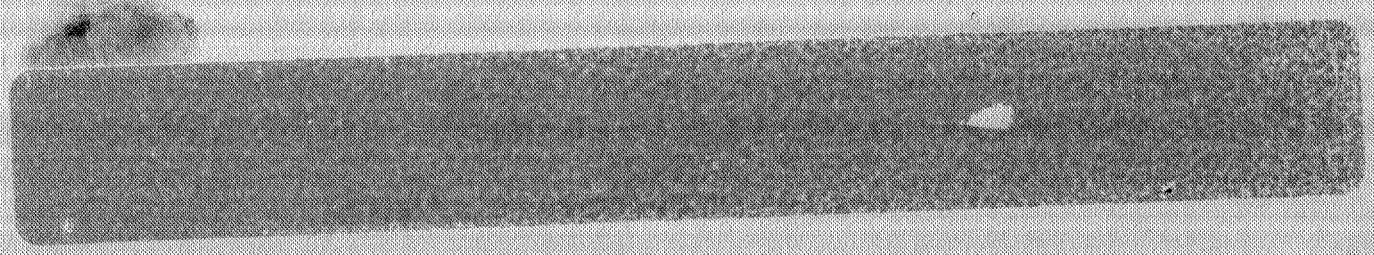
NASA Conference Publication 3102

Flight Mechanics/Estimation Theory Symposium 1990



Proceedings of a symposium held at
Goddard Space Flight Center
Greenbelt, Maryland
May 22–24, 1990

NASA

NASA Conference Publication 3102

Flight Mechanics/Estimation Theory Symposium 1990

Thomas Stengle, *Editor*
Goddard Space Flight Center
Greenbelt, Maryland

Proceedings of a symposium sponsored by
NASA Goddard Space Flight Center
and held at Goddard Space Flight Center
Greenbelt, Maryland
May 22–24, 1990

NASA
National Aeronautics and
Space Administration
Office of Management
Scientific and Technical
Information Division
1991

FOREWORD

The papers presented here have been derived primarily from speakers' summaries of talks presented at the Flight Mechanics/Estimation Theory Symposium held May 22-24, 1990 at Goddard Space Flight Center. For completeness, abstracts are included of those talks for which summaries were unavailable at press time. Papaers included in this document are presented as received from the authors with little or no editing.

CONTENTS

SESSION 3

SESSION 4

SESSION 5

FLIGHT MECHANICS/ESTIMATION THEORY SYMPOSIUM

MAY 22-24, 1990

SESSION 1

COMPARISON OF TWO ON-ORBIT ATTITUDE SENSOR ALIGNMENT METHODS

Kenneth Krack and Michael Lambertson
Computer Sciences Corporation
10110 Aerospace Road
Lanham-Seabrook, MD 20706

and

F. Landis Markley
National Aeronautics and Space Administration (NASA)
Goddard Space Flight Center
Greenbelt, MD 20771

Abstract

This paper compares two methods of on-orbit alignment of vector attitude sensors. The first method uses the angular difference between simultaneous measurements from two or more sensors. These angles are compared to the angular differences between the respective reference positions of the sensed objects. The alignments of the sensors are adjusted to minimize the difference between the two sets of angles. In the second method, the sensor alignment is part of a state vector that includes the attitude. The alignments are adjusted along with the attitude to minimize all observation residuals. It is shown that the latter method can result in much less alignment uncertainty when gyroscopes are used for attitude propagation during the alignment estimation. The additional information for this increased accuracy comes from knowledge of relative attitude obtained from the spacecraft gyroscopes. This paper presents the theoretical calculations of this difference in accuracy. Also presented are numerical estimates of the alignment uncertainties of the fixed-head star trackers on the Extreme Ultraviolet Explorer spacecraft using both methods.

1. INTRODUCTION

This paper compares two methods of determining the in-flight alignment estimation of vector-type attitude sensors. The two methods of alignment estimation will be referred to as attitude-independent and attitude-dependent and are outlined below. The estimated accuracy of the two approaches will be discussed.

Vector-type attitude sensors are those sensors whose output are vector measurements of the lines-of-sight to some reference object. Examples of this type of sensor are Sun sensors and star trackers. A minimum of two vector observations are needed for attitude determination. The analysis presented in this paper will be restricted to observations from two fixed-head star trackers (FHSTs) for simplicity.

Section 2 discusses the theoretical differences in the accuracy of these two approaches. Section 3 presents a numerical example of both methods using the Extreme Ultraviolet Explorer (EUVE) spacecraft as a typical mission case.

ATTITUDE-INDEPENDENT ALIGNMENT ESTIMATION

The attitude-independent method of sensor alignment estimation discussed here is based on an algorithm first presented in a paper by Shuster, Chitre, and Niebur (Reference 1) and later refined by Bierman and Shuster (Reference 2). This method of sensor alignment estimation uses the angle between two vector observations as its basic observation. This scalar observation is compared to the angle between the corresponding reference objects, and the sensor alignment is adjusted to minimize the difference. The method requires simultaneous measurements in each sensor. No a priori knowledge of the attitude is used nor does the algorithm solve for the attitude. The algorithm seeks to minimize the overall deviation of the sensor alignments from their prelaunch values.

ATTITUDE-DEPENDENT ALIGNMENT ESTIMATION

In an attitude-dependent alignment estimation method, the spacecraft attitude is part of a state vector that includes the sensor alignments. The attitude is either solved-for along with the alignment or is treated as a known quantity. The primary observation quantity is the unit vector measurement from the sensors as opposed to the scalar measurement of the attitude-independent

method. Residuals are computed between estimated observations based on the state vector and the actual observations. The state vector is adjusted to minimize the residuals. Two procedures are in common use to obtain this minimization: the sequential filter and the batch filter. The sequential filter updates its estimation of the state vector at discrete times using a previously estimated state that has been propagated from the time of the last sensor observation. The batch filter considers a collection of sensor observations in toto and seeks to minimize all residuals simultaneously. Both procedures require knowledge of the motion of the spacecraft between observations. This knowledge is usually obtained from gyroscope measurements but could be inferred from a dynamics model of the spacecraft. The analysis in this paper will restrict itself to the batch filter and will assume dynamics information is available.

2. THEORETICAL EVALUATION

As stated in the introduction, this analysis will assume a spacecraft with two FHSTs. The state vector (References 3 through 5) to be estimated is

$$\bar{x}(t) = \begin{bmatrix} \Delta\vec{\theta}(t) \\ \vec{m}_1 \\ \vec{m}_2 \end{bmatrix}$$

where $\Delta\vec{\theta}(t)$ is a vector of attitude error angles around the spacecraft body axes and $\vec{m}_i$ (for i = 1 and 2) is a vector of sensor misalignment angles around the i^{th} FHST axes (References 4 and 5).

A batch least-squares estimate of the state at an epoch time t_o, ignoring the effects of dynamics noise and consider parameters, has the covariance (Reference 3)

$$P(t_o) \equiv E\left[\bar{x}(t_o)\, \bar{x}^T(t_o) \right] = W_n^{-1}$$

where W_n is the normal matrix,

$$W_n = W_o + F^T W F$$

The matrix W_o is the inverse of the a priori covariance of $\bar{x}$ and $F^T W F$ represents the information contained in the measurements. Assuming no a priori attitude knowledge and a priori knowledge of sensor misalignments with standard deviation σ_a per axis gives

$$W_o = \sigma_a^{-2} \left[\begin{array}{c|c} 0_{3x3} & 0_{3x6} \\ \hline 0_{3x6} & I_{6x6} \end{array} \right]$$

where 0_{nxm} denotes an n x m matrix of zeros and I_{nxn} is an n x n identity matrix. The form of the matrices F and W depends on the measurements.

The i^{th} FHST returns the two-component measurement

$$\bar{Z}_i = -\begin{bmatrix} W_{ix} / W_{iz} \\ W_{iy} / W_{iz} \end{bmatrix}$$

where $\hat{W}_i$ is the unit vector measurement of the star in FHST coordinates (References 4 and 5). The errors in the two components of $\bar{Z}_i$ are assumed to be uncorrelated and to have equal measurement variances denoted by σ_m^2.

The observation at a given time processed by the estimator is a function of $\bar{Z}_1$ and $\bar{Z}_2$. The attitude-independent observation is a scalar given by the inner product of the two star vectors in the spacecraft body frame.

$$s = \hat{W}_1^T B_1 B_2^T \hat{W}_2$$

where B_1 and B_2 are the sensor alignment matrices that rotate a vector from the spacecraft body frame to FHST-1 and FHST-2 frames, respectively. This quantity contains no attitude information.

The attitude-dependent observation is a four-component vector

$$\bar{y} = \begin{bmatrix} \bar{Z}_1 \\ \bar{Z}_2 \end{bmatrix}$$

This quantity contains attitude information. For simplicity, assume the sensor alignments, B_1 and B_2, are given by

$$B_1 = I_{3x3} = 3 \times 3 \text{ identity matrix}$$

$$B_2 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \text{rotation by 90 degrees (deg) about the y-axis}$$

Thus, the sensor boresight, which is along the z-axis in the sensor frame, is along the spacecraft z- and x-axis for FHST-1 and FHST-2, respectively. The partial derivatives of the observations with respect to the state vector components are needed to compute the normal matrix. Again for simplicity, assume the observed stars are on the sensor boresights. The partial derivative matrix for the scalar observation is

$$G_s \equiv \frac{\partial s}{\partial \bar{x}} = \left[0_{1x3},\ (\hat{W}_1 \times B_1 B_2^T \hat{W}_2)^T,\ (\hat{W}_2 \times B_2 B_1^T \hat{W}_1)^T \right]$$

$$= [\ 0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 0,\ -1,\ 0\ ]$$

For the assumed geometry, the variance of the errors in the scalar observation, s, is $2\sigma_m^2$. The partial derivative matrix for the vector observation is (References 4 and 5)

$$G_v \equiv \frac{\partial \bar{y}}{\partial \bar{x}} = \left[\begin{array}{c|c|c} M & M & 0_{2x3} \\ \hline MB_2 & 0_{2x3} & M \end{array} \right]$$

where

$$M = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

The attitude and misalignment variances will be computed for four cases. For either a scalar or vector observation type, either a single measurement or else two measurements separated by a 90-deg attitude maneuver about the space-

craft y-axis will be considered. The estimate epoch time will be taken to be the time of the first observation. The measurement weight matrix, W, will be assumed to be the inverse of the measurement covariance in all cases. The measurement error propagation matrix between the two observation times is given by (References 3 through 5)

$$\Phi = \left[\begin{array}{c|c} \Phi_\theta & 0_{3x6} \\ \hline 0_{6x3} & I_{6x6} \end{array}\right]$$

where Φ_θ represents a 90-deg rotation about the y-axis. The four measurement scenarios and their associated weight and total derivative matrices follow:

1. A single scalar observation

$$F = G_s$$
$$W = \frac{1}{2}\sigma_m^{-2}$$

2. Two scalar observations separated by a 90-deg maneuver

$$F = \begin{bmatrix} G_s \\ G_s \Phi \end{bmatrix}$$
$$W = \frac{1}{2}\sigma_m^{-2} I_{2x2}$$

3. A single four-vector observation

$$F = G_v$$
$$W = \sigma_m^{-2} I_{4x4}$$

4. Two four-vector observations separated by a 90-deg maneuver

$$F = \begin{bmatrix} G_v \\ G_v \Phi \end{bmatrix}$$
$$W = \sigma_m^{-2} I_{8x8}$$

For each scenario the normal matrix will be computed and then inverted to obtain the covariance matrix. In all cases a permutation of the rows and

columns of the normal matrix will bring it to block diagonal form, facilitating matrix inversion. This permutation corresponds to rearranging the state vector as

$$\bar{x}' \equiv \left[\Delta\theta_x,\ m_{1x},\ \Delta\theta_z,\ m_{2x},\ \Delta\theta_y,\ m_{1y},\ m_{2y},\ m_{1z},\ m_{2z} \right]^T$$

The normal matrix with the permuted rows and columns will similarly be distinguished by a prime, W'_n.

SCENARIO 1: A SINGLE SCALAR OBSERVATION

First consider scenario 1, a single scalar observation, for which

$$F^T W\ F = \frac{1}{2}\,\sigma_m^{-2}\ G_s^T G_s = \frac{1}{2}\,\sigma_m^{-2} \left[\begin{array}{c|c|c} 0_{3x3} & 0_{3x3} & 0_{3x3} \\ \hline 0_{3x3} & D & -D \\ \hline 0_{3x3} & -D & D \end{array}\right]$$

where D = diag(0, 1, 0) and diag(...) denotes a diagonal matrix with the indicated elements along the diagonal. The permuted normal matrix in this scenario is

$$W'_n = \sigma_m^{-2} \left[\begin{array}{ccccccccc} 0 & & & & & & & & \\ & \alpha & & & & & & & \\ & & 0 & & & & & & \\ & & & \alpha & & & & & \\ & & & & 0 & & & & \\ & & & & & (\alpha + 1/2) & -1/2 & & \\ & & & & & -1/2 & (\alpha + 1/2) & & \\ & & & & & & & \alpha & \\ & & & & & & & & \alpha \end{array}\right]$$

where

$$\alpha = (\sigma_m\ /\ \sigma_a)^2$$

The variances of the state vector components are the diagonal elements of the inverse of W'_n. Thus, in this scenario the attitude error variances are all infinite, which is to be expected with attitude-independent observations. The four misalignment variances, $\sigma^2(m_{1x})$, $\sigma^2(m_{1z})$, $\sigma^2(m_{2x})$, and $\sigma^2(m_{2z})$, are all equal to their a priori values, σ_a^2, which means that the single scalar observation does not contain information about these misalignments. The y component of the misalignment variances for each of the two sensors is given by

$$\sigma^2(m_{1y}) = \sigma_a^2 \left(\frac{\sigma_m^2 + 1/2\ \sigma_a^2}{\sigma_m^2 + \sigma_a^2} \right)$$

It is instructive to examine two limiting cases: the case of sensor data uncertainty much less than the a priori state vector uncertainty ($\sigma_m \ll \sigma_a$) and the opposite case of poor measurement accuracy ($\sigma_m \gg \sigma_a$). The y-axis alignment variance for scenario 1 has the limit σ_a^2 for $\sigma_m \gg \sigma_a$ and $(1/2)\ \sigma_a^2$ for $\sigma_m \ll \sigma_a$. The former result is reasonable since the measurements do not improve the a priori estimates in this limit. In the opposite limit of accurate measurements, the y-axis misalignments of the two FHSTs are in some sense averaged. These results are independent of the number of observations.

SCENARIO 2: TWO SCALAR OBSERVATIONS SEPARATED BY A MANEUVER

For scenario 2, two scalar observations separated by a 90-deg maneuver,

$$F^T W\ F = \frac{1}{2}\ \sigma_m^{-2} \left(G_s^T G_s + \Phi^T G_s^T G_s \Phi \right) = \sigma_m^{-2}\ G_s^T G_s$$

since it is easily seen that

$$\Phi^T G_s^T G_s \Phi = G_s^T G_s$$

Thus, the only change in W'_n from scenario 1 is to replace each 1/2 in the matrix by 1, and the only variances that are modified are

$$\sigma^2(m_{1y}) = \sigma_a^2 \left(\frac{\sigma_m^2 + \sigma_a^2}{\sigma_m^2 + 2\ \sigma_a^2} \right)$$

These variances, however, have the same limits for $\sigma_m \gg \sigma_a$ and $\sigma_m \ll \sigma_a$ as in scenario 1.

SCENARIO 3: A SINGLE FOUR-VECTOR OBSERVATION

Scenario 3, a single four-vector observation, gives

$$F^T W F = \sigma_m^{-2} G_v^T G_v = \sigma_m^{-2} \left[\begin{array}{c|c} \Gamma_{\theta\theta} & \Gamma_{\theta m} \\ \hline \Gamma_{\theta m}^T & \Gamma_{mm} \end{array}\right]$$

where

$$\Gamma_{\theta\theta} = \mathrm{diag}(1,\ 2,\ 1)$$

$$\Gamma_{\theta m} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

$$\Gamma_{mm} = \mathrm{diag}(1,\ 1,\ 0,\ 1,\ 1,\ 0)$$

The permuted normal matrix is

$$W'_n = \sigma_m^{-2} \left[\begin{array}{cc|cc|ccc|c} 1 & 1 & & & & & & \\ 1 & (\alpha+1) & & & & 0 & & \\ \hline & & 1 & -1 & & & & \\ & & -1 & (\alpha+1) & & & & \\ \hline & & & & 2 & 1 & 1 & \\ & & & & 1 & (\alpha+1) & 0 & \\ & & & & 1 & 0 & (\alpha+1) & \\ \hline & & 0 & & & & & \alpha \\ & & & & & & & & \alpha \end{array}\right]$$

Inverting this matrix gives the attitude variances

$$\sigma^2(\Delta\theta_x) = \sigma^2(\Delta\theta_z) = \sigma_a^2 + \sigma_m^2$$

$$\sigma^2(\Delta\theta_y) = \frac{1}{2}\left(\sigma_a^2 + \sigma_m^2\right)$$

These results can be understood intuitively as follows. An observation of a star on the boresight of FHST-1, which is aligned with the spacecraft z-axis, provides information about the x- and y-axis attitude errors. Similarly, an observation of a star on the boresight of FHST-2, which is aligned with the spacecraft x-axis, provides information about the y- and z-axis attitude errors. As there is twice as much information about the y-axis attitude, the variance of the y-axis attitude is half as large as that of the other two axes. The x- and z-axis uncertainties are the root-sum-square of the a priori alignment uncertainty and the measurement uncertainty.

