., Washington, D.C. (1898).

5.19 Newcomb, S. "Tables of the Heliocentric Motion of Mars," Astronomical Papers. Vol. 6, Part 4, Nautical Almanac Office, U.S. Naval Observatory, G.P.O., Washington D.C. (1898).

5.20 Staff, Millstone Radar Observatory, Lincoln Laboratory, MIT, "The Scale of the Solar System," Nature. Vol. 190, p. 592 (May 13, 1961).

5.21 Block, N. Computation of Lunar Positions from the Improved Brown Lunar Theory. JPL, Research Summary 36-12, Vol. 1, pp. 125-128 (January 1962).

5.22 Muhleman, D. O., Holdridge, D. B. and Block N. The Astronomical Unit Determined by Radar Reflections from Venus. JPL Technical Report, No. 32-221 (March 1962).

5.23 Thompson, J. H., et al. "A New Determination of the Solar Paralax by Means of Radar Echoes from Venus," Nature. pp. 1950-1959 (May 1961).

5.24 Podobed, V. V. Fundamental Astrometry. Chicago: University of Chicago Press (1962).

APPENDIX A

The Expected Value Theorem

Throughout the text of this monograph, linear functions of variables (normally, the observed minus computed residuals, instrument errors, etc.) have been formed, and the question has been tacitly posed, "If the original variables themselves are normally distributed, what are the distributions of the functions which are formed?" and answered, "The functions themselves are normal." This appendix was prepared to substantiate this conclusion, and to develop the mathematics describing the moments of the resulting functions.

Consider any element of a multi-dimensional vector function defined to be a linear function of a set of normally distributed parameters (not necessarily independent)

$$u_i = \sum_{j=1}^{m} a_{ij} x_j \qquad i = 1 \ldots n \tag{1}$$

and form the moment generating function of this scalar

$$m(t_1 \ldots t_n) = E\left(e^{\sum_{i=1}^{n} t_i u_i}\right) \tag{2}$$

where E denotes the expected value. Now substituting for U_i yields

$$m(t_1 \ldots t_n) = E\left(e^{\sum_{i=1}^{n} t_i \sum_{j=1}^{m} a_{ij} x_j}\right) \tag{3a}$$

$$\equiv E\left(e^{\sum_{j=1}^{m} T_j x_j}\right)$$

where

$$T_j = \sum_{i=1}^{n} a_{ij} t_i \tag{3b}$$

Thus, upon replacing the notation E by its mathematical equivalent

$$m(t_1, \ldots t_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\sqrt{|\sigma^{-1}|}}{(2\pi)^{n/2}} \exp\left[\sum_{j=1}^{m} T_j x_j \right. \tag{4}$$

$$\left. -\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} \sigma^{ij} (x_i - \bar{x}_i)(x_j - \bar{x}_j) \right] \prod_{j=1}^{m} dx_j$$

where

σ^{-1} is the inverse of the matrix of variances and covariances for the m-vector X

σ^{ij} is the element in the ith row and jth column of

$\prod_{j=1}^{m}$ is a product function

$\bar{x}_k$ denotes the average value of the kth element of the m-vector X

Now, adopting the simplifying notation $Y = X - \bar{X}$ and completing the square of the terms in the exponent, will allow integration. This process will be perfored below. First note that

$$\sum_{j=1}^{m} r_j (y_i + \bar{x}_j) - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} \sigma^{ij} y_i y_j = -\frac{1}{2}\left[\sum_{i=1}^{n} \sum_{j=1}^{m} \sigma^{ij} y_i y_j - 2 \sum_{j=1}^{m} T_j y_j\right] + \sum_{j=1}^{m} T_j \bar{x}_j$$

Now note that the portion of exponent within the brackets will be a perfect square if a term involving T is added. This term is recognized to be

$\sum_{i=1}^{n} \sum_{j=1}^{m} \sigma_{ij} T_i T_j$ (where σ_{ij} is an element of the covariance matrix, i.e., $\sigma_{ij} = \rho \sigma_i \sigma_j$ which is required so that the product terms of T_k and y_k will not be a function of either σ_{ij} or σ^{ij}.)

Performing the required addition, the exponent becomes

$$\sum_{j=1}^{m} T_j(y_j - \bar{x}_j) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} \sigma^{ij} y_i y_j = -\frac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{m} \sigma^{ij} y_i y_j - 2\sum_{j=1}^{m} T_j y_j + \sum_{i=1}^{n}\sum_{j=1}^{m} \sigma_{ij} T_i T_j\right]$$

$$+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} \sigma_{ij} T_i T_j + \sum_{j=1}^{m} T_j \bar{x}_j$$

$$= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} \sigma^{ij}\left(y_i - \sum_{k}^{m} \sigma_{ki} T_k\right) \qquad (5)$$

$$\left(y_j - \sum_{l=1}^{m} \sigma_{lj} T_l\right)$$

$$+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} \sigma_{ij} T_i T_j + \sum_{j=1}^{m} T_j \bar{x}_j$$

At this point, if this form of the exponent is substituted into equation (4) and if the final two terms (those not involving y) are factored outside of the multiple integral, the result is

$$m(t_1 \ldots t_n) = e^{\sum_{i=1}^{n} T_i \bar{x}_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} \sigma_{ij} T_i T_j} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\sqrt{|\sigma^{ij}|}}{(2\pi)^{n/2}}$$

$$\exp\left[-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} \sigma^{ij}\left(y_i - \sum_{k=1}^{m} \sigma_{ki} T_i\right)\right. \qquad (6)$$

$$\left.\left(y_i - \sum_{l=1}^{m} \sigma_{lj} T_j\right)\right] \prod_{i=1}^{m} dy_i$$