The alignment variances are the same as in scenario 1; therefore, no alignment information is lost or gained by combining the four components of the vector observations into a single scalar observation for measurements taken at a single attitude.

SCENARIO 4: TWO FOUR-VECTOR OBSERVATIONS SEPARATED BY A MANEUVER

The final scenario, two four-vector observations separated by a 90-deg maneuver, has

$$F^T W F = \sigma_m^{-2}\left(G_v^T G_v + \Phi^T G_v^T G_v \Phi\right)$$

where

$$\Phi^T G_v^T G_v \Phi = \left[\begin{array}{c|c} \Phi_\theta^T \Gamma_{\theta\theta} \Phi_\theta & \Phi_\theta^T \Gamma_{\theta m} \\ \hline \Gamma_{\theta m}^T \Phi_\theta & \Gamma_{mm} \end{array}\right]$$

It is easy to see that

$$\Phi_\theta^T \Gamma_{\theta\theta} \Phi_\theta = \Gamma_{\theta\theta}$$

however,

$$\Phi_\theta^T \Gamma_{\theta m} = \begin{bmatrix} 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \neq \Gamma_{\theta m}$$

Thus, in contrast to the scalar observation case, the structure of the normal matrix will be different when a spacecraft maneuver is performed during alignment. The normal matrix is

$$W_n' = \sigma_m^{-2} \begin{bmatrix} \begin{matrix} 2 & 1 & 0 & -1 \\ 1 & (\alpha+2) & -1 & 0 \\ 0 & -1 & 2 & -1 \\ -1 & 0 & -1 & (\alpha+2) \end{matrix} & \begin{matrix} \\ \\ 0 \\ \\ \end{matrix} \\ \begin{matrix} \\ \\ 0 \\ \\ \end{matrix} & \begin{matrix} 4 & 2 & 2 & \alpha & & \\ 2 & (\alpha+2) & 0 & & \alpha & \\ 2 & 0 & (\alpha+2) & & & \alpha \end{matrix} \end{bmatrix}$$

It can be seen that $\sigma^2(m_{iz}) = \sigma_a^2$ in this scenario as in all the others. This result is simply that observations of stars on the boresight can never improve knowledge of the misalignment component about the boresight. The y-axis attitude and misalignment variances are the same as scenario 3 with σ_m^2 being replaced by (1/2) σ_m^2, reflecting the presence of twice as many measurements. Thus

$$\sigma^2(\Delta\theta_y) = \frac{1}{2}\left(\sigma_a^2 + \frac{1}{2}\sigma_m^2\right)$$

The y-axis misalignment uncertainties are the same as those in scenario 2, using scalar observations formed from the same measurements. The new feature of scenario 4 is the coupling of $\Delta\theta_x$, $\Delta\theta_z$, m_{1x}, and m_{2x} in the 4 x 4 block in the upper-left corner of W_n'. Inverting this submatrix gives the variances

$$\sigma^2(\Delta\theta_x) = \sigma^2(\Delta\theta_z) = \sigma_m^2 \left(\frac{\sigma_m^2 + 2\,\sigma_a^2}{2\,\sigma_m^2 + 2\,\sigma_a^2} \right)$$

$$\sigma^2(m_{ix}) = (\sigma_a^{-2} + \sigma_m^{-2})^{-1}$$

The x-axis misalignment variance of the FHST is the harmonic mean of the measurement variance and the a priori alignment variance. It is, therefore, less than either variance and approaches the smaller of the two in the limit that the other becomes very large. The limits of the x- and z-axis attitude variances are (1/2) σ_m^2 for $\sigma_m \gg \sigma_a$ and σ_m^2 for $\sigma_m \ll \sigma_a$. The latter is, like $\sigma^2(m_{ix})$, finite as σ_a tends to infinity. Thus, unlike any other scenario considered, the attitude and alignment knowledge obtained from the FHSTs is much better than the a priori knowledge of the misalignments.

THEORETICAL EVALUATION SUMMARY

Table 1 summarizes the alignment variance results for the four scenarios. As the table shows, the only significant difference in the alignment variance results between the attitude-independent and the attitude-dependent methods is due to the attitude maneuver.

Table 1. Alignment Variance Results Summary

	$\sigma^2(m_{ix})$	$\sigma^2(m_{iy})$	$\sigma^2(m_{iz})$
Scenario 1	σ_a^2	$\sigma_a^2 \left(\frac{\sigma_m^2 + 1/2\,\sigma_a^2}{\sigma_m^2 + \sigma_a^2} \right)$	σ_a^2
Scenario 2	σ_a^2	$\sigma_a^2 \left(\frac{\sigma_m^2 + \sigma_a^2}{\sigma_m^2 + 2\,\sigma_a^2} \right)$	σ_a^2
Scenario 3	σ_a^2	$\sigma_a^2 \left(\frac{\sigma_m^2 + 1/2\,\sigma_a^2}{\sigma_m^2 + \sigma_a^2} \right)$	σ_a^2
Scenario 4	$(\sigma_a^{-2} + \sigma_m^{-2})^{-1}$	$\sigma_a^2 \left(\frac{\sigma_m^2 + \sigma_a^2}{\sigma_m^2 + 2\,\sigma_a^2} \right)$	σ_a^2

3. A NUMERICAL EXAMPLE

This section presents a numerical example of the results obtained in Section 2. The spacecraft that will be used in this example is EUVE, an astronomical satellite that is scheduled to launch in August 1991. Its primary attitude sensors are two FHSTs and three-axis gyroscopes. During the first 6 months, EUVE will perform a full-sky ultraviolet survey and will be spinning at rate of three revolutions per orbit around the spacecraft x-axis. After this survey phase, the spacecraft will be held at constant attitudes to measure the spectrum of various targets. The mission profile of EUVE, therefore, provides the opportunity of FHST alignment calibration under both maneuvering and stationary conditions. Although the alignment and attitude motion of EUVE differs from the analytical model presented in Section 2, they are sufficiently similar to substantiate the analytical results.

NUMERICAL MODEL

The two EUVE FHSTs are NASA standard star trackers with an 8-by-8 deg field-of-view (FOV). Their alignment is given as a 3-2-3 Euler rotation sequence. The rotation angles are given in Table 2.

Table 2. Nominal FHST Alignments

Sensor	Rotation Angles (Degrees)		
	θ_1	θ_2	θ_3
FHST-1	59.90	105.60	0.0
FHST-2	128.10	105.60	0.0

The assumed a priori alignment uncertainties are 2.68×10^{-4} radians (55.3 arc-seconds) (3σ) for each FHST axis. The measurement noise assumed is 1.944×10^{-4} radians (40.1 arc-seconds) (3σ). In the attitude-dependent case, the a priori attitude uncertainties are assumed to be 1 deg (3σ) for each axis. The measurement rate for the EUVE FHSTs is one observation per 0.256 seconds. To simplify the interpretation of the results, the gyroscopes are assumed to be perfect.

The attitude-independent alignment accuracy estimate is made using a program that implements the equations for the alignment covariance given in References 1 and 2. The attitude-dependent alignment accuracy estimate is performed using the Attitude Determination Error Analysis System (ADEAS), as described in Reference 5. In both cases, a one-orbit simulation is run. Although the sensor coverage is not continuous because a realistic star catalog is used, the length of the simulation run and the data rate are sufficient to give several thousand observations.

RESULTS

The results for the four simulation runs corresponding to the four scenarios are given in Table 3. For simplicity, the arithmetic mean of the uncertainties for the two FHSTs are given. The actual differences are in all cases less than 1 percent. All values are three-sigma.

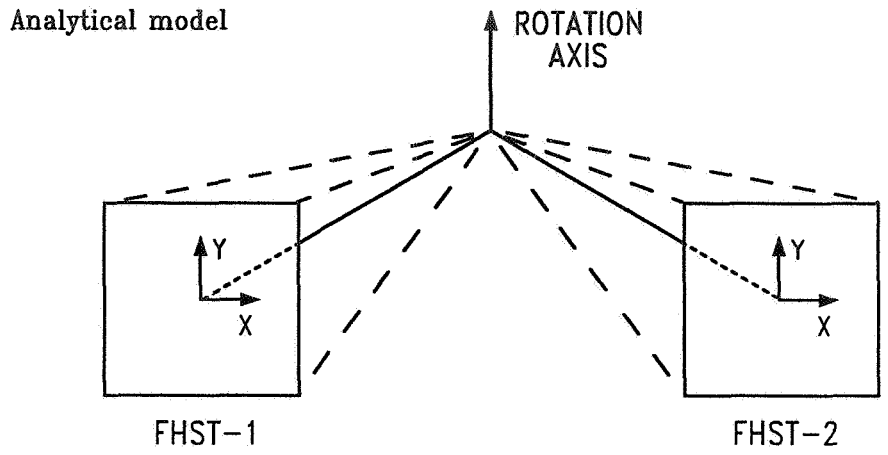
Table 3. EUVE Alignment Uncertainties

Axis	Attitude-Independent, Nonrotating (Scenario 1)	Attitude-Independent, Rotating (Scenario 2)	Attitude-Dependent, Nonrotating (Scenario 3)	Attitude-Dependent, Rotating (Scenario 4)
X	1.94×10^{-4} (40.0)	1.90×10^{-4} (39.2)	1.93×10^{-4} (39.8)	0.32×10^{-4} (6.52)
Y	2.65×10^{-4} (54.7)	2.03×10^{-4} (41.8)	2.66×10^{-4} (54.8)	1.50×10^{-4} (30.9)
Z	2.68×10^{-4} (55.3)	2.07×10^{-4} (42.6)	2.68×10^{-4} (55.3)	1.17×10^{-4} (24.1)

NOTE: Values are in radians (arc-seconds).

OBSERVATIONS

Overall, the results in Table 3 and Section 2 agree. The smallest uncertainties are in the case of using an attitude-dependent alignment method during a maneuver. The differences between the results presented in Table 3 and the results from Section 2 are mainly due to the different geometry of the two cases. Figure 1 shows the fundamental geometry of the analytical and numerical models.



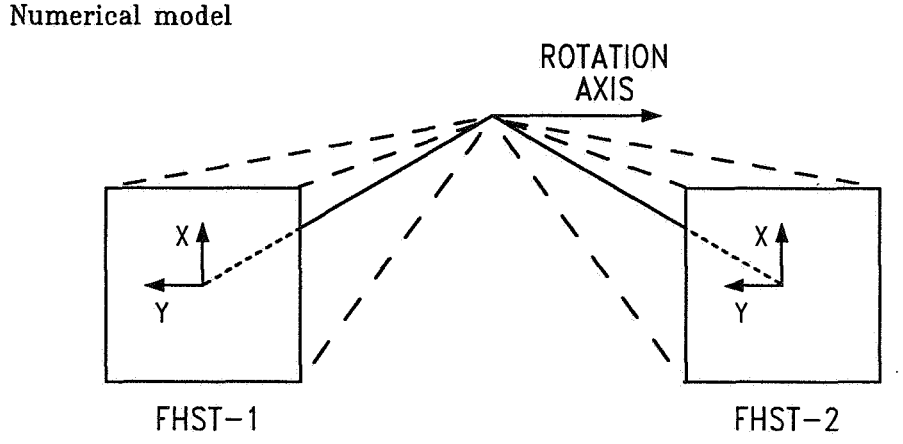


Figure 1. Geometry of the Analytical and Numerical Models

In the analytical model, a rotation about the FHST y-axis is equivalent to a change in the angle between the sensors. The corresponding axis in the numerical model is the x-axis. Table 1 indicates that the uncertainty of this axis should be 1.90×10^{-4} radians (39.1 arc-seconds). This result is in very good agreement with the first three cases. Allowing for the difference in sensor coordinate definition, the uncertainties of the y- and z-axes for the two nonrotating cases are also in good agreement with Table 1 expected results. The smaller uncertainties in the rotating cases are due to the finite FHST FOV size in the numerical example. In the nonrotating cases, the FHST observes the same star for the entire simulation. Under this condition, the y- and z-axes alignments are unobservable. This unobservability is also the consequence of assuming the star to be on the boresight as in the analytical evaluation. In the rotating numerical examples, however, the realistic star catalog used results in star measurements over the whole FOV and the alignments of the y- and z-axes are, therefore, observable. The variances of all three axes alignments in the attitude-independent case are larger than $(1/2)\ \sigma_a^2$.

4. CONCLUSION

The reason for the superior performance in the fourth measurement scenario of Section 2 is that the axis of the 90-deg attitude maneuver has been assumed to coincide exactly with the spacecraft y-axis. Since an actual estimator would obtain the angular rotation from a set of rate-integrating gyroscopes, this assumption is equivalent to the definition that the gyroscope axes are perfectly aligned with the spacecraft body axes. An overall rotation of all the sensors, including unspecified payload instruments, is indistinguishable from a spacecraft attitude rotation. Because of this perfect correlation, if the sensors and payload instruments can be coaligned, the freedom exists to arbitrarily choose one sensor to be perfectly aligned, which is equivalent to defining the spacecraft body axes in terms of this sensor's axes. The attitude-dependent algorithm, as modeled in Section 2, implicitly defines the spacecraft axes in relation to the gyroscope axes, so the improved alignments of the FHSTs in scenario 4 are actually owing to their alignment with respect to the rotation vector as measured by the gyroscopes.

An estimator that uses scalar observations as in Section 2 to align the attitude sensors does not include any attitude information. There is, therefore,

no measurement of the sensor alignment with respect to perfectly aligned gyroscopes. Because of the correlation between an overall sensor rotation and the attitude, there are three degrees of freedom in the sensor alignment that are unobservable. The resulting uncertainties in the alignment estimates, therefore, cannot be made arbitrarily small and will strongly depend on the prelaunch measurements. If the alignment uncertainty of one of the sensors is assumed to be zero, this sensor would become a reference sensor similar to the gyroscope in the attitude-dependent model. The alignment uncertainty of the other sensor should then be similar to the attitude-dependent results. To illustrate this effect, the scalar observation simulation software was executed assuming one of the sensors to be perfectly aligned. The results of this simulation are given in Table 4. Also given in Table 4 is the attitude-dependent, rotating case results from Section 3 for comparison. All values are three-sigma.

Table 4. EUVE Alignment Uncertainties Showing Reference Sensor Case Results

Axis	Attitude-Dependent, Rotating (Scenario 4)	Attitude-Independent, Rotating Reference Sensor
X	0.32×10^{-4} (6.52)	0.30×10^{-4} (6.19)
Y	1.50×10^{-4} (30.9)	1.32×10^{-4} (27.2)
Z	1.17×10^{-4} (24.1)	1.50×10^{-4} (31.0)

NOTE: Values are in radians (arc-seconds).

It can be seen that the two cases are in good agreement. This result would indicate that the distinguishing factor for producing small alignment uncertainties is not attitude-dependence or -independence but whether there is a reference sensor that serves to define an on-orbit spacecraft coordinate frame.

Whether, for a particular mission, an alignment algorithm that defines an on-orbit spacecraft frame or one that maintains the prelaunch frame should be

used depends on the type of payload. When the payload can align itself with respect to the attitude sensors, there is no need to maintain the prelaunch spacecraft reference. The ambiguity of the attitude reference can be removed by defining a new on-orbit reference, e. g., the gyroscope axes, and much improved attitude accuracies can result. However, when the payload is attitude sensitive but not sufficiently so to allow coalignment, an algorithm that preserves the prelaunch frame is preferable. In this case, the alignment of the payload is known relative only to the prelaunch reference. A method that does not make use of a reference sensor minimizes the deviation of the on-orbit sensor alignment from the prelaunch measurement and, therefore, will result in the best estimate of the payload attitude.

REFERENCES

1. Shuster, M. D., Chitre, D. M., and Niebur, D. P.,"In-Flight Estimation of Spacecraft Attitude Sensor Accuracies and Alignments," *Journal of Guidance and Control*, Vol. 5, No. 4, pp 339-343, 1982

2. Bierman, G. J. and Shuster, M. D., "Spacecraft Alignment Estimation," 27th IEEE Conference on Decision and Control, 1988

3. Markley, F. L., Seidewitz, E., and Nicholson, M., "A General Model for Attitude Determination Error Analysis," Flight Mechanics/Estimation Theory Symposium 1988, NASA Conference Publication 3011, 1988

4. Markley, F. L., Seidewitz, E., and Deutschmann, J., "Attitude Determination Error Analysis: General Model and Specific Application," presented at the CNES International Symposium on Space Dynamics, Toulouse, France, 1989

5. Nicholson, M., Markley, F. L., and Seidewitz, E., Attitude Determination Error Analysis System (ADEAS) Mathematical Specifications Document, Computer Sciences Corporation, CSC/TM-88/6001, October 1988

Attitude Sensor Alignment Calibration for the Solar Maximum Mission

Daniel S. Pitone

Computer Sciences Corporation

and

Malcolm D. Shuster

The Johns Hopkins University Applied Physics Laboratory

ABSTRACT

An earlier heuristic study of the fine attitude sensors for the Solar Maximum Mission (SMM) revealed a temperature dependence of the alignment about the yaw axis of the pair of fixed-head star trackers relative to the fine pointing Sun sensor. In the present work, new sensor alignment algorithms which better quantify the dependence of the alignments on the temperature are developed and applied to the SMM data. Comparison with the results from the previous study reveals the limitations of the heuristic approach. In addition, some of the basic assumptions made in the prelaunch analysis of the alignments of the SMM are examined. The results of this work have important consequences for future missions with stringent attitude requirements and where misalignment variations due to variations in the temperature will be significant.

1. INTRODUCTION

Because of the stringent attitude accuracy requirements, the temperature dependence of the alignments of the fine attitude sensors of the Solar Maximum Mission (SMM) has been closely studied.[1-5] These works have attempted to quantify the relationship between the variations in the SMM structural temperatures and the variations in the alignments of the fixed-head star trackers (FHSTs) relative to the fine pointing Sun sensors (FPSSs). The present work, which is an extension of Refs. 3 and 4, attempts to quantify more completely the relationship between the temperature and the alignments. This is done by using newly developed alignment estimation algorithms that can estimate the alignments better than the work in Refs. 3 and 4. In addition, the limiting assumptions made in the prelaunch analysis of the SMM with regard to the sensor alignments are examined.

One of the contributions of this work is a consistent framework for estimating the inflight alignments of spacecraft attitude sensors and investigating the nature of the changing alignments. An algorithm is provided for computing the alignments at a single temperature using inflight sensor measurements without the need to compute the spacecraft attitude and angular velocity. These alignment estimates at different temperatures are then input to a second algorithm which computes an optimal estimate for the temperature dependence. These methods can provide clues to the specific causes of the alignment changes and aid in the design of future structural configurations to minimize temperature effects.

This work begins by reviewing the heuristic analysis and its limitations as presented in Refs. 3 and 4. A brief outline of the new alignment algorithms is then given and simulations are provided to demonstrate their capabilities. Then the application of the algorithms to SMM data is presented and the dependence of the alignments on temperature is completely discussed. The last section discusses the assumptions made pertaining to the actual inflight SMM alignment estimation algorithms and possible modifications that would have made these algorithms more effective.

2. AN HEURISTIC EXAMINATION OF THE SMM ALIGNMENTS

SMM History and Configuration

The SMM was launched in February of 1980 from the Eastern Test Range into a low Earth orbit to study solar radiation at several wavelengths. In November of 1980 the reaction wheels that controlled the spacecraft failed. Thereafter, the spacecraft was put into a stabilizing spin mode to preserve the mission. During this time little scientific work was accomplished. In April of 1984 the mission was repaired in orbit by the Space Transportation System, after which it was returned to normal mission operation. The spacecraft functioned normally until December of 1989 when it reentered the Earth's atmosphere. Further details on the history of the SMM are given in Ref. 6.

The SMM was the first of the Multimission Modular Spacecraft (MMS) which were designed for repair and adaptation. It consisted basically of two separate components. The first component was the MMS component which consisted of the communications, power, and attitude control modules. The attitude control module included two FHSTs and a complete set of gyros which were used for fine attitude determination. The FHSTs were mounted together on a rigid structure inside the attitude control module to minimize alignment variations.

The second main component of the SMM was the payload component. It consisted of the SMM scientific payload and the SMM-specific attitude sensors. The scientific payload consisted of the instruments used to study the Sun. The SMM-specific attitude sensors were a redundant set of FPSSs and a set of coarse Sun sensors. The scientific instruments and the FPSSs were comounted on a rigid plate to minimize misalignment between them. The basic configuration of the SMM is shown in Fig. 1.

SMM Attitude Determination Configuration

The goal of the SMM attitude determination and control system was to point the boresights of the scientific instruments as accurately as possible, nominally to within 5 arc-sec, at specific locations on the Sun. To this end, the scientific instruments and the FPSSs were mounted on the same rigid instrument support plate with their boresights parallel. It was assumed that they would remain parallel and that the relative misalignments about all three axes would be null throughout the mission. Thus, the goal of the attitude determination system became to point the FPSSs at locations on the Sun to within 5 arc-sec (3σ).

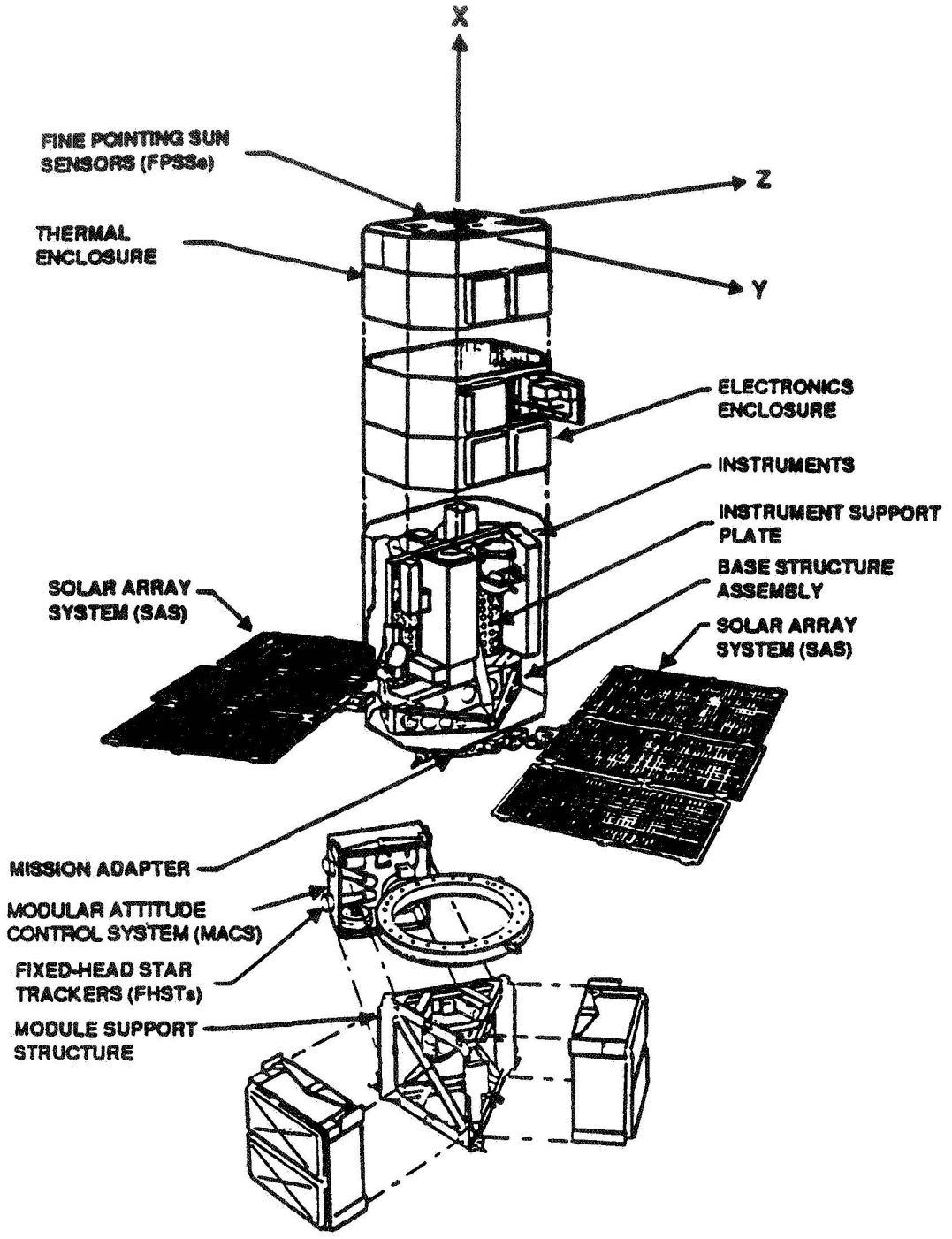


Figure 1. SMM Structure

The spacecraft attitude reference and body coordinate frames were defined so that the SMM attitude would be directly meaningful in terms of solar pointing. The reference frame was a noninertial frame in which the x-axis was defined to be the unit vector from the spacecraft to the Sun, the y-axis was defined as the unitized cross product of the x-axis and the direction of the solar north pole, and the z-axis completed the right-handed

orthogonal system. The body frame was defined by the FPSS coordinate axes. The x-axis of the body frame was defined as being parallel to the FPSS boresight. The y-, and z-axes of the body frame were defined as parallel to the z and negative y FPSS sensor axes. Thus, the output angles from the FPSS directly provided the attitude about the y- (pitch) and z-axes (yaw), which corresponded to the spherical coordinates of the pointing location on the Sun. Pointing the boresight of the FPSS (the x-axis of the spacecraft) to a specific location on the Sun required only successive pitch and yaw maneuvers.

The SMM FPSSs were vector sensors that measured rotations about two sensor axes and they had a field-of-view (FOV) of 2 deg $\times$ 2 deg and a specified accuracy of 5 arc-sec (3σ). Since the boresights of the two FPSSs were parallel, they provided attitude information only about two axes, corresponding to the pitch and yaw axes of the body reference system. The SMM FHSTs were also vector sensors that measured rotations about two sensor axes. These had an FOV of 8 deg $\times$ 8 deg and a specified accuracy of 30 arc-sec (3σ). The FOV's of the FHSTs did not overlap (i.e., their boresights were well separated). Thus, they could be used to compute a full three-axis attitude. Since the FPSSs provided only two axes of attitude information, the FHSTs were the prime source of roll attitude information.

The SMM Inflight Alignment Determination System

Since it was believed that the misalignment of the scientific instruments relative to the FPSSs would be negligible, the only alignment calibration needed was of the FHSTs relative to the FPSSs. Thus, a complete alignment calibration could be accomplished in orbit since the relative alignments are completely observable.[7] Therefore, for the SMM, the alignment calibration system should have been relatively simple and should have provided all the necessary alignment information necessary to compute the most accurate attitude of the spacecraft.

However, two assumptions were made in the prelaunch analysis of the alignments for the SMM that restricted the accuracy with which the alignments could be determined. The first assumption was, as stated above, that the alignment of the FPSSs relative to the scientific payload would not change. Since they were all mounted on a single rigid plate, it would seem that this was a valid assumption. However, because of it, no alignment calibration of the two FPSSs relative to each other was ever performed. If this assumption had not been made, the alignment of FPSS2 relative to FPSS1 could have been easily computed and the resulting value of the misalignment would have provided insight into the general level of misalignment between the FPSSs and the scientific payload.

The second assumption was that no roll alignment information of the two FHSTs relative to the FPSSs would be available. This followed supposedly from the fact that the FPSS boresights were parallel to the spacecraft body x-axis. Instead, it was assumed that the only roll alignment information available was the roll alignment of FHST2 relative to the roll alignment of FHST1. For this reason, the only alignment information ever calculated for the SMM were the pitch and yaw alignments of the FHSTs relative to FPSS1 and the roll alignment of FHST2 relative to FHST1.

The other notable point about the SMM alignment scheme was that the full inflight calibration algorithm generally required that the SMM be switched to an operating mode that eliminated the scientific data from the telemetry. The SMM scientists, however, were

very hesitant to switch to this mode except for absolutely necessary calibrations. When the alignment variation on the SMM was recognized, it was decided the alignment would be monitored each day. However, because the scientists could not be expected to agree to the mode switch every day, a pseudo-alignment algorithm that provided even less complete alignment information was devised in its place.

The pseudo-alignment algorithm was designed to monitor the pitch and yaw alignments of the FHSTs each day. It worked by subtracting the pitch and yaw attitude computed by the FHSTs from the pitch and yaw attitude computed by the FPSSs. The difference was attributed to misalignment of the FHSTs. Since the method worked by subtracting the computed attitudes, the alignments of the FHSTs were lumped together, i.e., it was assumed they moved as a single rigid element. Even though use of the pseudo-alignment algorithm provided very incomplete alignment information, it provided sufficiently good results that a characterization of the yaw alignment variation was possible. This will be further explored in the next section.

Review of Previous Work

Using the pseudo-alignment algorithm, it was discovered that the SMM FHSTs yaw alignment relative to FPSS1 varied over time and could be as large as 120 arc-sec. However, the pseudo-alignment algorithm showed no such variation for the alignments about the pitch axis. As mentioned earlier, it was thought that misalignment about the roll axis was unobservable.

An investigation into the possible causes of the alignment variation showed that the variation was correlated with the changing structural temperatures of the instrument support plate on which the FPSSs were mounted. After the data were filtered to remove some of the noise, a scatter plot could be constructed that showed the variation to be nearly linear. This scatter plot is shown in Fig 2. A least-squares straight line fit to the data yielded the following model for the yaw misalignment.

$$\mathrm{M} = -130.\ \text{arc-sec} + 11.\,\mathrm{T}\ \text{arc-sec}/^{o}\mathrm{C} \quad , \tag{2-1}$$

where M is the yaw misalignment and T is the spacecraft structural temperature which could be obtained from the regular telemetry. The computed accuracy of this equation, assuming the errors were normal and uncorrelated was 8.5 arc-sec (1σ).

Some questions still remained. For example, why was the yaw alignment temperature dependent while the pitch alignment seemingly was not? The previous work postulated an answer to this question which, however, was not demonstrated convincingly. In addition, because of the large amount of noise inherent in the pseudo-alignment calibration scheme, was it possible to better characterize the alignment dependence? In order to understand more fully the nature of the SMM alignments, it was decided to utilize more complete and rigorous methods of alignment determination to recalibrate the SMM alignments. This work is the subject of the remainder of this paper.

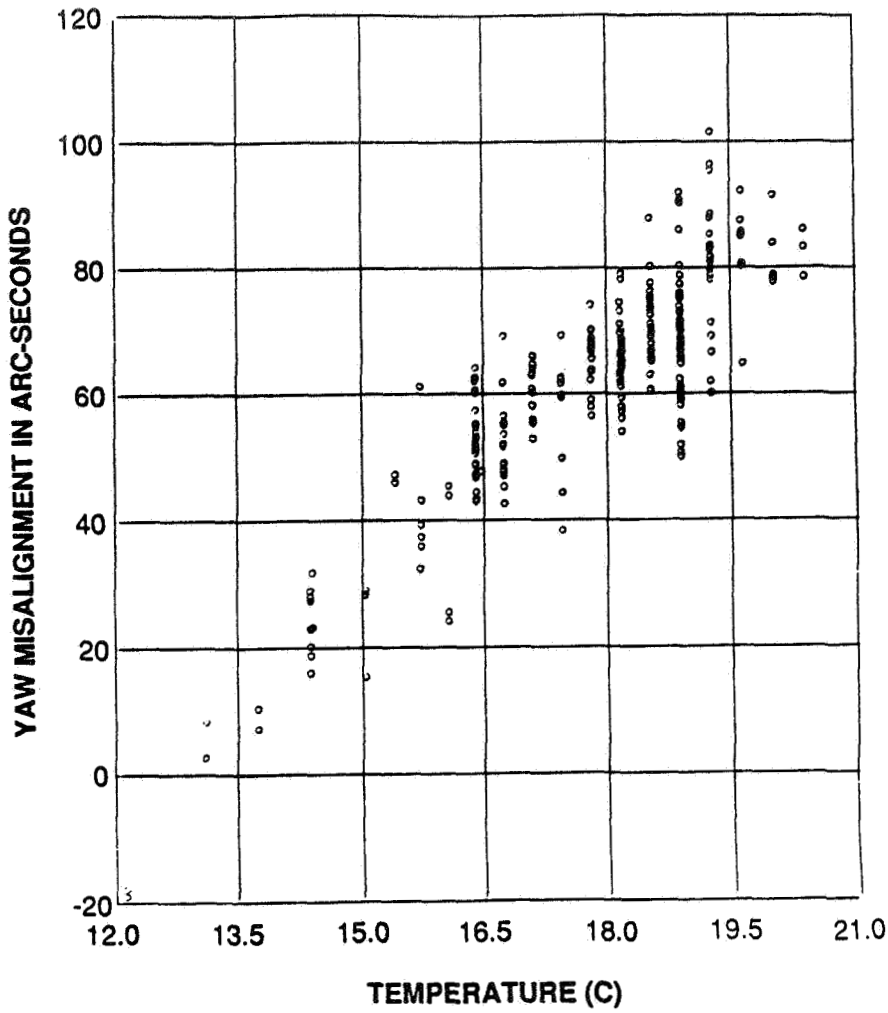


Figure 2. Scatter Plot of Yaw Misalignment and Temperature

3. A NEW BATCH ESTIMATOR OF SENSOR ALIGNMENTS

The batch alignment calibration algorithm used in this work[7–9] estimates alignments from derived attitude-independent effective measurements. (See also Ref. 10 for a comparison of this batch algorithm with sequential attitude-dependent algorithms.) This section reviews this algorithm and provides guidelines for its application. The reader is referred to Refs. 7 through 9 for a more complete description and derivation of the algorithm. A statistically correct method for computing the temperature dependence of the alignments will also be presented here. Realistic simulations are presented as well to demonstrate the power of these methods.