But this expression reduces to

$$m(t_1 \ldots t_n) = \exp\left[\sum_{i=1}^{n} T_i \bar{x}_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} \sigma_{ij} T_i T_j\right]$$

since the density function being integrated represents a multivariate normal distribution (with means of $\bar{y}_i = \Sigma\, \sigma_{ki} T_i \quad (k = 1, m)$ and $\bar{y}_j = \Sigma\, \sigma_{lj} T_j \quad (l = 1, m)$ and with variance σ_{ij}), and since this function is integrated over the entire region of definition.

The rth moment of the distribution for the variable U_K can now be obtained by differentiating the moment generating function with respect to t_K r times and evaluating the result for all t = 0. For the distribution just developed, this process yields

$$m(t_1 \ldots t_n) = e^{V}$$

$$\mu'_k = \left.\frac{dm(t_1 \ldots t_n)}{dt_k}\right|_{t_i=0} = \left. e^{V}\frac{dV}{dt_k}\right|_{t_i=0} = \left.\frac{dV}{dt_k}\right|_{t_i=0} = \sum_{i=1}^{n} \bar{x}_i \frac{dT_i}{dt_k} = \sum_{i=1}^{n} \bar{x}_i\, a_{ki} \qquad (7)$$

Thus,

$$\{\bar{y}\} = [A]\{\bar{x}\}$$

The variances and covariance are obtained in the same manner when it is noted that $\sigma_{kl} = \mu''_{kl} - \mu'_k \mu'_l$

But

$$\mu''_{kl} = \left.\frac{d^2 m(t_1 \ldots t_n)}{dt_k\, dt_l}\right|_{t_i=0} = e^v \left.\frac{dv}{dt_k}\frac{dv}{dt_l}\right|_{t_i=0} + e^v \left.\frac{d^2 v}{dt_k\, dt_l}\right|_{t_i=0}$$

$$= \mu'_k \mu'_l + \left.\frac{d^2 v}{dt_k\, dt_l}\right|_{t_i=0}$$

Thus

$$\sigma_{kl} = \left.\frac{d^2 v}{dt_k\, dt_l}\right|_{t_i=0}$$

$$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m}\sigma_{ij}\left[\frac{dT_i}{dt_k}\frac{dT_j}{dt_l} + \frac{dT_j}{dt_k}\frac{dT_i}{dt_l}\right]$$

$$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m}\sigma_{ij}\left[a_{ki}\, a_{lj} + a_{kj}\, a_{li}\right]$$

But since the summation is performed over all i and all j, this result reduces to

$$\sigma_{kl} = \sum_{i=1}^{n}\sum_{j=1}^{m}\sigma_{ij}\left[a_{ki}\, a_{lj}\right]$$

Finally, this result is equivalent to

$$[\sigma]_y = A\,[\sigma]_x\,A^T \qquad (8)$$

where $[\sigma]$ = a real symmetric matrix of variances and covariances

$$\sigma_{ii} = \sigma_i^2$$

$$\sigma_{ij} = \rho \sigma_i \sigma_j$$

Equations (7) and (8) are basic to most of the discussion of this monograph and studies of error propagation in general. However, their application requires knowledge that the process being analyzed approximates the assumptions made in this derivation to an acceptable degree. This assumption was that all of the variables X_i were normally distributed (had no higher moments).

It is interesting to note that the general linear transformation

$$\underline{u} = A\underline{X}$$

conserves the moments such that

$$\bar{\underline{u}} = A\bar{\underline{X}}$$

and

$$E\left[(\underline{u}-\bar{\underline{u}})(\underline{u}-\bar{\underline{u}})^T\right] = AE\left[(\underline{X}-\bar{\underline{X}})(\underline{X}-\bar{\underline{X}})^T\right]A^T = A\sigma A^T$$

However, in this case, higher order moments become extremely difficult to compute. This fact is the underlying reason that most of the "simple" estimation formulations and propagation techniques concern themselves with only the first and second moments.

APPENDIX B

COMPUTATION OF MARGINAL DENSITY FUNCTION

For problems in which the probability associated with a particular subset of the variates of the multi-variate normal density function is desired, it is necessary to modify the approach of the previous analysis. This approach evolves from the fact that the desired probability is defined by integrating the density function $f(X_1 \ldots X_m)$ over the volume of the m-dimensional volume element; i.e.,

$$P(|x_1|<a\ ,\ |x_2|<b \cdots |x_m|<c) = \int_V f(x_1 \cdots x_s)\, dx_1 \cdots dx_m$$

where the x_i are not statistically independent. The question at hand is thus "how can the density function $f(X_1 \ldots X_m)$ be obtained from the more general form $f(X_1 \ldots X_n)$, $n > m$" or "what is the marginal density function $f(X_1 \ldots X_m)$."

First it is noted that the definition of a marginal density is

$$f(x_1 \ldots x_m) = \int_{-\infty}^{\infty} \cdots \int f(x_1 \ldots x_n)\, dx_{m+1} \ldots dx_n$$

where

$$f(x_1 \ldots x_n) = \left(\frac{1}{2n}\right)^3 \sqrt{|\sigma^{ij}|}\ e^{-\frac{1}{2}\left[(\underline{x}-\bar{\underline{x}})^T \sigma^{ij} (\underline{x}-\bar{\underline{x}})\right]}$$

$$\sigma^{ij} = [\sigma_{ij}]^{-1}$$

$\underline{x} - \bar{\underline{x}}$ = vector composed of $x_1, \ldots, x_n$

Because of problems of functional complexity, consider the transformation $\underline{y} = \underline{x} - \bar{\underline{x}}$, and complete the square in the exponent of the terms involving $y_K\ (m+1 \le K \le n)$

$$(\underline{x}-\bar{\underline{x}})^T \sigma^{ij} (\underline{x}-\bar{\underline{x}}) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma^{ij} y_i y_j$$

$$= \sigma^{KK} y_K^2 + \sum_{\substack{i=1 \\ \neq K}}^{n} \sigma^{Ki} y_K y_i$$

$$+ \sum_{\substack{i=1 \\ \neq K}}^{n} \sigma^{iK} y_i y_K + \sum_{\substack{i=1 \\ \neq K}}^{n} \sum_{\substack{j=1 \\ \neq K}}^{n} \sigma^{ij} y_i y_j$$

$$= \sigma^{KK}\left(y_K^2 + 2 y_K \frac{1}{\sigma_{KK}} \sum_{\substack{i=1 \\ \neq K}}^{n} \sigma^{Ki} y_i\right) + \sum_{\substack{i=1 \\ \neq K}}^{n} \sum_{\substack{j=1 \\ \neq K}}^{n} \sigma^{ij} y_i y_j$$

$$(\underline{x}-\underline{\bar{x}})^T \sigma^{ij} (\underline{x}-\underline{\bar{x}}) = \sigma^{KK}\Big(y_K + \frac{1}{\sigma^{KK}} \sum_{\substack{i=1 \\ \neq K}}^{n} \sigma^{Ki}\Big)^2$$

$$- \frac{1}{\sigma^{KK}} \Big(\sum_{\substack{i=1 \\ \neq K}}^{n} \sigma^{Ki} y_i \Big)^2 + \sum_{\substack{i=1 \\ \neq K}}^{n} \sum_{\substack{j=1 \\ \neq K}}^{n} \sigma^{ij} y_i y_j$$

But the second term can be written as

$$\frac{1}{\sigma^{KK}} \Big(\sum_{\substack{i=1 \\ \neq K}}^{n} \sigma^{Ki} y_i \Big)^2 = \frac{1}{\sigma^{KK}} \sum_{\substack{i=1 \\ \neq K}}^{n} \sum_{\substack{j=1 \\ \neq K}}^{n} \sigma^{Ki} \sigma^{Kj} y_i y_j$$

so that the exponent becomes

$$\sigma^{KK}\Big(y_K + \frac{1}{\sigma^{KK}} \sum_{\substack{i=1 \\ \neq K}}^{n} \sigma^{Ki}\Big)^2 + \sum_{\substack{i=1 \\ \neq K}}^{n} \sum_{\substack{j=1 \\ \neq K}}^{n} \Big(\sigma^{ij} - \frac{\sigma^{Ki}\sigma^{Kj}}{\sigma^{KK}}\Big) y_i y_j$$

Now again transforming variables,

$$Z \equiv y_K + \frac{1}{\sigma^{KK}} \sum_{\substack{i=1 \\ \neq K}}^{6} \sigma^{Ki}$$

$$\tilde{\sigma}^{ij} \equiv \sigma^{ij} - \left[(\sigma^{Ki}\sigma^{Kj})/\sigma^{KK}\right]$$

the exponent becomes

$$y^T \sigma^{ij} y = \sigma^{KK} Z + \sum_{\substack{i=1 \\ \neq K}}^{n} \sum_{\substack{j=1 \\ \neq K}}^{n} \tilde{\sigma}^{ij} y_i y_j \ .$$

But the function Z is involved in such a manner that it can now be conveniently integrated out of the density function since

$$f(y \cdots y_n) = \int_{-\infty}^{\infty} f(y_1 \cdots y_m, z, y_{m+1} \cdots y_n)\, dz$$

$$= \Big(\frac{1}{2\pi}\Big)^{5/2} \frac{\sqrt{|\sigma^{ij}|}}{\sigma^{KK}} \, e^{-1/2\left[\sum_{\substack{i=1 \\ \neq K}}^{6} \sum_{\substack{j=1 \\ \neq K}}^{6} \tilde{\sigma}^{ij} y_i y_j\right]}$$