Basic Principles

The alignment matrix of a specific sensor i, denoted by S_i, may be defined as the orthogonal rotation matrix between the sensor coordinate frame and the body coordinate frame. This alignment matrix may be decomposed into two component rotation matrices. Thus, the alignment of sensor i at any time during the flight of a spacecraft may be written as

$$S_i = M_i S_i^o \quad , \tag{3-1}$$

where S_i^o is the prelaunch estimate of the sensor alignment matrix and M_i is the misalignment matrix, which represents the change in the alignment from the prelaunch value. The misalignment matrix may be represented by the rotation vector,[11] $\boldsymbol{\theta}_i$. If it is assumed that M_i represents a very small rotation, then

$$M_i = I + [[\boldsymbol{\theta}_i]] \quad , \tag{3-2}$$

where $[[\boldsymbol{\theta}_i]]$ is the antisymmetric matrix representation of a vector and can be defined as

$$[[\boldsymbol{\theta}_i]] = \begin{bmatrix} 0 & \theta_{i3} & -\theta_{i2} \\ -\theta_{i3} & 0 & \theta_{i1} \\ \theta_{i2} & -\theta_{i1} & 0 \end{bmatrix} \quad . \tag{3-3}$$

The vector, $\boldsymbol{\theta}_i$, which represents the misalignment of the sensor from its prelaunch value, is the quantity which is computed by the alignment algorithm.

To simplify the equations which follow, we define an uncalibrated observed vector as

$$\hat{\mathbf{W}}_{i,k}^o \equiv S_i^o \hat{\mathbf{U}}_{i,k} \quad , \tag{3-4}$$

where $\hat{\mathbf{U}}_{i,k}$ is the observed vector by sensor i, $i = 1, 2, \ldots, n$, at time t_k, $k = 1, 2, \ldots, N$, in the sensor coordinate frame. Thus, $\hat{\mathbf{W}}_{i,k}^o$ is the representation of the observed vector in the body frame uncorrected for the misalignment of the sensor from the prelaunch alignment estimate.

The effective attitude-independent scalar measurement used in this work is the difference of the cosine of the angle between the observed vectors of sensors i and j from the cosine of the angle between the corresponding reference vectors. Assuming that the misalignments and the sensor noise is small, the effective measurement equation can be derived as

$$z_{ij,k} = \hat{\mathbf{W}}_{i,k}^o \cdot \hat{\mathbf{W}}_{j,k}^o - \hat{\mathbf{V}}_{i,k} \cdot \hat{\mathbf{V}}_{j,k} = (\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{j,k}^o) \cdot (\boldsymbol{\theta}_i - \boldsymbol{\theta}_j) + \Delta z_{ij,k} \quad , \tag{3-5}$$

where $\Delta z_{ij,k}$, the effective measurement noise, is

$$\Delta z_{ij,k} = \hat{\mathbf{W}}_{i,k}^o \cdot \Delta\hat{\mathbf{W}}_{j,k}^o + \hat{\mathbf{W}}_{j,k}^o \cdot \Delta\hat{\mathbf{W}}_{i,k}^o \quad , \tag{3-6}$$

where we have assumed that the errors in the observations are much larger than those in the reference vectors.

For the system of n attitude sensors, these scalar measurements at a common time t_k may be collected in a single column vector which we may write as

$$\mathbf{Z}_k = H_k \, \Theta + \Delta \mathbf{Z}_k \quad , \tag{3-7}$$

where H_k and $\Delta \mathbf{Z}_k$ are determined from equations (3-5) and (3-6). Since at any time t_k, n sensed unit vectors correspond to $2n$ equivalent one-dimensional measurements, and three measurements are needed to determine the attitude, there can be only $2n-3$ independent $z_{ij,k}$ at t_k. Thus, if we wish the covariance matrix of $\Delta \mathbf{Z}_k$ to be full-rank, then $\mathbf{Z}_k$ and $\Delta \mathbf{Z}_k$ are $(2n-3)$-dimensional column vectors. Θ is a $3n$-dimensional column vector, and H_k is a $(2n-3) \times 3n$ matrix. The covariances of $\Delta z_{ij,k}$, if we assume the QUEST model for the sensor measurement errors,[12] are given by

$$E\{\Delta z_{ij,k}\} = 0 \quad , \tag{3-8}$$

$$E\{\Delta z_{ij,k} \Delta z_{ij,k}\} = (\sigma_i^2 + \sigma_j^2)(\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{j,k}^o) \cdot (\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{j,k}^o) \quad , \tag{3-9}$$

$$E\{\Delta z_{ij,k} \Delta z_{i\ell,k}\} = \sigma_i^2 \, (\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{j,k}^o) \cdot (\hat{\mathbf{W}}_{i,k}^o \times \hat{\mathbf{W}}_{\ell,k}^o) \quad , \tag{3-10}$$

$$E\{\Delta z_{ij,k} \Delta z_{\ell m,k}\} = 0 \quad . \tag{3-11}$$

To demonstrate how the components of the measurement equation are formed, a simple three sensor example will be presented. The three scalar measurements are

$$z_{12,k} = \hat{\mathbf{W}}_{1,k}^o \times \hat{\mathbf{W}}_{2,k}^o \cdot (\boldsymbol{\theta}_1 - \boldsymbol{\theta}_2) + \Delta z_{12,k} \quad , \tag{3-12}$$

$$z_{13,k} = \hat{\mathbf{W}}_{1,k}^o \times \hat{\mathbf{W}}_{3,k}^o \cdot (\boldsymbol{\theta}_1 - \boldsymbol{\theta}_3) + \Delta z_{13,k} \quad , \tag{3-13}$$

$$z_{23,k} = \hat{\mathbf{W}}_{2,k}^o \times \hat{\mathbf{W}}_{3,k}^o \cdot (\boldsymbol{\theta}_2 - \boldsymbol{\theta}_3) + \Delta z_{23,k} \quad . \tag{3-14}$$

Thus, the 3×1 matrix $\mathbf{Z}_k$ is

$$\mathbf{Z}_k = [\, z_{12,k}, z_{13,k}, z_{23,k} \,]^T \quad , \tag{3-15}$$

and the 9×1 matrix Θ is

$$\Theta = [\, \boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \boldsymbol{\theta}_3^T \,]^T = [\, \theta_{11}, \theta_{12}, \theta_{13}, \theta_{21}, \theta_{22}, \theta_{23}, \theta_{31}, \theta_{32}, \theta_{33} \,]^T \quad , \tag{3-16}$$

where the misalignment angle, θ_{ij}, in equation (3-16) refers to misalignment angle j of sensor i. The matrix H_k is formed as

$$H_k = \begin{bmatrix} (\hat{\mathbf{W}}_{1,k}^o \times \hat{\mathbf{W}}_{2,k}^o)^T & -(\hat{\mathbf{W}}_{1,k}^o \times \hat{\mathbf{W}}_{2,k}^o)^T & \mathbf{0}^T \\ (\hat{\mathbf{W}}_{1,k}^o \times \hat{\mathbf{W}}_{3,k}^o)^T & \mathbf{0}^T & -(\hat{\mathbf{W}}_{1,k}^o \times \hat{\mathbf{W}}_{3,k}^o)^T \\ \mathbf{0}^T & (\hat{\mathbf{W}}_{2,k}^o \times \hat{\mathbf{W}}_{3,k}^o)^T & -(\hat{\mathbf{W}}_{2,k}^o \times \hat{\mathbf{W}}_{3,k}^o)^T \end{bmatrix} \quad . \tag{3-17}$$

Relative alignments are defined as the alignment of an attitude sensor relative to another attitude sensor. For convenience, the sensor to which the relative alignments are being determined will be designated as sensor 1. Thus,

$$\boldsymbol{\psi}_i \equiv \boldsymbol{\theta}_i - \boldsymbol{\theta}_1 \quad , \tag{3-18}$$

where ψ_i is the alignment of sensor i relative to sensor 1. The $(3n-3)$-dimensional vector of relative alignments can likewise be found from the absolute alignment vector as

$$\Psi \equiv [\psi_2^T, \psi_3^T, \ldots, \psi_n^T]^T = F\Theta \quad , \tag{3-19}$$

where F is the $(3n-3) \times 3n$ matrix and is defined as

$$F \equiv \begin{bmatrix} -I_{3\times3} & I_{3\times3} & 0_{3\times3} & \cdots & 0_{3\times3} \\ -I_{3\times3} & 0_{3\times3} & I_{3\times3} & \cdots & 0_{3\times3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -I_{3\times3} & 0_{3\times3} & 0_{3\times3} & \cdots & I_{3\times3} \end{bmatrix} \quad . \tag{3-20}$$

From equation (3-5) it is clear that $\mathbf{Z}_k$ is sensitive only to the relative alignments and we may write

$$\mathbf{Z}_k = H'_k \Psi + \Delta\mathbf{Z}_k \quad , \tag{3-21}$$

where H'_k is the $(2n-3) \times (3n-3)$ matrix obtained by deleting the first three columns of H_k. As stated earlier, the relative alignment vector has dimension $3n-3$ and is completely observable from the inflight data.

The Attitude-Independent Inflight Estimator

Using the measurement model developed above the negative-log-likelihood function[13] may be written as

$$J_\Psi(\Psi) = \frac{1}{2}\sum_{k=1}^{N} \left\{ \left[(\mathbf{Z}_k - H'_k\Psi)^T P_{Z_k}^{-1} (\mathbf{Z}_k - H'_k\Psi)\right] + \log\det P_{Z_k} + (2n-3)\log 2\pi \right\} \quad . \tag{3-22}$$

which is minimized to obtain the maximum likelihood estimate of the relative alignments. Minimizing this expression leads to the usual normal equations

$$P_{\Psi\Psi}^{-1}(\mathrm{PF}) = \sum_{k=1}^{N} H_k'^T P_{Z_k}^{-1} H'_k \quad , \tag{3-23}$$

$$\Psi^*(\mathrm{PF}) = P_{\Psi\Psi}(\mathrm{PF}) \sum_{k=1}^{N} H_k'^T P_{Z_k}^{-1} \mathbf{Z}_k \quad , \tag{3-24}$$

where "PF" denotes *prior-free* and indicates that the estimate is based only on the inflight data and not on any prior (i.e., prelaunch) estimate of the alignments.

Determination of the Temperature Dependence of the Alignments

As was shown in section 2, the SMM yaw alignments depend on the spacecraft structural temperature variations. Thus, we seek to use maximum-likelihood estimation methods to

better quantify the dependence of the relative alignments on the temperature. The relative misalignments are assumed to depend linearly on the temperature. Thus, we write

$$\Psi_{T_i}^{\text{true}} = \mathbf{a} + \mathbf{b}(T_i - T_o) \quad , \tag{3-25}$$

where T_o is the reference temperature at which the alignments were determined during the prelaunch alignment calibration and T_i is the spacecraft structural temperature at which the inflight misalignment is estimated to yield $\Psi_{T_i}^*(\text{PF})$. We wish to compute the values of the $(3n-3) \times 1$ coefficient vectors, **a** and **b**. Thus we write,

$$\Psi_{T_i}^*(\text{PF}) = \Psi_{T_i}^{\text{true}} + \Delta\Psi_{T_i}^*(\text{PF}) \quad , \tag{3-26}$$

where

$$E\{\Delta\Psi_{T_i}^*(\text{PF})\} = \mathbf{0} \quad , \tag{3-27}$$

$$E\{\Delta\Psi_{T_i}^*(\text{PF})\,\Delta\Psi_{T_i}^{*T}(\text{PF})\} = P_{\Psi_{T_i}}(\text{PF}) \quad , \tag{3-28}$$

where at each temperature T_i, $\Psi_{T_i}^*(\text{PF})$ and $P_{\Psi_{T_i}}(\text{PF})$ are obtained from equations (3-23) and (3-24).

Substituting equation (3-25) into equation (3-26) yields

$$\Psi_{T_i}^*(\text{PF}) = \mathbf{a} + \mathbf{b}(T_i - T_o) + \Delta\Psi_{T_i}^*(\text{PF}) \quad , \tag{3-29}$$

and $\Psi_{T_i}^*(\text{PF})$ serves now as an effective measurement.

Defining the parameter vector

$$\Lambda = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix} \quad , \tag{3-30}$$

the measurement equation becomes

$$\Psi_{T_i}^*(\text{PF}) = H'_{T_i}\Lambda + \Delta\Psi_{T_i}^*(\text{PF}) \quad , \tag{3-31}$$

where

$$H'_{T_i} = [I \quad I(T_i - T_o)] \quad . \tag{3-32}$$

The coefficient vector Λ may be obtained by minimizing the negative-log-likelihood function

$$J_\Lambda(\Lambda) = \frac{1}{2}\sum_{T_i}\Big\{\Big[(\Psi_{T_i}^*(\text{PF}) - H'_{T_i}\Lambda)^T P_{\Psi_{T_i}}^{-1}(\text{PF})\,(\Psi_{T_i}^*(\text{PF}) - H'_{T_i}\Lambda)\Big] + \log\det P_{\Psi_{T_i}}) + N_T \log 2\pi\Big\} \quad , \tag{3-33}$$

where N_T is the number of temperatures at which the inflight estimates of the relative alignments has been estimated. Carrying out the minimization leads to[5]

$$P_{\Lambda\Lambda}^{-1} = \sum_{T_i} H_{T_i}^{\prime T} P_{\Psi_{T_i}}^{-1}(\text{PF})\, H'_{T_i} \quad , \tag{3-34}$$

$$\Lambda^* = P_{\Lambda\Lambda} \sum_{T_i} H_{T_i}^{\prime T} P_{\Psi_{T_i}}^{-1}(\text{PF})\Psi_{T_i}^*(\text{PF}) \quad . \tag{3-35}$$

Simulations

The algorithms developed above were applied to two simulations, the first a very simple spacecraft in which the alignments do not depend on the temperature and the second a spacecraft configured like the SMM and whose alignments are temperature-dependent. For each simulation a set of model misalignments was computed in order to generate the observed and reference vectors and to judge the effectiveness of the simulated calibrations. The distribution of the misalignments was assumed to be Gaussian and zero-mean. This particular model was chosen to test an algorithm for estimating launch-shock error levels[7,9] not relevant to the SMM application as presented here.

The first simulation is of a spacecraft with three attitude sensors with well separated boresights. The model sensors were taken to have an FOV of 20 deg by 20 deg and an accuracy of 10 arc-sec. One hundred frames of simulated data were generated in which each sensor was assumed always to have valid data. The results are shown in Table 1. The agreement between the model misalignments and the estimates is consistent with the computed standard deviations.

TABLE 1

Comparison of Model and Estimated
Relative Misalignments for Simulation 1

Model Relative Misalignment (arc-sec)	Estimated Relative Misalignment (arc-sec)
−73.	−73. ± 1.
−40.	−50. ± 8.
63.	78. ± 13.
−14.	−7. ± 7.
−43.	−45. ± 1.
131.	146. ± 12.

For the second simulation, where model misalignments were needed at a range of distinct temperatures, the coefficient vector **b** was given arbitrary values while **a** was sampled from a zero-mean Gaussian distribution. The model relative misalignments at the remaining temperatures were computed from

$$\Psi_{T_i} = \mathbf{a} + \mathbf{b}\,(T_i - T_o) \quad . \tag{3-36}$$

The types of sensors and their size and accuracies were modeled after the SMM FPSSs and FHSTs. The temperature range of the plate on which the FPSSs were mounted

TABLE 2
Comparison of Model and Estimated Coefficients of the
Temperature Dependence of Misalignments For Simulation 2

Model Constant Term **a** (arc-sec)	Estimated Constant Term **a** (arc-sec)	Model Linear Term **b** (arc-sec/°C)	Estimated Linear Term **b** (arc-sec/°C)
−43.	$-47. \pm 14.$	0.	0.0 ± 5.0
0.	$1. \pm 1.$	0.	0.2 ± 0.1
0.	$-1. \pm 1.$	0.	0.2 ± 0.1
−44.	$-52. \pm 10.$	60.	60.1 ± 4.0
100.	$91. \pm 5.$	−30.	$-29.9. \pm 2.0$
−128.	$-119. \pm 5.$	30.	$29.9. \pm 2.0$
158.	$131. \pm 11.$	60.	60.1 ± 4.0
−39.	$-48. \pm 5.$	−30.	-29.9 ± 2.0
−189.	$-179. \pm 5.$	30.	29.9 ± 2.0

was modeled as varying between 2°C and 10°C with data generated at 2°C intervals. The reference temperature was taken as 6°C. As for the SMM, the misalignments were computed relative to the first sensor. The simulated calibration was performed at each temperature and the results were collected and used in the algorithm of equations (3-34) and (3-35) to estimate the dependence of the misalignments on temperature.

The results are shown in Table 2. Again the agreement between the model and estimated alignments is consistent with the computed standard deviations. Thus, for both simulations the algorithms worked well and provided meaningful error bounds.

4. APPLICATION TO THE SMM

In this section the newly derived algorithms will be applied to data from the SMM in order to better quantify the dependence of the SMM alignments on temperature variations. In addition, the limiting assumptions made during the SMM prelaunch analysis will also be examined in greater detail.