The question is now what is $\tilde{\sigma}^{ij}$, and can it be constructed in a simple manner from σ^{ij} or σ_{ij}. Consider the matrix product

$$\left[\sigma^{ij}\right]\left[\sigma_{ij}\right] = \left[I\right]$$

or

$$\sum_{m=1}^{n} \sigma^{im} \sigma_{mj} = \delta_{ij}$$

Now consider the definition of $\tilde{\sigma}^{ij}$ and the same produce. Assume the sum is δ_{ij}

$$\sum_{\substack{m=1 \\ \neq K}}^{n} \sigma_{mj} \tilde{\sigma}^{im} = \sum_{\substack{m=1 \\ \neq K}}^{n} \sigma_{mj} \Big(\sigma^{im} - \frac{\sigma^{Ki}\sigma^{Km}}{\sigma^{KK}}\Big) = \delta'_{ij}$$

$$\sum_{\substack{m=1\\ \neq K}}^{n} \sigma_{mj}\,\tilde{\sigma}^{im} = \sum_{\substack{m=1\\ \neq K}}^{n} \sigma_{mj}\,\sigma^{im} - \frac{\sigma^{Ki}}{\sigma^{KK}} \sum_{\substack{m=1\\ \neq K}}^{n} \sigma_{mj}\,\sigma^{Km}$$

$$= \left(\delta_{ij} - \sigma_{Kj}\,\sigma^{iK}\right) - \frac{\sigma^{Ki}}{\sigma^{KK}}\left(\delta_{Kj} - \sigma_{Kj}\,\sigma^{KK}\right)$$

where $j \neq K$. Thus

$$\sum_{\substack{m=1\\ \neq K}}^{n} \sigma_{mj}\,\tilde{\sigma}^{im} = \delta_{ij} - \sigma_{Kj}\,\sigma^{iK} + \sigma_{Kj}\,\sigma^{iK} \qquad \begin{matrix} i \neq K \\ j \neq K \end{matrix}$$

or

$$\left[\delta_{ij}\right]_{\substack{i\neq K\\ j\neq K}} = \left[\delta'_{ij}\right]$$

This relation requires that the matrix $\tilde{\sigma}^{ij}$ be the matrix formed by inverting the matrix constructed from σ_{ij} by deleting the row and column containing σ^{KK}.

Two further observations are also possible. First, the resultant density function is that of a multi-variate normal distribution; and second, since σ^{KK} is the ratio of the cofactor of σ^{KK} divided by the determinant σ_{ij},

$$\sqrt{\frac{|\sigma^{ij}|}{\sigma^{KK}}} = \sqrt{\frac{1}{\sigma^{KK}\,|\sigma_{ij}|}} = \sqrt{\frac{1}{|\tilde{\sigma}_{ij}|}} = \sqrt{|\tilde{\sigma}^{ij}|}\ .$$

This process can be continued indefinately to construct the marginal density function for any set of variables of interest. As an example, the marginal covariances for position (or velocity) errors can be constructed from the total covariance matrix for both position and velocity errors as follows:

$$\left[\sigma_{ij}\right] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \cdots \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \cdots \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \cdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \end{bmatrix} \qquad \left[\tilde{\sigma}_{ij}\right] = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$$

Since the result of this process is a normal distribution, a 3-dimensional surface which contains the variables ($\Delta\vec{r}$ or $\Delta\vec{V}$) to a specified probability level can be generated by

1. Performing a coordinate transformation to produce a set of uncorrelated error axes (diagonalization of the covariance matrix discussed in Appendix C).

2. Integrating the resultant distribution (Appendix C) to evaluate the radius of the equivalent sphere.

3. Using the radius of the equivalent sphere as a constraint on the values, which can be attained by the variables measured along the three principal axes, generate the 3-D ellipsoid.

4. Transforming the coordinate system to construct the error ellipsoid in the desired coordinate system.

APPENDIX C

DIAGONALIZATION OF REAL SYMMETRIC MATRICES AND THE DEVELOPMENT OF PROBABILITY ELLISOIDS

Previous appendices have proved that deviations resulting from any linear combination of Gaussian errors will be elements of a multivariate, normal distribution. Since this is the case, the probability that a given value of a variable lies in a given region can be found by integrating the joint probability density function and the variance, covariance matrix which is real and symmetric can be diagonalized by an orthogonal transformation to display the eigen values on the principal diagonal. To accomplish these objectives (in reverse order), consider the mathematical representation of the density function of an error vector:

$$f(\bar{X}) = \left(\frac{1}{2\pi}\right)^{k/2} \sqrt{|\sigma^{ij}|}\, e^{-\frac{1}{2}\left[(x-\bar{x})^T \sigma^{ij} (x-\bar{x})\right]}$$

where: σ^{ij} is the inverse of σ_{ij} and is also symmetric.

Now, consider a general symmetric matrix M (M can be σ_{ij}, etc.) and the set of equations

$$(m - \lambda I) X = 0 \tag{1}$$

where M is an (M, M) symmetric matrix; I the corresponding identity matrix; λ is a vector of scalar parameters; and X is an (m, 1) column matrix. Let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be the eigen values of the matrix M and let the corresponding eigen vectors be denoted by $\underline{X}_1, \underline{X}_2, \ldots, \underline{X}_n$. That is to say, the λ_α and $\underline{X}_\alpha$ satisfy the equations

$$(m - \lambda_\alpha I) X_\alpha = 0 \qquad \alpha = 1, 2, \ldots, n. \tag{2}$$

a corresponding set of equations is represented by

$$(m - \lambda_\beta I) X_\beta = 0 \qquad \beta = 1, 2, \ldots, n. \tag{3}$$

Thus, from Equations (2) and (3)

$$m X_\alpha = \lambda_\alpha X_\alpha$$
$$m X_\beta = \lambda_\beta X_\beta$$

If the first of these last equations is multiplied by $\underline{X}^T_\beta$ (the transpose of X_β) and the second $\underline{X}^T_\alpha$, the result is:

$$X_\beta^T M X_\alpha = \lambda_\alpha X_\beta^T X_\alpha \tag{4}$$

$$X_\alpha^T M X_\beta = \lambda_\beta X_\alpha^T X_\beta \tag{5}$$

but Equation (5) can be transposed as

$$X_\beta^T M^T X_\alpha = \lambda_\beta X_\beta^T X_\alpha \tag{6}$$

which, since M is symmetric, reduce to

$$X_\beta^T M X_\alpha = \lambda_\beta X_\beta^T X_\alpha \tag{7}$$

Subtracting Equation (7) from Equation (4) now gives

$$(\lambda_\alpha - \lambda_\beta)(X_\beta^T X_\alpha) = 0 \tag{8}$$

Hence, for $\lambda_\alpha \neq \lambda_\beta$

$$X_\beta^T X_\alpha = 0 \tag{9}$$

This last equation is expressed by the statement "the column vectors X_α and X_β are orthogonal if $\lambda_\alpha \neq \lambda_\beta$ for $\alpha \neq \beta$. Since Equations (2) and (3) do not determine the values of X_α and X_β uniquely, that is, the equations are still satisfied if X_α and X_β are multiplied by arbitrary scalar constants, it is possible to select those arbitrary scalars so as to normalize the vectors. That is

$$X_\alpha^T X_\alpha = 1 \qquad X_\beta^T X_\beta = 1 \qquad \text{for } \alpha, \beta = 1, 2, \ldots, n. \tag{10}$$

Equations (9) and (10) can be expressed by the single equation

$$X_\alpha^T X_\beta = \delta_{\alpha\beta} \quad ; \quad \alpha, \beta = 1, 2, \ldots, n \tag{11}$$

where the Kronecker delta ($\delta_{\alpha\beta}$) is defined by

$$\delta_{\alpha\beta} = 1 \qquad \text{for} \qquad \alpha = \beta$$

and (12)

$$\delta_{\alpha\beta} = 0 \qquad \text{for} \qquad \alpha = \beta$$

Substituting into Equation (5) from Equation (11) now yields

$$\underline{X}_{\alpha}^{T} m \underline{X}_{\beta} = \lambda_{\beta} \delta_{\alpha\beta} \qquad (\alpha, \beta = 1, 2, \ldots, n) \tag{13}$$

The relationships expressed by Equations (11) and (13) can be exhibited in the more conventional matrix form by deferring a compound matrix of eigen vectors

$$\mathbb{X} \equiv (X_1, X_2, \ldots, X_n) \tag{14}$$

in terms of Equation (11) becomes

$$\mathbb{X}^{T} \mathbb{X} = I \tag{15}$$

From Equation (15),

$$\mathbb{X}^{T} = \mathbb{X}^{-1} \tag{16}$$

where $\mathbb{X}^{-1}$ is the inverse of the matrix $\mathbb{X}$. This last equation states that $\mathbb{X}$ is an orthogonal matrix.

Equation (13) may be written in terms of this compound matrix as

$$\mathbb{X}^{T} m \mathbb{X} = \Delta \tag{17}$$

where

$$\Delta \equiv \lambda_{\beta} \delta_{\alpha\beta} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & & \\ 0 & \cdots & \cdots & \lambda_n \end{pmatrix} \tag{18}$$

Equation (17) indicates that the matrix M is transformed to diagonal form by means of the matrix of its eigen vectors. Moreover, from Equations (15), (16), and (17)

$$m = \mathbb{X} \Delta \mathbb{X}^{T} \tag{19}$$

Equation (19) expresses the symmetric matrix M in terms of its eigen values and eigen vectors.

The model matrix ($\mathbb{X}$) can thus be computed. The steps required in this process are:

1. the characteristic equation is found by expanding the determinant of $\delta ij - \lambda[I]$ and equating the result to zero
2. the roots to the characteristic equation are found ($\lambda_1, \ldots \lambda_6$)
3. the column vectors$\{\beta\}$are found by equating the result of the following

matrix multiplications to zero

there will be n vectors $\{\beta_{ij}\}$ because there are n values of λ_j and these vectors are linearly independent if the λ_j are distinct

4. construct the matrix $\mathbb{X}$ by ordering the column vectors

$$\mathbb{X} = \left[\{\beta_{ij}\} \cdots \{\beta_{ij}\}\right]$$

Consider now the quadratic form

$$q \equiv X^T m X \tag{20}$$

where X is an arbitrary column matrix of M elements and M is a symmetric matrix as previously discussed. If the scalar q is assigned a fixed value and if the vector X is considered to be a variable of M components, Equation (20) describes a "surface" in M-dimensional "space." Now, performing the coordinate transformation

$$X = \mathbb{X} y \tag{21}$$

transforms the quadratic form q to

$$q = y^T \mathbb{X}^T m \mathbb{X} y = y^T \Delta y \tag{22}$$

Thus, defining the M-vectors Y by

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \tag{23}$$

and performing the indicated matrix multiplications of Equation (22) yields q as

$$q = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 \tag{24}$$

Now, denoting the value assigned to q as K^2 requires that Y_1, Y_2, ..., y_n be chosen to satisfy the equation

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 = K^2 \tag{25}$$

For the case in which $Y_1 = Y_2 = \cdots Y_{i-1} = Y_{i+1} = y_n = 0$ or

$$y_i^2 = \frac{K}{\lambda_i} \tag{26}$$

$$y_i = \pm \frac{K}{\sqrt{\lambda_i}}$$

Hence, $K/\sqrt{\lambda_i}$ is the ith semi-principal axis of the surface represented by Equation (25).

Now, getting back to the task at hand, since the correlations in the covariance matrix can be eliminated with a coordinate transformation, the exponent of the density function contains the term

$$\sum_{i=1}^{n} z_i^2 \quad \text{where} \quad z_i = y_i/\lambda_i$$

But, this is the form for the square of the radius of an n-dimensional hypersphere (if considering only the cases for thich the summation is constant). Thus, a further transformation suggests itself in the evaluation of the probability that a given random sample from the statistically described distributions will fall within the specified radius. This transformation results from the fact that

$$P = \int_{-K}^{K} \cdots \int_{-f(z_i)}^{f(z_i)} \left(\frac{1}{2\pi}\right)^{n/2} e^{-\frac{1}{2}\sum z_i^2} \prod_{i=1}^{n} dz_i \tag{27}$$

(This equation is the integral of f(x) dx after transforming coordinates) is, in fact, the integral over the volume of hypersphere. This being the case, the integral can be written as

$$P = \int_{0}^{k} \left(\frac{1}{2\pi}\right)^{n/2} e^{-\frac{1}{2}r^2} f(r)dr \tag{28}$$

where f(r)dr is the spherically symmetric volume element of n-dimensions

$$V = \frac{\pi^{n/2}\, r^n}{\Gamma\left(\frac{n+2}{2}\right)} = \text{volume of hypersphere}$$

Γ = gamma function

n = dimensionality of hypersphere

n = 1	f(r)dr	=	2dr	(29a)
= 2		=	$2\pi r\, dr$	(29b)
= 3		=	$4\pi r^2\, dr$	(29c)
= 4		=	$2\pi^2 r^3\, dr$	(29d)
= 5		=	$\frac{8}{3}\pi^2 r^4\, dr$	(29e)
= 6		=	$\pi^3 r^5\, dr$	(29f)

Thus, the probabilities that the resultant error will be within a given "distance" of the center of the hypersphere can be computed as follows:

$$
\begin{aligned}
n=1 \quad P &= 2\left[\operatorname{erf}(K) - .5\right] \\
=2 \quad &= 1 - e^{-K^2/2} \\
=3 \quad &= 2\left[\operatorname{erf}(K) - .5\right] - Ke^{-K^2/2}\left[\sqrt{\tfrac{2}{\pi}}\right] \\
=4 \quad &= 1 - e^{-K^2/2}\left(1 + K^2/2\right) \\
=5 \quad &= 2\left[\operatorname{erf}(K) - .5\right] - Ke^{-K^2/2}\left[\sqrt{\tfrac{2}{\pi}} + \sqrt{\tfrac{2}{\pi}}\,\tfrac{K^2}{3}\right] \\
=6 \quad &= 1 - e^{-K^2/2}\left(1 + K^2/2 + K^4/8\right)
\end{aligned}
$$

where erf(K) denotes the error function.

Numerical data are tabulated for these six cases as a function of the radius, K, in Table C.1.

TABLE C.1 PROBABILITY OF INCLUSION AS A FUNCTION OF K AND N

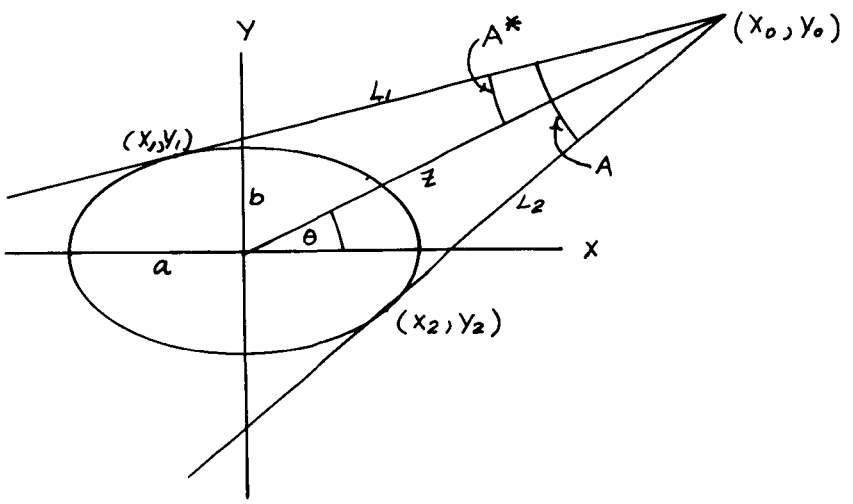
PROBABILTY						
K	N = 1	N = 2	N = 3	N = 4	N = 5	N = 6
= 0	0	0	0	0	0	0
.2	.1586	.0198	.0021	.0002	.0000	.0000
.4	.3108	.0769	.0162	.0030	.0005	.0001
.6	.4515	.1647	.0515	.0144	.0035	.0008
.8	.5762	.2739	.1127	.0415	.0138	.0043
= 1.0	6826	.3935	.1987	.0902	.0374	.0144
.2	.7698	.5133	.3038	.1628	.0801	.0366
.4	.8384	.6247	.4192	.2569	.1452	.0767
.6	.8904	.7220	.5355	.3661	.2326	.1383
.8	.9282	.8021	.6440	.4815	.3370	.2218
= 2.0	.9544	.8647	.7384	.5940	.4504	.3233
.2	.9722	.9111	.8161	.6959	.5643	.4355
.4	.9836	.9314	.8523	.7477	.6002	.4634
.6	.9906	.9676	.9234	.8581	.7720	.6731
.8	.9984	.9802	.9506	.9085	.8351	.7567
= 3.0	.9974	.9889	.9708	.9389	.8911	.8265
.2	.9986	.9940	.9833	.9633	.9310	.8847
.4	.9994	.9969	.9910	.9790	.9587	.9274
.6	.9998	.9985	.9954	.9886	.9765	.9727
.8	.9999	.9993	.9977	.9940	.9870	.9844
= 4.0	.99995	.9997	.9988	.9969	.9930	.9915
.2	.99998	.9999	.9995	.9985	.9965	.9956
.4	.99999	.9999	.9998	.9993	.9985	.9980