Application of Algorithms to SMM Data

Since this study was conducted nearly five years after the data from the heuristic analysis was collected, provisions had to be made to collect sufficient statistics. In order to make the best comparisons, it was decided to examine the data from after the repair of the SMM

until the time of the first null redefinition in December of 1984. During this time, the SMM structural temperatures ranged from 12.12°C to 20.16°C. It was attempted to collect three orbits of data at each discrete temperature point. However, to obtain good observability of the alignments, data must be well spread throughout a sensor field-of-view. Since the spacecraft was pointed so that the Sun would be very close to the FPSS boresight, it was difficult to obtain much variation in the FPSS observations, even by combining all data sets at the same temperature. Thus, it was impossible to gain any observability about the roll axes of the FPSSs (which in turn prevented any observability of the roll axes of the FHSTs relative to the FPSSs).

Despite this limitation of the data, the calibrations were carried out and the results are shown in Fig. 3 through Fig. 6. These graphs show several very interesting results. First, as unseen from the heuristic analysis, a temperature dependence about the pitch axis is evident. The slope of the dependence seems to be nearly the same as that of the yaw dependence but opposite in sign. Thus, it remained to determine why this was occurring. It is possible that noise from other effects obscured the pitch dependence during the heuristic analysis.

The main source of noise in the pseudo-alignment solutions was from the FHST attitude solutions. The accuracy of these solutions depended on the quality of the star fields observed by the FHSTs. Thus, in order to resolve the discrepancy between the two results, the pitch data from the heuristic analysis was reduced. The first step was to identify periods of time when poor star fields were used by the pseudo-alignment algorithm. It turned out that periods could be found, of length one to two weeks, where poor star data was used that greatly increased the noise level in the solutions. Elimination of these periods and replotting of the remaining data as a scatter plot of temperature and pitch misalignment results in a nearly linear dependence with slope opposite that of the yaw misalignment. This plot is shown in Fig. 7. Thus, after this further analysis, it appears the two methods agree.

The second point to notice from Figs. 3 through 6 is that the slopes of the dependence using the new algorithms are opposite those gained from the heuristic analysis. This can be seen to be correct when it is recalled that the pseudo-alignments were calculated by subtracting the FHST attitude solution from the FPSS attitude solution. Thus, the pseudo-alignments were actually, according to equation (3-18), alignments of the FPSSs relative to the FHSTs. Since this is opposite to what was computed by the new algorithms, it would be expected that the earlier slopes would be opposite in sign.

Using the algorithms developed above to compute the actual temperature dependence of the alignments, the coefficients were solved for and the results are shown in Table 3. As can be seen, the linear dependence was significant about both the pitch and yaw axes. The magnitude of the dependence as calculated by the new algorithms was several times greater than that computed by the heuristic methods because the latter was masked by the larger random errors in the alignments.

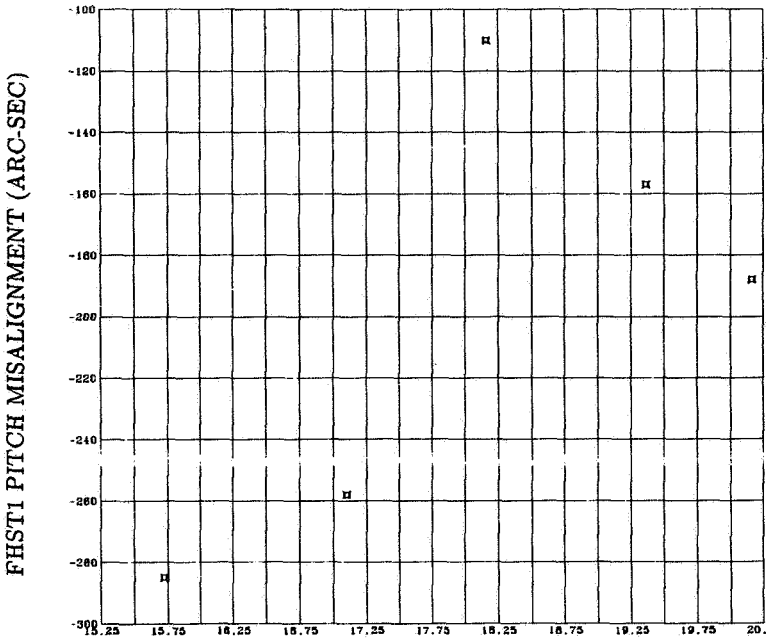


Figure 3. FHST1 Pitch Misalignment

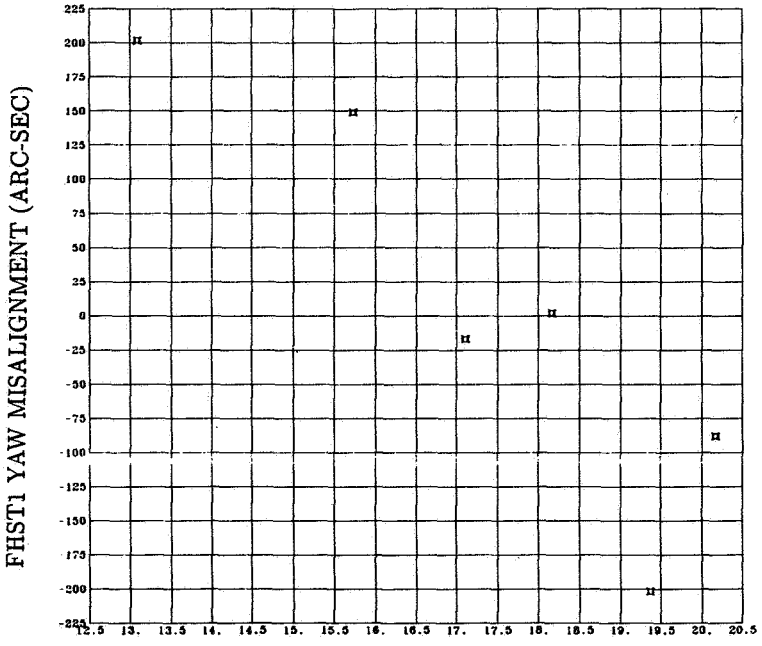


Figure 4. FHST1 Yaw Misalignment

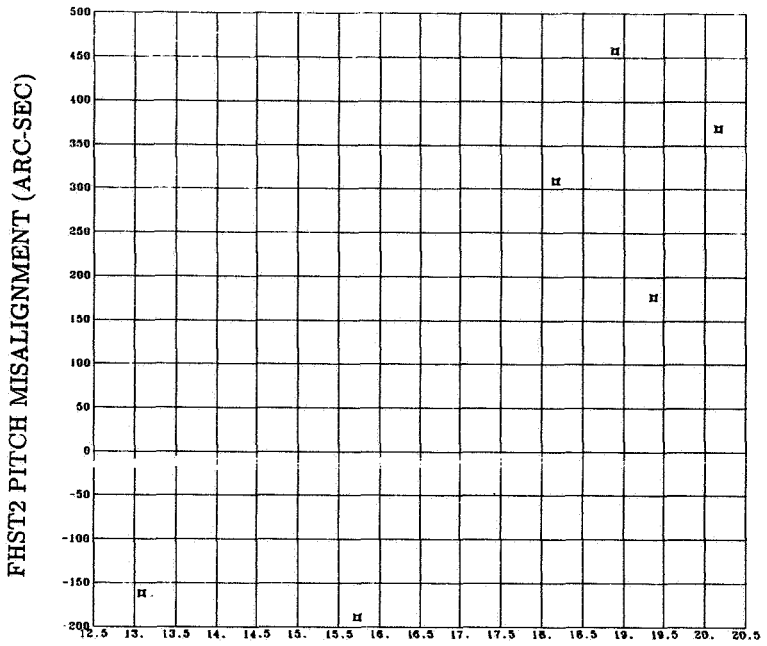


Figure 5. FHST2 Pitch Misalignment

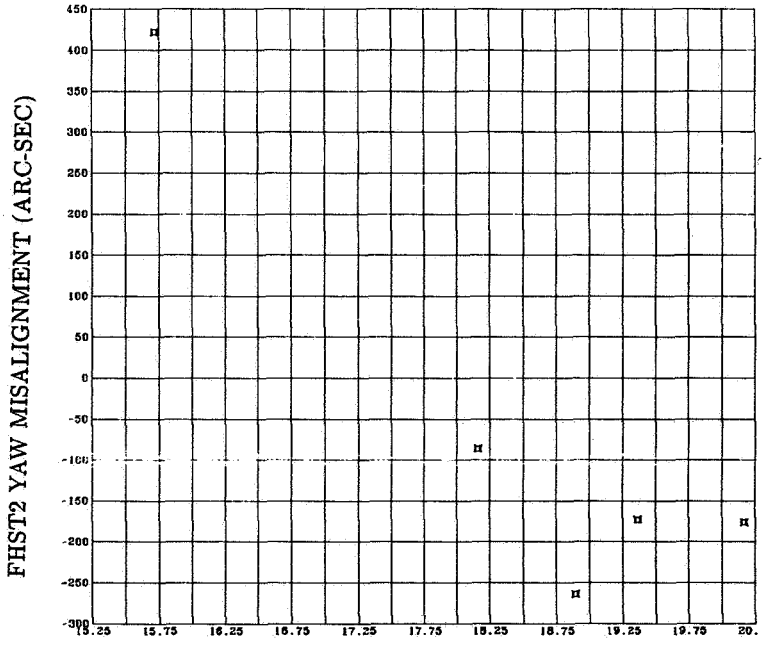


Figure 6. FHST2 Yaw Misalignment

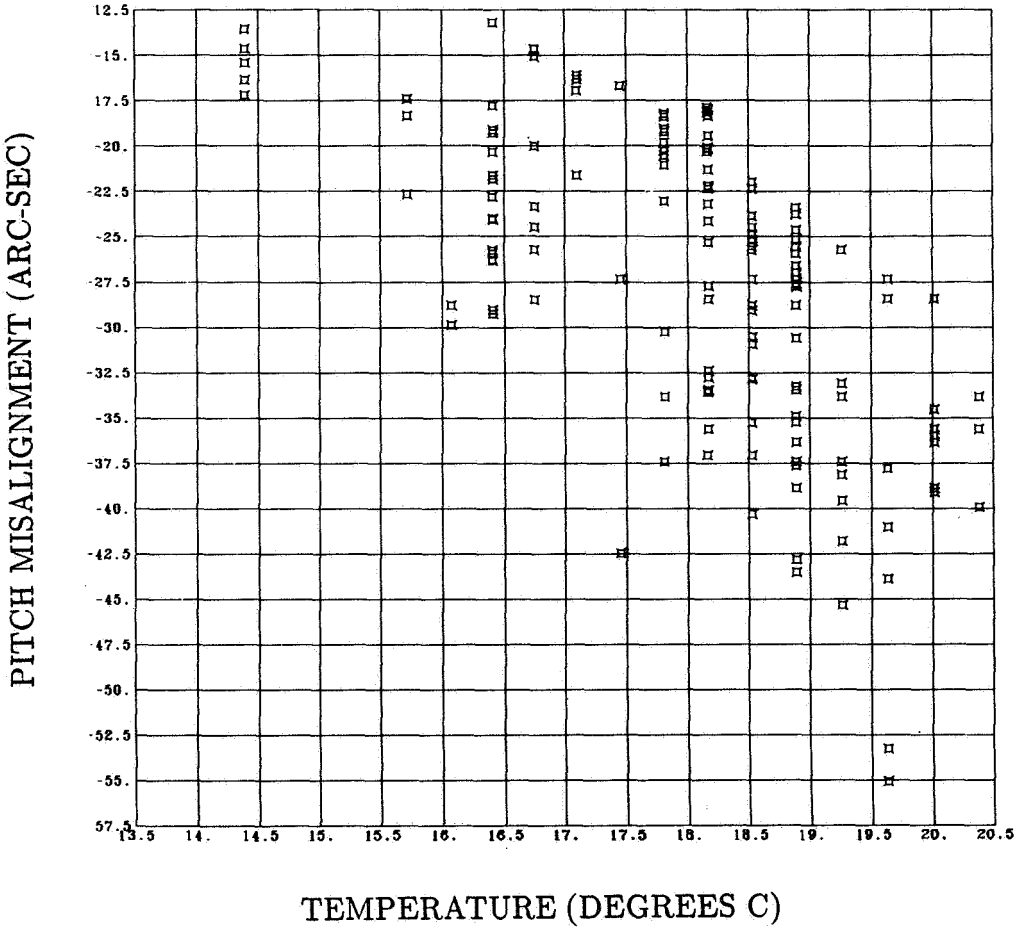


Figure 7. Reduced Pitch Misalignment Data

TABLE 3

Coefficients of the Temperature Dependence of the Alignments
Relative to Solar Maximum Mission Fine Pointing Sun Sensor 1

Sensor Axis	Constant Term a (arc-sec)	Linear Term b (arc-sec/°C)
FPSS2 PITCH	0.7 ± 0.1	0.0 ± 0.0
FPSS2 YAW	0.5 ± 0.1	0.0 ± 0.0
FHST1 PITCH	-164.7 ± 3.7	37.1 ± 1.4
FHST1 YAW	17.8 ± 4.3	-33.5 ± 1.7
FHST2 PITCH	278.2 ± 2.9	48.0 ± 1.1
FHST2 YAW	-100.8 ± 3.4	-40.9 ± 1.3

Thus, there is good agreement between the two methods about the nature and dependence of the SMM FHST pitch and yaw alignments relative to the FPSSs. The new algorithms clearly show better results because they avoid the large error sources that caused the heuristic methods to miss the pitch alignment dependence and they treated the alignments of the FHSTs separately which provided better information as to the actual behavior of each one.

Examination of SMM Alignment Assumptions

As stated earlier, two assumptions were made regarding the expected nature of the SMM fine attitude sensor alignments. In this section, each assumption will be examined to ascertain its validity.

The first assumption was that the alignment of the SMM payload relative to the FPSSs would not change from its prelaunch value of zero. While no real data is available from the SMM scientific instruments to be used in any alignment algorithm, this assumption could still be tested. Since there were two FPSSs that were mounted on the instrument support plate along with the scientific payload, any alignment variation between the FPSSs should be of the same order of magnitude as that between the FPSSs and the payload. Thus, the alignment of the FPSSs relative to each other was studied using the new algorithms. This study showed that the alignment of FPSS2 relative to FPSS1 showed no significant variation, and the values of the misalignment angles were never greater than 1 or 2 arc-sec. Thus, the assumption that the payload alignment never varied can be judged as valid from the limited information available.

The second assumption was that no roll alignment information involving the FPSSs was observable. As stated in Ref. 6, all the relative alignments are observable, and since the alignments of the SMM were computed relative to FPSS1, they should all be observable. However, the observability of the alignment about an axis depended on the spread of data

in the sensors that are sensitive to that axis. Because all the data that was examined in the previous section had a poor spread in the FPSS, poor observability was obtained about the FPSS roll axis. In fact, this is the crux of the assumption that no observability is possible about this axis because the spread could never be good. However, it is asserted here that if a good spread could be obtained, the SMM roll axis misalignment could have been determined within a reasonable error. Thus, data with a good spread in the FPSS FOV had to be found.

It was discovered that data from the SMM that had a good spread in the FPSS FOV was available. During the times of the FPSS electronic response calibrations, that calibrated the transformation of FPSS measurements in units of FPSS counts to degrees, the SMM was rotated in such a way that the observed Sun vector had a significant variation in the FPSS FOV. This data was obtained and the calibration was performed. The computed alignments had variances of 60 arc-sec for the determination of the FHST roll alignments. During the normal calibrations when the spread of data was poor, the variance was close to ten thousand arc-sec. Thus, the simple rotations during the electronic response calibrations provided great improvement in the misalignment determination capability.

During the electronic response calibrations, the rotations were as large as one quarter of a degree; thus, the angle between the spacecraft to Sun vector and the FPSS boresight vector reached one quarter of a degree. Since the FPSS field-of-view was 2 deg by 2 deg, this Sunline angle could have conceivably reached one degree. However, no usable data was available where the Sunline angle was greater than one quarter of a degree. Thus, as a test of the possible observability of the SMM roll misalignments, simulations were done, using the exact SMM configuration, that increased the Sunline angle to one half degree and one degree to improve the spread of data in the FPSS field-of-view. The misalignment calibrations at one half degree decreased the variance of the roll misalignment solutions to 30 arc-sec and the misalignment calibrations at one degree lowered the variance to 16 arc-sec. Thus, by simply rotating the spacecraft to the full capability of the SMM FPSSs, full observability of all the misalignments could have been obtained.

5. CONCLUSIONS

A full analysis of the nature of the fine attitude sensor alignment behavior on the Solar Maximum Mission has been presented. The dependence of the alignments on the spacecraft structural temperature variations has been investigated, and results have shown that simple equations could be derived to fully account for the dependence. In addition, an assumption on the observability of the alignments has been shown to be unnecessary. The results of this work can be applied by other missions so that they may be able to explain the variations in the sensor alignments.

6. REFERENCES

[1] Hewitt, D., *Multimission Modular Spacecraft (MMS) Thermal System Specification*, Goddard Space Flight Center, S-700-12, February 1977.

[2] Patt, F., *Analysis of the Solar Maximum Mission (SMM) Fine Pointing Sun Sensor (FPSS) Response Changes*, Computer Sciences Corporation, CSC/TM-85/6714, May 1985.

[3] Pitone, D., Eudell, A., and Patt, F., "Temperature Dependence of Attitude Sensor Coalignments on the Solar Maximum Mission," *Proceedings, Flight Mechanics/Estimation Theory Symposium*, NASA Goddard Space Flight Center, Greenbelt, Maryland, May 1989.

[4] Pitone, D., Eudell, A., and Patt, F., "Temperature Dependence of Attitude Sensor Coalignments on the Solar Maximum Mission," *Proceedings, AIAA/AAS Astrodynamics Specialist Conference*, Stowe, Vt., August 1989.

[5] Pitone, D., *The Estimation of Attitude Sensor Misalignment Due to Spacecraft Bending on the Solar Maximum Mission*, Masters Thesis, Howard University, May 1990.

[6] Pitone, D., *Solar Maximum Mission End-of-Mission Document*, Computer Sciences Corporation, CSC/TM-90/6058, April 1990.

[7] Shuster, M. D., Pitone, D. S., and Bierman, G. J., "Batch Estimation of Spacecraft Sensor Alignments, "submitted to *the Journal of the Astronautical Sciences.*

[8] Bierman, G. J., and Shuster, M. D., "Spacecraft Alignment Estimation," *Proceeedings, 27th IEEE Conference on Decision and Control*, Austin, Texas, December 1988.

[9] Shuster, M. D., and Pitone, D. S., "Consistent Estimation of Spacecraft Sensor Alignments," *Proceedings, American Control Conference*, San Diego, California, May 1990.

[10] Shuster, M. D., "Inflight Estimation of Spacecraft Sensor Alignment," *Proceedings, 13th AAS Guidance and Control Conference*, Keystone, Colorado, February, 1990.

[11] Wertz, J., (Ed.), *Spacecraft Attitude Determination and Control*, D. Reidel, Dordrecht, the Netherlands, 1978.

[12] Shuster, M. D. and Oh, S. D., "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, pp. 70–77, 1981.

[13] Sorenson, Harold, *Parameter Estimation*, Marcel Dekker, New York and Basel, 1980.

INFLIGHT ESTIMATION OF GYRO NOISE*

O. H. Filla, T. Z. Willard, and D. Chu
Computer Sciences Corporation

J. Deutschmann
Flight Dynamics Analysis Branch
NASA/Goddard Space Flight Center

ABSTRACT

A method is described and demonstrated for estimating single-axis gyro noise levels in terms of the Farrenkopf model parameters. This is accomplished for the Cosmic Background Explorer (COBE) by comparing gyro-propagated attitudes with less accurate single-frame solutions and fitting the squared differences to a third-order polynomial in time. Initial results are consistent with the gyro specifications, and these results are used to determine limits on the duration of batches used to determine attitude. Sources of error are discussed, and guidelines for a more elegant implementation, as part of a batch estimator or filter, are included for future work.

*This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, Contract NAS 5-31500.

1. INTRODUCTION

1.1 WHY WORRY ABOUT GYRO NOISE?

The usual batch approach to attitude estimation ignores the contribution of propagation to solution error. As long as the gyro angular rates are perfect, such simplification is justified. Of course, true perfection is never achieved. Even if systematic errors due to biases, scale factors, and misalignments are removed, there is still a random component of the error, which is called "noise." Some noise is completely random. Its value at one instant says nothing about that at the next. This is "white noise," and it shows up as a jittery line (Figure 1). Closer inspection reveals slowly varying biases on the gyro rates. Although both of these random errors average to zero, at least over infinite time, their instantaneous effects are not zero and tend to increase in size with propagation time.

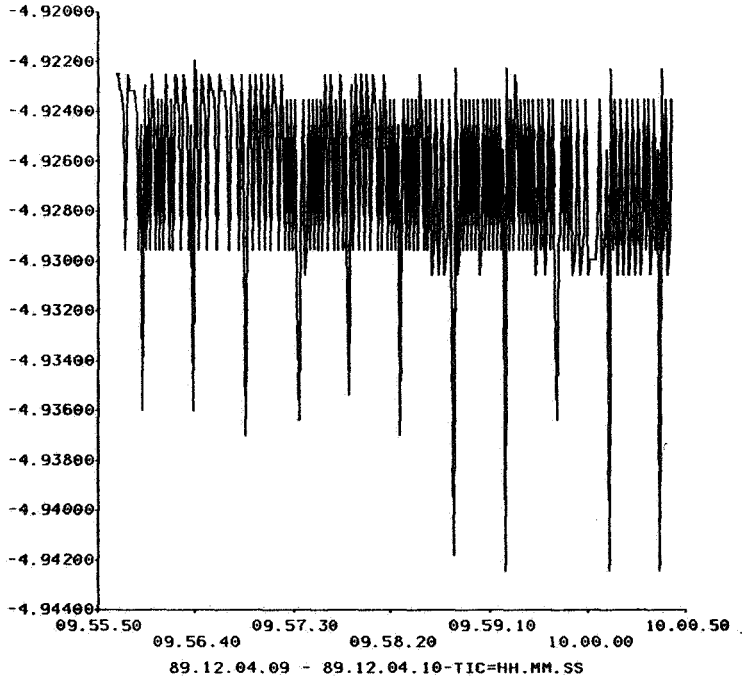


Figure 1. Noisy COBE X-Gyro Rates

To control the effect of gyro noise, the propagation time and batch length must be limited. Batch length depends on the frequency and accuracy of the observations and the gyro noise levels. Kalman filtering addresses the propagation error problem more elegantly by gradually forgetting past observations, and it can incorporate models that automatically account for random process noise. Whether a batch estimator or a filter is used, to obtain the best results the gyro noise level must be known.

When the impact of gyro noise was studied prior to launch of the Cosmic Background Explorer (COBE), the X-gyro scale factor instability appeared to limit batch duration to

20 minutes, or one-fifth of an orbit (Reference 1). This was a serious constraint, because the batch estimator is often used to average out sensor errors that vary at the orbit rate. Short batches would not be as effective for this purpose. To quantify the impact of gyro noise systematically, the COBE ground support software was enhanced to estimate noise parameters, attitude, and calibrations.

1.2 THE APPROACH

Gyro noise shows up in several ways. It is directly observable in the rate data, as Figure 1 illustrated. The most obvious noise measure is the variance of the rate, but variance includes real motion and so overestimates the noise. Another method is to set some cutoff frequency above which the spacecraft cannot move and then find the variance of the rate minus the components below that frequency. However, this leaves out any random drift of the rate and so underestimates the noise. A third method is to obtain limits on the total rate noise. The long-term variation of the rate is also observable in the history of the drift rate bias. Computing the bias for many orbits over several days, as in Figure 2, provides information from which to estimate the very low frequency part of the rate noise.

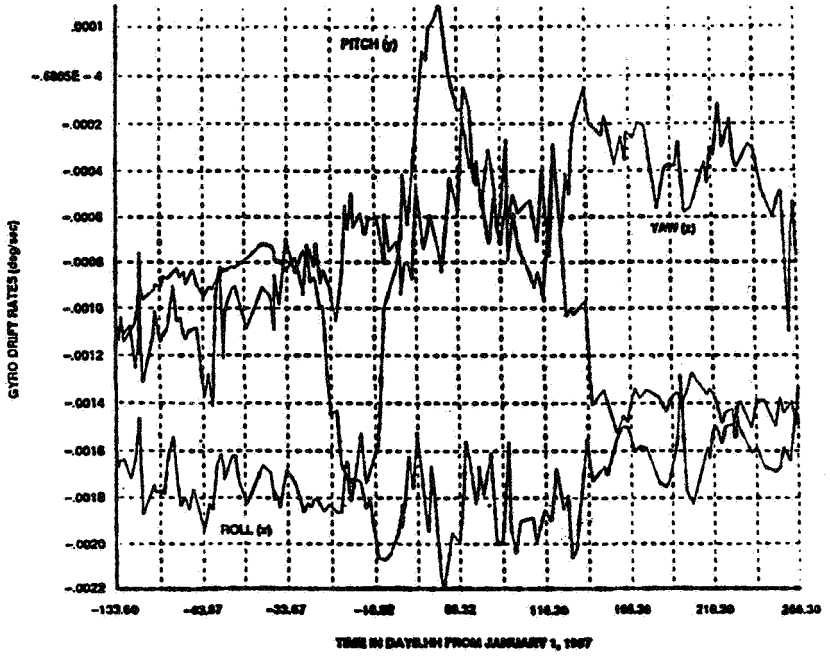


Figure 2. History of the ERBS Gyro Biases

Unfortunately, estimating noise directly from the rates does not give a very complete picture of the gyro noise. This is particularly true over the frequency midrange from 100 hertz (Hz) to 1 revolution per orbit, where it is difficult to distinguish between motion and noise. To distinguish motion from noise, one has to take advantage of the sensor observations, and this means working with attitudes instead of rates.

COBE has two attitude estimators; the batch estimator uses the gyros, while the single-frame estimator does not. By comparing the two estimates, one can isolate the effects of gyro noise. The squared difference between them becomes an observation of the combined constant variance in each and the growing variance due to propagation error. With a model for how that variance grows with time, one can construct an estimator for the model parameters.

2. THE GYRO NOISE MODEL

2.1 OVERVIEW

One popular method of estimating the effect of gyro noise is the Farrenkopf model (Reference 2). For the COBE rate-integrating gyros operating in rate mode, the output is integrated only over the 1-second interval between gyro samples. These incremental angles, through which the spacecraft rotates between samples, are just like samples of the angular rate times the constant sampling interval. For this reason, the gyros are treated here as rate gyros rather than rate-integrating gyros. The model includes four noise sources, and for each noise source it computes a transfer function relating the noise to the output angle. From the inverse transfer functions, an expression is obtained to relate the variance of the noise sources to that of the total rotation angle obtained from the sum of the incremental angles.

2.2 THE PHYSICS OF RATE GYROS—OPEN LOOP

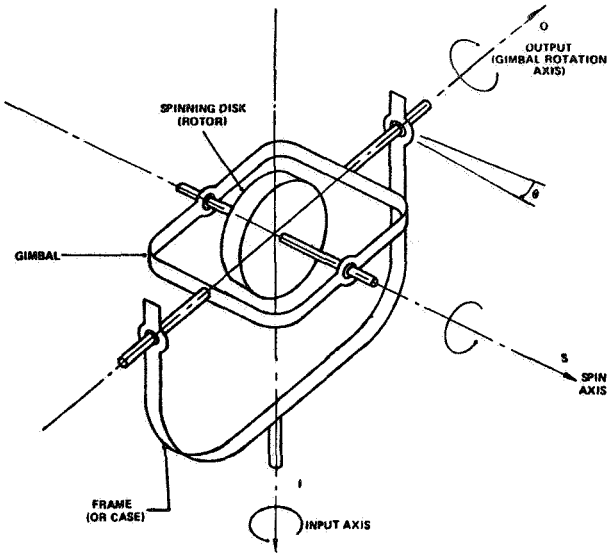
Rate gyros do not have their angular momentum vectors fixed in inertial space but are forced to rotate with the body, as shown in Figure 3. A torque must be applied to change the pointing direction of the rotor, and this torque tilts the rotor up or down in proportion to the input rotation angle.

Euler's equation for rigid body motion describes the behavior of the open loop non-rebalanced gyro. Here, M is used for the torque about the output axis, and H is used for the angular momentum of the rotor, which is assumed to have constant magnitude. The dot denotes differentiation with respect to time.

$$M = \dot{H} \tag{2-1}$$

The applied torque M consists of two parts. Viscous damping from the gimbal bearing is equal to the product of the damping coefficient C with the output angular rate $\dot{\theta}$. In addition, torque is applied by the gimbal assembly to change the rotor direction. For small θ, this torque is equal to the product of the input angular rate ω and the rotor momentum H.

$$\omega H \cos(\theta) \approx \omega H \tag{2-2}$$



Source: L. Fallon, III, "Gyroscopes" (Section 6.4), *Spacecraft Attitude Determination and Control*, J. R. Wertz, ed. Dordrecht, Holland: D. Reidel Publishing Company, 1980.

Figure 3. Single Axis Gyro Without Rebalance

The time derivative of the angular momentum is the transverse moment of inertia for the rotor I times the angular acceleration of output angle $\ddot{\theta}$.

$$-C\dot{\theta} + \omega H = I\ddot{\theta} \tag{2-3}$$

Because the system is linear, it is convenient to use Laplace transform notation where "s" indicates differentiation and "1/s" indicates integration. Written in Laplace transform notation, the open loop equation is

$$-sC\theta(s) + H\omega(s) = s^2 I\theta(s) \tag{2-4}$$

This can also be represented in a block diagram, as in Figure 4.

The transfer function of $\theta(s)$ with respect to $\omega(s)$ is

$$\frac{\theta(s)}{\omega(s)} = \frac{H}{sC(1 + s\tau_o)} \tag{2-5}$$

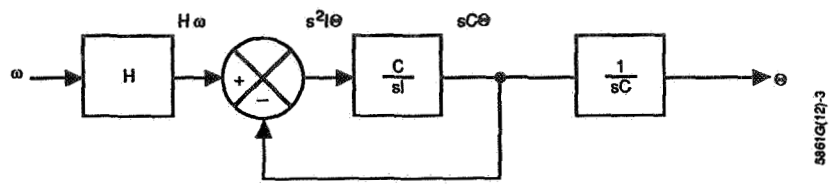


Figure 4. Open Loop Gyro Model

where τ_o is the open loop gyro time constant and, for COBE, is equal to 0.7 x 10^{-3} seconds (Reference 3).

$$\tau_o = I/C = 0.7 \times 10^{-3} \text{ seconds} \tag{2-6}$$

The impulse response or inverse transform for this open loop transfer function is

$$\theta(t) = (1 - \tau_o e^{-t/\tau_o}) \frac{H}{C} \tag{2-7}$$

Because the integration time of 1 second is much longer than τ_o, the second term can be neglected. This corresponds to dropping the inertial term from the original differential equation and the transfer function to give

$$\omega H \approx C\dot{\theta} \tag{2-8}$$

and

$$\frac{\theta(s)}{\omega(s)} \approx \frac{H}{sC} \tag{2-9}$$

The tilt angle θ, however, is not the rotation about the input axis that is sought. To this end, if ω is recognized as the time derivative of the input rotation angle ϕ

$$\omega = \dot{\phi} \tag{2-10}$$

ϕ can be obtained by integrating the differential equation.

$$\phi = \int_0^T \omega \, dt \approx \int_0^T C\dot{\theta}/H \, dt \tag{2-11}$$

If the input rate ω is assumed constant over the 1-second sampling interval, the integral can be replaced by a simple proportional relationship between the input and output angles.

$$\phi \approx C\dot{\theta}/H \tag{2-12}$$

Because this constant of proportionality changes rapidly when the tilt angle is large, most rate gyros have a feedback control loop to rebalance the rotor and keep the tilt small.

2.3 TORQUE REBALANCING—CLOSED LOOP

In torque-rebalanced rate gyros, such as those on COBE, the response of the gyro to an input rotation is damped by a viscous torque, and a restoring torque is applied to return the gyro to the null position in the spacecraft frame. The control loop for rebalancing the gyros has as input the incremental output angle $\Delta\theta$, or the difference between the current tilt of the rotor and the tilt at the previous time step.

This incremental tilt angle $\Delta\theta$ is divided by the time between samples, multiplied by the rotor angular momentum, and is fed back as a torque. The output is actually a pulse-width modulated (PWM) signal in which the pulse width is proportional to the tilt of the gyro from the null position. To calculate the total input angle ϕ through which the spacecraft has moved, the pulses are summed and scaled appropriately.

Adding feedback proportional to the output angle $\Delta\theta$ and making the simplifications discussed above gives the Farrenkopf model shown in Figure 5.