APPENDIX D

FLATTENED PLANET MEASUREMENT

Section 2.5.2.1.1 discussed the measurement of the angle subtended by a planet. The analysis was based on a perfectly spherical model of a planet. The following is a derivation of the measurement of the angle subtended by a flattened planet. The subtended angle is expressed as a function of the equatorial and polar radii, the relative attitude of the planet with respect to the observer, and the distance from the planet center to the observer. Although Section 2.5.2.1.1 presented the variation in the measured angle due to distance and diameter uncertainties, it must be remembered that these analyses were only for a spherical planet, though in most cases the results are acceptable. If an "exact" analysis is desired, the results of this appendix can be extended to find the variations in the angle measurement due to uncertainties in the distance and the planet radii. The general expression for the subtended angle will be found, and its variation with respect to the relative attitude will be derived. Other variations can be found in a completely analogous manner.

The geometry of the analysis is shown below:



The previous sketch shows the intersection of the plane of the measurement and the ellipsoid model of the flattened planet. Since the exact shape of the planet is not known and the most accurately known parameters of the size are the equatorial and polar radii, the best results are obtained if the measurement plane is selected to be that which is determined by either of the two radii. Thus, the planar case may by analyzed for purposes of determing flattening effects on the measurement.

If the planar analysis is pursued, the slopes of L_1 and L_2 can be found from the negative reciprocal slope of the gradient at $P(X_1, X_2)$ and $P(X_2, Y_2)$. Since

$$f(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \quad ,$$

then

$$f = \frac{2x}{a^2}\hat{a}_x + \frac{2y}{b^2}\hat{a}_y$$

so the slope of $L_1 = -\frac{b^2}{a^2}\frac{x_1}{y_1}$ and the slope of $L_2 = -\frac{b^2}{a^2}\frac{x_2}{y_2}$

The equations of L_1 and L_2 can now be written as

$$\frac{x_1 z\cos\theta}{a^2} + \frac{y_1 z\sin\theta}{b^2} = 1$$

$$\frac{x_2 z\cos\theta}{a^2} + \frac{y_2 z\sin\theta}{b^2} = 1$$

These equations must be combined with the restricting equation of the ellipse. A quadratic equation results, the solutions of which are

$$x_1 = \frac{\frac{b^2\cos\theta}{z} - a\sqrt{\sin^4\theta + \frac{b^2}{a^2}\cos^2\theta\sin^2\theta - \frac{b^2}{z^2}\sin^2\theta}}{\sin^2\theta + \frac{b^2}{a^2}\cos^2\theta}$$

$$y_1 = \frac{\frac{a^2\sin\theta}{z} + b\sqrt{\cos^4\theta + \frac{a^2}{b^2}\sin^2\theta\cos^2\theta - \frac{a^2}{z^2}\cos\theta}}{\frac{a^2}{b^2}\sin^2\theta + \cos^2\theta}$$

$$x_2 = \frac{\frac{b^2 \cos\theta}{z} + a\sqrt{\sin^4\theta + \frac{b^2}{a^2}\cos^2\theta\sin^2\theta - \frac{b^2}{z^2}\sin^2\theta}}{\sin^2\theta + \frac{b^2}{a^2}\cos^2\theta}$$

$$y_2 = \frac{\frac{a^2 \sin\theta}{z} - b\sqrt{\cos^4\theta + \frac{a^2}{b^2}\sin^2\theta\cos^2\theta - \frac{a^2}{z^2}\cos^2\theta}}{\frac{a^2}{b^2}\sin^2\theta + \cos^2\theta}$$

The slopes of L_1 and L_2 can now be written explicitly in terms of a, b, z and θ.

$$\text{Slope of } L_1 = -\left[\frac{b^2\cos\theta - za\sin\theta\sqrt{\sin^2\theta + \frac{b^2}{a^2}\cos^2\theta - \frac{b^2}{z^2}}}{a^2\sin\theta + zb\cos\theta\sqrt{\cos^2\theta + \frac{a^2}{b^2}\sin^2\theta - \frac{a^2}{z^2}}}\right]$$

$$\text{Slope of } L_2 = -\left[\frac{b^2\cos\theta + za\sin\theta\sqrt{\sin^2\theta + \frac{b^2}{a^2}\cos^2\theta - \frac{b^2}{z^2}}}{a^2\sin\theta - zb\cos\theta\sqrt{\cos^2\theta + \frac{a^2}{b^2}\sin^2\theta - \frac{a^2}{z^2}}}\right]$$