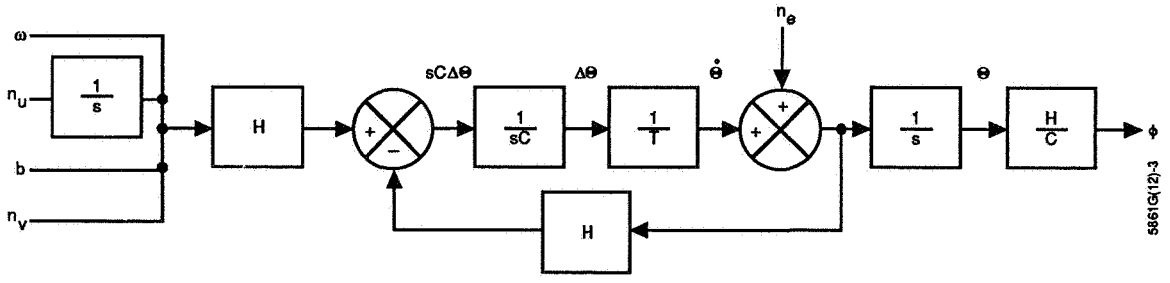


Figure 5. Closed Loop Rate Gyro Model

Ignoring the error sources for the moment, the system equation is

$$H\omega(s) - \frac{H}{T}\Delta\theta(s) = sC\Delta\theta(s) \tag{2-13}$$

As shown in Figure 5, $H\Delta\theta(s)/T$ is subtracted from $H\omega(s)$ leaving $sC\theta(s)$. This quantity is divided by sC to give the incremental angle $\Delta\theta(s)$. Because the feedback signal must be a torque, $\Delta\theta(s)$ is divided by the sampling interval T to produce an average rate $\dot{\theta}(s)$ and is then multiplied by the rotor angular momentum H.

$$\dot{\theta} \equiv \Delta\theta/T \tag{2-14}$$

The resulting transfer function relating the input rate $\omega(s)$ to the output rate $\dot{\theta}(s)$ is

$$\frac{\dot{\theta}(s)}{\omega(s)} = \frac{1}{1 + sCT/H} = \frac{1}{1 + s\tau_c} \tag{2-15}$$

The impulse response for this system is

$$\dot{\theta}(t) = e^{-t/\tau_c} \tag{2-16}$$

where the closed loop time constant τ_c is

$$\tau_c \equiv CT/H \tag{2-17}$$

For COBE, the closed-loop time constant is 0.03 second, which yields a bandwidth of 5 Hz for the rebalanced system. For timescales much longer than the closed-loop time constant τ_c, the rebalanced rate gyro passes the input rate straight through. This can be represented by eliminating the loop from Figure 5.

2.4 ERROR SOURCES AND EXPECTATIONS

The gyro model includes the following four error sources that enter the control loop as either input or output errors (Reference 4):

- Rate bias (b), which is a constant that is imprecisely known
- Float torque noise (n_v), which is white noise on the rate bias
- Float torque derivative noise (n_u), which is the "strength" of the rate bias random walk
- Electronic noise (n_e), which is white noise on the incremental angle θ

Although the rate bias is constant, it is imprecisely known and when used for propagation introduces error. It can be estimated, but it can never be known perfectly. Because it is constant, the bias autocorrelation, which is needed for predicting the variance of the integrated angle ϕ, is also a constant. The variance is obtained by setting the two times λ and τ equal in the expression for the autocorrelation. E[...] is used here as the expectation operator.

$$E[b(\lambda)b(\tau)] = \sigma_b^2 \tag{2-18}$$

The float torque, float torque derivative, and electronic noise sources are zero mean and white, meaning that their value at one time is independent of their value at any other time. This is an idealization and is not true if the sampling is done at a rate above the highest frequency component of the white noise. Although the idealization is used here, it is understood that this applies to the COBE gyro sampling rate of 1 Hz.

The white noise approximation is indicated by the Dirac delta function δ (t), which is zero for all t except zero, and which integrates to 1 over an interval including zero. In addition, these three white noise sources are completely uncorrelated among themselves. This is reflected in the equation below using the Kronecker delta function δ_{ij}, which is zero when i is not equal to j and 1 when it is.

$$E[n_i(\tau)] = 0 \quad (i = v, u, e) \tag{2-19}$$

$$E[n_i(\lambda)\, n_j(\tau)] = \sigma_i^2\, \delta_{ij}\, \delta(\lambda - \tau) \tag{2-20}$$

The names float torque and float torque derivative noise are holdovers from the days when the rotors were floated in a liquid for rebalance. Although the physical interpretation of these sources is different for the COBE dry-tuned gyros, the effect is the same and the traditonal names are preserved.

The transfer function for the summed output angle θ with respect to n_v and b is 1/s times the transfer function for the output angular rate $\dot{\theta}$ derived above. This corresponds to the final integration of the sampled rates to give the total rotation.

$$\frac{\theta(s)}{\omega(s)} = \frac{\theta(s)}{N_v(s)} = \frac{\theta(s)}{B(s)} = \frac{\dot{\theta}(s)}{s\omega(s)} = \frac{1}{s(1 + s\tau_c)} \tag{2-21}$$

The corresponding impulse response is

$$g_v(t) = g_b(t) = 1 - e^{t/\tau_c} \tag{2-22}$$

The transfer function for n_u is 1/s times that for n_v.

$$\frac{\theta(s)}{N_u(s)} = \frac{\theta(s)}{sN_v(s)} = \frac{1}{s^2(1 + \tau_c s)} \tag{2-23}$$

which has the impulse response

$$g_u(t) = \tau_c e^{-t/t_c} + t - \tau_c \tag{2-24}$$

The output transfer function for n_e is

$$\frac{\theta(s)}{N_e(s)} = \frac{\tau_c}{1 + \tau_c s} \tag{2-25}$$

with impulse response

$$g_e(t) = e^{-t/\tau_c} \tag{2-26}$$

The response of the system to each of these inputs is the convolution integral of the input with the appropriate impulse response function. Because the system is linear, the combined response is the sum of the individual responses. This and the fact that the noise sources are independent imply that the variance of the total rotation angle equals the sum of the individual variances.

$$\text{var}(\theta) = \text{var}(\theta_e) + \text{var}(\theta_v) + \text{var}(\theta_b) + \text{var}(\theta_u) \tag{2-27}$$

Since each of the noise sources has zero mean, the expected response is also zero mean. This is because the expectation operator can be brought inside the convolution integral to act upon n_i, as in the example below. The expected value of n_i is zero, and this makes the expected response zero as well.

$$E[\theta_i] = E\left[\int_0^t n_i(\tau)\, g_i(t - \tau)\, d\tau\right] = 0 \tag{2-28}$$

The expected autocorrelation is the expectation of the product of two convolution integrals. Because the expected value is zero, there is no square of the average to subtract. As before, the variance is obtained by evaluating the autocorrelation at a single time t.

$$\text{var}(\theta_i) = E\left[\left(\int_0^t n_i(\lambda)\, g_i(t - \lambda)\, d\lambda\right)\left(\int_0^t n_i(\tau)\, g_i(t - \tau)\, d\tau\right)\right] \tag{2-29}$$

These integrals can be combined and the expectation operator brought inside to give

$$\text{var}\,(\theta_i) = \int_0^t g_i(t - \lambda) \int_0^t E[n_i(\lambda)\, n_i(\tau)]\, g_i(t - \tau)\, d\tau d\lambda \tag{2-30}$$

Formal integration of these expressions using the appropriate autocorrelations for the four error sources yields the following expressions for their contributions.

$$\text{var}(\theta_b) = \sigma_b^2[t - \tau_c\,(1 - e^{-t/\tau_c})]^2 \tag{2-31}$$

$$\text{var}(\theta_v) = \sigma_v^2(t - 3\tau_c/2 + 2e^{-t/\tau_c} - \tau_c e^{-2t/\tau_c}/2) \tag{2-32}$$

$$\text{var}(\theta_u) = \sigma_u^2(t^3/3 - t^2\tau_c + t\tau_c^2 + \tau_c^3 - 2t\tau_c^2\, e^{-t/\tau_c} - t_c^3\, e^{-2t/\tau_c}) \tag{2-33}$$

$$\text{var}(\theta_e) = \sigma_e^2\, \tau_c(1 - e^{-2t/\tau_c}) \tag{2-34}$$

Given that the closed-loop time constant τ_c is 0.03 second and a typical propagation time-span is thousands of seconds long, the total variance is well approximated by a cubic polynomial in time.

$$\text{var}(\theta) \approx \tau_c \sigma_e^2/2 + \sigma_v^2 t + \sigma_b^2 t^2 + \sigma_u^2 t^3/3 \qquad (2\text{-}35)$$

Finally, this variance of the summed gyro output angle should be multiplied by (C/H)2 to convert the variance of θ to the variance of ϕ, the angle through which the spacecraft has rotated. For the COBE gyros, this factor is 1 and so the variances of θ and ϕ are equal.

$$\text{var}(\theta) = \text{var}(\phi) \qquad (2\text{-}36)$$

3. ESTIMATES OF THE GYRO NOISE

3.1 IMPLEMENTATION AND PRACTICE

Because gyro noise estimation was not planned for from the beginning, it is done with the existing COBE software, as shown in Figure 6. As a result, instead of using sensor observations directly, the fine attitude determination subsystem (FADS) gyro-propagated attitude is used as the observation and the coarse attitude determination subsystem (CADS) nongyro single-frame solution is used as a reference. The two attitudes are converted to their Euler angle representations in the quality assurance subsystem (QA), and the squared difference in the yaw angle serves as a sample of the propagation error variance at that time.

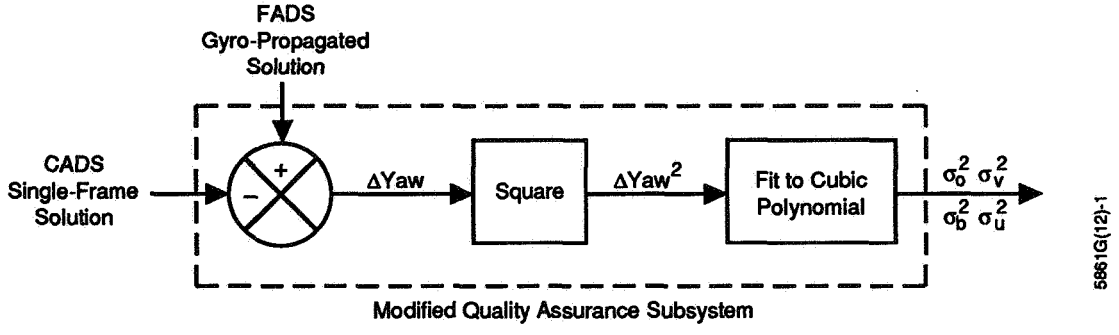


Figure 6. Spin Gyro Noise Estimator

The variation of this sampled variance is fit to a cubic polynomial function of time, in keeping with the model developed in the last section. The one difference is that due to the large single-frame solution noise, the constant term $\tau_c \sigma_e^2$ is more a reflection of this

single-frame error than of electronic signal generator noise as in the Farrenkopf paper (Reference 2). For this reason, the constant variance is referred to hereafter as σ_0^2.

One problem with this approach is that, due to the 4-degree roll of the spacecraft, the spin axis gyro also picks up some of the pitch motion. If gyro noise estimation had been planned for before the ground support software was written, it could have been implemented more cleanly.

Gyro noise is usually measured by putting the gyro on a turntable and comparing the integrated displacement angle with the measured angle (Reference 5). For accuracy, the test can be conducted over many days. Although not impossible in flight, such long tests are awkward since they require large data sets and long processing times. Instead, shorter timespans are employed and the results averaged to improve accuracy.

Averaging helps more in finding σ_v than it does in finding σ_u. This is because the variance from σ_v grows linearly with time, while that from σ_u grows as the third power of time. One-tenth the propagation time provides one-tenth the variance and therefore one-tenth the information about σ_v. Repeating the short test 10 times provides the same information as one long test. For σ_u, however, one-tenth the time provides one-thousandth the information. Repeating the test over 10 intervals provides only a small portion of the information about σ_u that the longer run provides.

As described above, gyro noise is assumed to cause growing errors in the propagation, which are seen by comparing the gyro-propagated solutions with the single-frame solutions. The single-frame solutions are themselves noisy and susceptible to large localized errors, as shown in Figure 7. Thus, long timespans covering several orbits are necessary. As long as the large errors are constant from orbit to orbit, the gradual increase in error due to gyro noise is still discernable.

Another problem arises because of the least squares optimization criterion used in producing the gyro-propagated attitude. Because two medium-size errors at the ends of a batch are preferable to no error at one end and a large error at the other, the minimum propagation error is apt to be at the center rather than the start of the batch. This is also visible in Figure 7.

Instead of fitting one monotonically increasing polynomial to the entire timespan, the batches are broken into halves. In finding coefficients for the first half of the data, time is thought of as going backward from the middle to the start of the batch. This corrects the signs of the odd power coefficients, which would otherwise have the wrong sign. An alternative to this would be to propagate from a single-frame solution. In this case, the constant term σ_0^2 would be set to zero.

3.2 RESULTS AND COMPARISON WITH SPECIFICATIONS

Few results have been collected as of this time. The following coefficients come from the example of Figure 7. They represent the estimated noise from a 5-hour, 38-minute timespan.

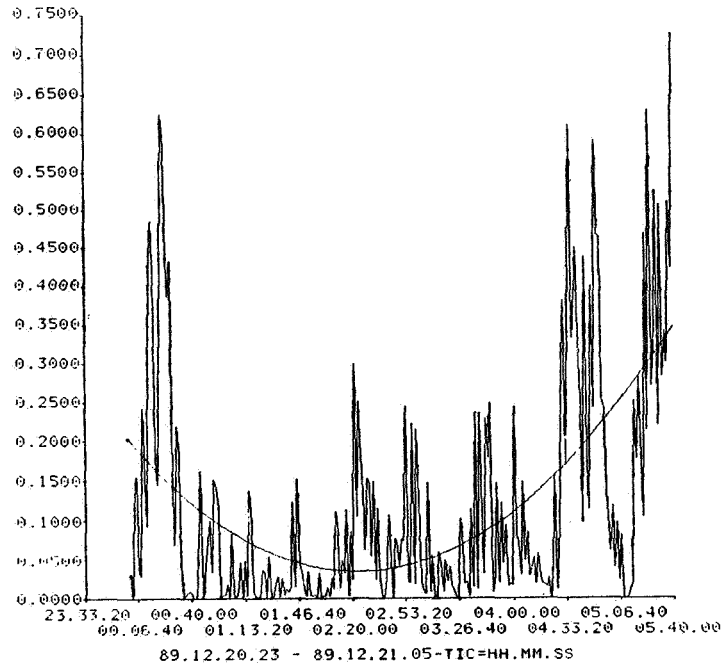


Figure 7. Typical Yaw Differences and Their Cubic Fit

Single Frame Noise	σ_0^2	$0.1464\ \text{deg}^2$
Float Torque Noise	σ_v^2	$0.2196 \times 10^{-4}\ \text{deg}^2/\text{s}$
Gyro Bias	σ_b^2	$0.4006 \times 10^{-8}\ \text{deg}^2/\text{s}^2$
Float Torque Derivative Noise	σ_u^2	$-0.4458 \times 10^{-12}\ \text{deg}^2/\text{s}^3$

The first number represents the combined variance of the batch and single-frame yaw solutions, and appears to be dominated by the large spikes in the yaw differences. The second coefficient represents the white noise on the gyro rate. For samples taken every second, this implies an error of 0.005 degree per second$^{1/2}$. The 24 arc-second telemetry quantization explains 0.002 degree per second$^{1/2}$ of that number, and the rest comes from other sources. The third coefficient is the variance on the assumed gyro bias and is 0.2 degree per hour, which is about the accuracy of the typical bias solution. Finally, the fourth coefficient is negative, indicating that it is not well estimated over this timespan. A longer batch may be necessary to observe it.

Expected worst case values of σ_v^2 and σ_u^2 can be backed out of the following COBE specifications (Reference 1).

- Root mean square gyro drift rate noise error will be less than 0.002 degree per second
- Gyro rates will be sampled at 1-second intervals (science format at 4 kilobits per second)

- X-gyro telemetry quantization will be 24 arc-seconds per count, sampled once per second
- X-gyro scale factor will be stable to 50 parts per million over 30 minutes
- Nominal spin rate will be 4.8 degrees per second

If the drift error is assumed to be the result of integrating the float torque noise over the 1-second sampling interval, it implies that the inherent float torque noise variance is 0.4 x 10^{-5} deg^2/s. The telemetry quantization contributes the same amount to the effective float torque noise, thus doubling its strength to 0.8 x 10^{-5} deg^2/s.

To estimate the float torque derivative noise, the scale factor stability is multiplied by the spin rate to give the spin rate stability. Assuming this error to be the standard deviation of the rate at the end of the specified 30-minute period, the float torque derivative noise would have to be 0.3 x 10^{-11} deg^2/s^3.

Float Torque Noise	σ_v^2	0.8×10^{-5} deg^2/s
Float Torque Derivative Noise	σ_u^2	0.3×10^{-11} deg^2/s^3

The estimated value of σ_v^2 is 2-1/2 times that specified, but more sample propagations are needed before the gyro performance is questioned. Besides, there may be other sources of error outside the gyro that may behave like the float torque noise. The estimate of σ_u^2, while not valid because of the sign, is the right order of magnitude. It is likely that the scale factor stability and other float torque derivative noise sources are well within specifications and are too small to be observed over a 5-hour, 38-minute propagation.

3.3 USING THE RESULTS

For a Kalman filter that estimates the attitude and rate bias, the noise standard deviations can be used to compute the state noise covariance matrix (Q), which is added to the state covariance matrix (P) after each propagation step (T) (Reference 1). This controls how quickly the old state estimate is forgotten.

$$Q = \begin{bmatrix} \sigma_v^2 T + \sigma_u^2 T^3/3 & -\sigma_u^2 T^2/2 \\ -\sigma_u^2 T^2/2 & \sigma_u^2 T \end{bmatrix} \qquad (3\text{-}1)$$

For the batch estimator, the only way to control propagation error is to limit the batch length τ_B. The estimator can be modeled as an averaging operation and incorporated into a simplified random process model as shown in Figure 8.