An exact solution of A could be found by taking the difference in the arctangents of the above slopes. Since the partial derivative of A with respect to θ is desired, the algebra will be significantly simplified if an approximation to b^2/a^2 is made. Since the flattening of the planets is not

very severe (particularly of the earth), the following is a good approximation

$$\frac{b^2}{a^2} = 1 - 2f - f^2 \approx 1$$

where f is the flattening ($f \approx 1/298$ for the earth). Now the slopes of L_1 and L_2 can be written as

$$\text{Slope of } L_1 = -\left[\frac{b^2 \cos\theta - a \sin\theta \sqrt{z^2 - b^2}}{a^2 \sin\theta + b \cos\theta \sqrt{z^2 - a^2}}\right] = U_1$$

$$\text{Slope of } L_2 = -\left[\frac{b^2 \cos\theta + a \sin\theta \sqrt{z^2 - b^2}}{a^2 \sin\theta - b \cos\theta \sqrt{z^2 - a^2}}\right] = U_2$$

and A can be written as $A = \arctan(U_2) - \arctan(U_1)$. At this point the uncertainty in A can be found by taking partial derivatives with respect to the uncertain parameters; i.e., $\frac{\partial A}{\partial z}$, $\frac{\partial A}{\partial a}$, $\frac{\partial A}{\partial b}$ and $\frac{\partial A}{\partial \theta}$. The first three quantities have been derived for the spherical planet in Section 2.5.2.1.1. For this reason, they will not be pursued here. Instead, $\frac{\partial A}{\partial \theta}$ will be determined. If more accurate results are desired for $\frac{\partial A}{\partial z}$, $\frac{\partial A}{\partial a}$, and $\frac{\partial A}{\partial b}$, the method is exactly analogous.

The chain rule is employed in order to determine $\frac{\partial A}{\partial \theta}$.

$$\frac{\partial A}{\partial \theta} = \frac{1}{1+U_2^2}\frac{\partial U}{\partial \theta} - \frac{1}{1+U_1^2}\frac{\partial U_1}{\partial \theta}$$

The result is

$$\frac{\partial A}{\partial \theta} = -\left[\frac{(a^2\sin\theta - b\cos\theta\sqrt{z^2-a^2})(-b^2\sin\theta + a\cos\theta\sqrt{z^2-b^2}) - (b^2\cos\theta + a\sin\theta\sqrt{z^2-b^2})(a^2\cos\theta + b\sin\theta\sqrt{z^2-a^2})}{(a^2\sin\theta - b\cos\theta\sqrt{z^2-a^2})^2 + (b^2\cos\theta + a\sin\theta\sqrt{z^2-b^2})^2}\right]$$

$$+\left[\frac{(a^2\sin\theta + b\cos\theta\sqrt{z^2-a^2})(-b^2\sin\theta - a\cos\theta\sqrt{z^2-b^2}) - (b^2\cos\theta - a\sin\theta\sqrt{z^2-b^2})(a^2\cos\theta - b\sin\theta\sqrt{z^2-a^2})}{(a^2\sin\theta + b\cos\theta\sqrt{z^2-a^2})^2 + (b^2\cos\theta - a\sin\theta\sqrt{z^2-b^2})^2}\right]$$

The preceding equation gives the uncertainty in the measurement of the angle subtended by a flattened planet due to the uncertainty in the relative attitude uncertainty of the observer. It is noted that the analysis simplifies if the planet is assumed to be an oblate spheroid. In this case, the flattening effect can be neglected if the choice of the measurement plane is the one determined by the equatorial radius. This can be seen if the variations in the measurement angle due to planet attitude are examined.

Due to the fact that part of the planet that would normally be used for an angle measurement may not be experiencing sunlight, some modification must be made to the conventional measurement. Unless suitable sensors that can detect the dark horizon are used, the measurement will resort to some estimation of the center of the planet, and the angle to be measured becomes the angle between one line of tangency and the estimated center point of the planet. Of course, the estimation of the center point introduces still another inaccuracy into the measurement. The expression for the measurement angle becomes

$$A^* = \theta - \arctan(U_1)$$

The variation of A* due to planet attitude uncertainty becomes

$$\frac{\partial A^*}{\partial \theta} = 1 - \frac{1}{1+U_1^2}\frac{\partial U_1}{\partial \theta}$$

This must be added to the uncertainty of the estimated center point of the planet. The total model uncertainty for the measurement of the angle between the center point and the planet edge becomes

$$\Delta A^* = \frac{\partial A^*}{\partial z}\Delta z + \frac{\partial A^*}{\partial a}\Delta a + \frac{\partial A^*}{\partial b}\Delta b + \frac{\partial A^*}{\partial \theta}\Delta\theta + \Delta\theta_{c.p.}$$

where the first three terms may be determined from the results of the spherical planet in Section 2.5.2.1.1 (using $A^* = A/2$), and $\Delta\theta_{c.p.}$ is the uncertainty angle of the centerpoint of the planet.

A similar expression results for the measurement of the total angle subtended by the planet:

$$\Delta A = \frac{\partial A}{\partial z}\Delta z + \frac{\partial A}{\partial a}\Delta a + \frac{\partial A}{\partial b}\Delta b + \frac{\partial A}{\partial \theta}\Delta\theta$$

The uncertainty in planet center does not enter into this expression since it is not used.