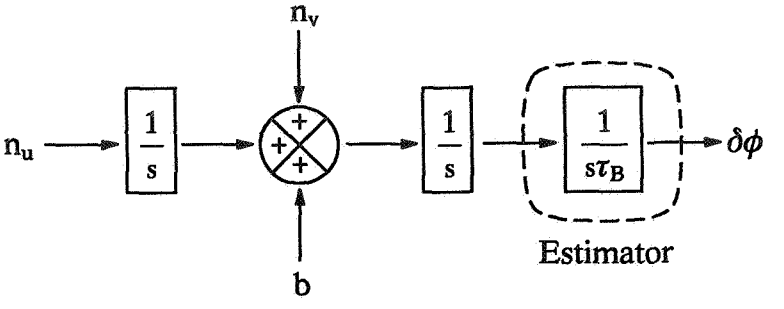


Figure 8. Propagation Error and Batch Estimation

If the observations have variance σ_0^2 and come at a rate of k per second, the epoch error $\delta\phi$ has the following variance:

$$\text{var}\,(\delta\phi) = \frac{\sigma_0^2}{k\,\tau_B} + \frac{\sigma_v^2\,\tau_B}{3} + \frac{\sigma_b^2\,\tau_B^2}{4} + \frac{\sigma_u^2\,\tau_B^3}{20} \qquad (3\text{-}2)$$

For the COBE float torque noise estimate and azimuth observations with variance 0.02 deg^2 made every 0.5 sec, the epoch solution variance is minimal when the batch is 55 seconds long. This criterion ignores the inconvenience of doing many short batch solutions. A more realistic way to use this estimate is to choose batch length to keep the epoch solution variance within acceptable limits. For the COBE definitive accuracy requirement of 3 arc-minutes (3σ), this restricts batches to 70 sec. For routine attitude determination where the requirement is 1 deg (3σ), batches of up to 50 minutes are acceptable.

4. HOW NOISE ESTIMATION COULD BE DONE BETTER

4.1 THE DIFFICULTY OF OBSERVING σ_u^2

To observe σ_u, one must propagate for a long time. The current batch estimator, which stores the data for the entire batch in memory, makes this very expensive. Skipping points permits longer batches but may introduce additional error that is not really in the gyro data. A better way would be to propagate open loop and estimate the noise parameters as one goes along. In this way, one could process long timespans without as much overhead.

4.2 A FILTER IMPLEMENTATION

One solution is to incorporate the noise estimation in a Kalman filter, as shown in Figure 9. This would facilitate long propagations, because the end state could be used as the

starting point for the next timespan. This could also be done with a modified COBE batch estimator if it were operated so as to produce a continuous attitude history. In either case, the partial derivatives of the observations with respect to the noise parameters would also have to be passed in with the attitude.

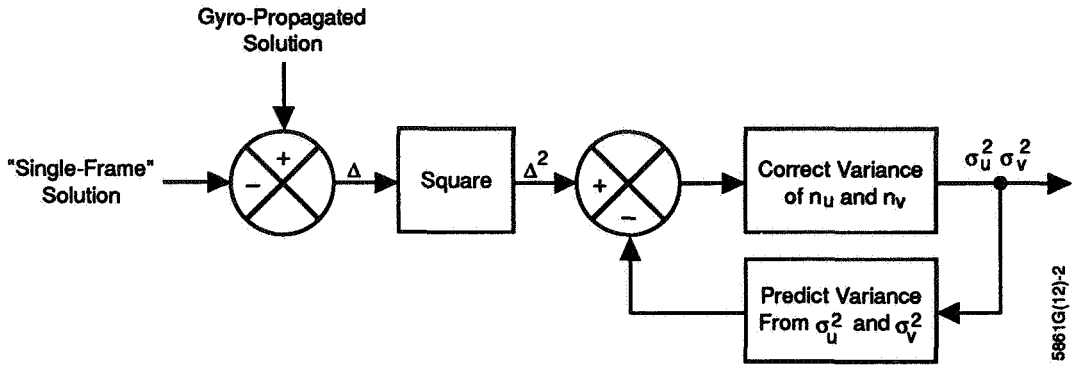


Figure 9. Filter Noise Estimator

The observations of the gyro noise source variances would come from the propagated attitude (a_c). The reference would be the attitude estimated from sensor observations (a). The noise model would be the same as that used earlier.

$$[a - a_c]_i^2 = \sigma_0^2 + \sigma_v^2 t_i + \sigma_b^2 t_i^2 + \sigma_u^2 t_i^3/3 \qquad (4\text{-}1)$$

Ideally, the estimated attitude would be computed independently of the gyro rates. This is no problem for COBE, which has continuous attitude observability. For other spacecraft, one might have to settle for fewer observations of the propagation error due to the lack of a complete reference attitude. Alternatively, one might accept use of the gyros to produce the reference attitude as long as the state noise covariance (Q) was large enough that old observations were forgotten quickly.

The approach could also be extended to all three axes if a model could be developed for the contribution of the noise on each gyro axis to each propagated angle. Those contributions would depend on the actual angular velocity and so would have to be computed numerically over the duration of the propagation, just as the epoch-to-current-attitude propagation matrix and gyro bias variational matrix are computed now in the batch estimator.

5. CONCLUSIONS

The method presented here for estimating gyro noise levels has been derived from a simple but theoretically based gyro model and has been shown to provide reasonable

results. Although these results are quite believable, this single sample should not be given too much weight. Many such runs, or even longer ones, should be made and the results averaged before being accepted. After all, ground tests of the gyros last for days, and at least this much time is needed in flight, where the reference attitude is so much less accurate. To this end, the filter version of the noise estimator should be considered.

The next step in estimating noise should be to look at the other manifestations of noise, such as the drifting of gyro biases over the course of days and weeks and the spectral decomposition of the gyro rates themselves. The former is a comparatively long timescale variation, which should be compatible with the shorter effects examined here. The latter is a shorter timescale approach, which should also agree with the present estimates of the gyro noise.

Perhaps even more important than continuing to study the problem is putting these noise estimates to use in tuning Kalman filters and setting batch lengths for more enlightened attitude determination. If these results prove helpful in this regard, a new system that estimates gyro noise along with attitude could be built that would provide more information about gyro performance more conveniently than is now possible.

REFERENCES

1. Computer Sciences Corporation, CSC/TM-89/6007, *COBE Attitude Error Analysis and Calibration Plan*, D. Chu and T. Z. Willard, August 1989

2. R. L. Farrenkopf, *Generalized Results for Precision Attitude Reference Systems Using Gyros*, AIAA Paper 74-903, AIAA Mechanics and Control of Flight Conference, Anaheim, CA, August 1974

3. M. Femiano, COBE Rate-Measuring Assembly CDR—Electrical Design, National Aeronautics and Space Administration, Goddard Space Flight Center, Code 700, 6 May 1985

4. F. L. Markley, *Modeling Gyro Noise*, GSFC Code 550 lecture notes, 18 February 1986

5. The Analytical Science Corporation, *Applied Optimal Estimation*, edited by A. Gelb. Cambridge, MA: MIT Press, 1974

Fixed-Head Star Tracker Magnitude Calibration on the Solar Maximum Mission

D.S. Pitone, B.J. Twambly, A.H. Eudell, and D.A. Roberts
Computer Sciences Corporation (CSC)

ABSTRACT

The sensitivity of the fixed-head star trackers (FHSTs) on the Solar Maximum Mission (SMM) is defined as the accuracy of the electronic response to the magnitude of a star in the sensor field-of-view, which is measured as intensity in volts. To identify stars during attitude determination and control processes, a transformation equation is required to convert from star intensity in volts to units of magnitude and vice versa. To maintain high accuracy standards, this transformation is calibrated frequently. A sensitivity index is defined as the observed intensity in volts divided by the predicted intensity in volts; thus, the sensitivity index is a measure of the accuracy of the calibration. Using the sensitivity index, analysis is presented that compares the strengths and weaknesses of two possible transformation equations. The effect on the transformation equations of variables, such as position in the sensor field-of-view, star color, and star magnitude, is investigated. In addition, results are given that evaluate the aging process of each sensor. The results in this work can be used by future missions as an aid to employing data from star cameras as effectively as possible.

1. INTRODUCTION

From the time of the repair of the Solar Maximum Mission (SMM) in April 1984 to its reentry in December 1989, an enormous amount of attitude sensor data were collected. The fixed-head star trackers (FHSTs) on the SMM during this time primarily collected two types of data. The first, star position in the FHST field-of-view (FOV), provided after transformation by the FHST alignment matrix, observed star position vectors in the body frame for SMM attitude determination. The second, star intensity data, were used for star identification purposes. Star intensity is the FHST measurement of the magnitude of the star. This paper is concerned with this second type of data, primarily what is the most effective way to calibrate the FHSTs' electronic response to stellar magnitude and how may the FHST magnitude data be used to give some insight into how the electro-optical components of the FHSTs aged over the lifetime of the SMM. These data, in turn, may then be used to aid analysis for future missions that employ similar types of star cameras for attitude determination and control.

For the SMM, visual stellar magnitude was used for star identification. This identification was necessary for both attitude determination and attitude control of the spacecraft. For attitude determination, the observed intensity was taken from the raw FHST data, converted to visual magnitude using a calibrated transformation equation, and then used

to help identify the observed star. Using the identity of the observed star, the corresponding representation in Geocentric Inertial (GCI) coordinates from the star catalog are used to form a reference vector. The reference vector and the observed vector are then used along with other pairs of reference and observed vectors in an algorithm to determine the attitude.

Stellar magnitude was used for attitude control to maintain the current attitude of the spacecraft. Using ground based software, stars useful for control at specific attitudes and time periods were selected from the star catalog. The expected positions of these stars in the FHST FOV and their predicted magnitudes were then uplinked to the spacecraft. The catalog magnitude was transformed to intensity in volts, again using the calibrated transformation equation. The FHST then searched for these stars, called guide stars, at the predicted coordinates in the FOV. If a star was observed within a specified angular distance, usually one tenth of a degree (deg), of the predicted location, its intensity was compared with the predicted intensity. If they matched, again within a specified tolerance, it was assumed that the correct guide star had been located. Discrepancies in the position of the guide star in the FHST FOV were then attributed to error in the spacecraft attitude, and the spacecraft would be rotated to null the discrepancy. Thus, the spacecraft attitude was maintained.

The preceding paragraphs show that stellar magnitude is important for accurate spacecraft attitude determination and control. Thus, the equation that related magnitude and intensity was calibrated periodically to maintain an accurate transformation.

This work has two major goals. The first is to evaluate two possible transformation equations for accuracy and ease of application. In addition, the various parameters that effect the optical sensitivity of the FHST will be examined and correlated to sensitivity. The second goal is to establish a method for investigating the aging process of star cameras. It is apparent that as the sensor ages, it will become less sensitive. Methods for examining the rate of aging are presented along with results for the SMM FHSTs.

The paper is organized as follows: Section 2 provides a brief history of the mission, along with the basic configuration of the attitude determination and control system of the spacecraft. Section 3 includes a description of the FHST hardware aboard the SMM and the mathematical models used to reduce the FHST data. Section 4 outlines the analysis performed to calibrate the intensity-to-magnitude transformation equation, including an investigation of two possible transformation equations and the actual calibration results. Section 5 considers how the results of the calibrations may be used to gain insight into the aging process of the electro-optical components of the FHST. This provides information about the duration of reliability these type of star cameras have in space.

2. THE SOLAR MAXIMUM MISSION

Mission History

The SMM was launched in February 1980 from the Eastern Test Range at Kennedy Space Center into an approximately circular low-Earth orbit, with an inclination near 28 deg. The scientific objective of the mission was the study of solar phenomena. The spacecraft attitude system provided three-axis stabilization and supported solar feature

targeting. The spacecraft functioned normally until November 1980, when the standard reaction wheel (SRW) package that provided the controlling torques to the spacecraft began to fail. To preserve the mission, the SMM was put into a spin (approximately 1 deg per second (sec)) about the minor principal axis. During April 1984, the spacecraft was repaired in-orbit as part of a Space Transportation System (STS) mission. The entire attitude control system was replaced, and the spacecraft was returned to the nominal scientific observing mode. The spacecraft functioned normally until August 1987, when one of the two Ball Aerospace CT401 FHSTs sustained a failure in the power supply to the sensor. Since the spacecraft had two FHSTs, this loss had only a minimal effect on the attitude determination and control of the spacecraft. On December 2, 1989, the SMM reentered the Earth's atmosphere. The surviving remnants crashed into the Indian Ocean. Further details on the history of the SMM are available in Ref. 1.

Mission Configuration

The SMM was the first of the multimission modular spacecraft (MMS) series that were modular to facilitate mission repair and adaptation. The SMM basically consisted of two parts: the MMS itself and the experiment module. The components that came with the MMS series were a communications module, a power module, and a modular attitude control system (MACS). The SMM FHSTs were located in the MACS. The experiment module contained all the SMM mission-specific components, including the SMM scientific instruments and the mission-specific attitude sensors. The instruments comprising the scientific payload were designed to study the Sun at several different wavelengths of light, including gamma, X-ray, and ultraviolet.

3. SMM FHST DESCRIPTION

Hardware Description

The SMM FHSTs were National Aeronautics and Space Administration (NASA) Standard Star Trackers (SST), manufactured by Ball Aerospace Systems Division. An SST is an electro-optical instrument that uses an image dissector to search for and track stars[2]. It provides star position and magnitude information about a two-axis coordinate system in an eight-by-eight deg FOV. The accuracy of the two-axis position information is 10 arc-sec (1σ) when fully calibrated and is output as counts. The transformation equation from counts to degrees is calibrated for star temperature and intensity, and the magnetic field.

The SST is protected against bright light by a shutter. When the bright object alert sensor (BOAS) is triggered, the shutter closes, protecting the SST. When the BOAS senses that the bright object is out of the FOV of the SST, the shutter is reopened.

After the shutter is opened, the SST starts out in a search mode. The instrument will search the total FOV by scanning, as shown in Fig. 1. This scan takes approximately 10 sec to search the complete FOV. The search scan continues until a star is sensed. At this point the SST enters the track scan mode. The track scan mode forms a smaller cross-pattern that repeats approximately 100 times per sec. This pattern is also shown in Fig. 1. The track scan mode provides many observations of the same star. A threshold is set that

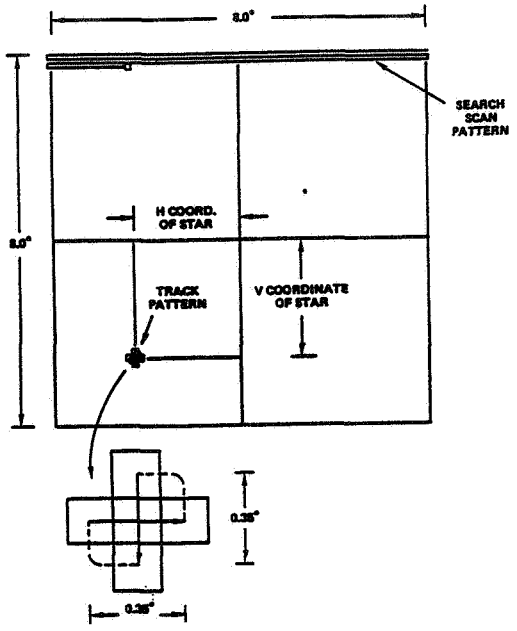


Figure 1. FHST Field-of-View

provides a break-track command that forces the SST back to the search mode. Thus, the SST operates by searching for a star, entering track mode, providing many observations of that star, and then reentering search mode until another star is found. This pattern is repeated until the BOAS is triggered, and the shutter is closed.

When a star is tracked, light from the star enters the lens. The image is focused on the photocathode of the image dissector tube (IDT). The optical image is converted to an electron image at the photocathode, and this image is refocused at the IDT aperture plate. A small aperture passes a portion of the electron image to the electron multiplier. The signal is multiplied and sent to the video processing electronics where it is demodulated. At this point, star intensity and position information are obtained. Further details on the hardware of the SMM FHSTs are available in Ref. 2.

Mathematical Description

Each FHST provided star position about two sensor axes and star magnitude information to the SMM onboard computer (OBC), where it was time tagged and subsequently telemetered to the ground. Each star observation was telemetered as three pieces of data. The first piece was the angle from the first sensor coordinate axis for a specific star at the time tagged by the OBC. The second piece of data was the angle from the second sensor coordinate axis for the same star at the same time. The sensed stellar intensity of the

observation, the third piece of data, had units of voltage.

For use in attitude determination and control algorithms, the raw FHST data are converted to usable quantities. As shown in Fig. 1, the SMM FHSTs had two output axes, horizontal and vertical. The first piece of position data was the angular length of the observed star vector's projection onto the horizontal axis, labeled the H coordinate. The second piece of data was the angular length of the observed star vector's projection onto the vertical axis, labeled the V coordinate. The H and V coordinates were then calibrated for the effects mentioned earlier and transformed to degrees[3]. From the H and V coordinates, a vector representing the direction of the observed star in the FHST sensor coordinate frame was formed. This vector was then transformed to the body frame by the FHST alignment matrix for use in the attitude determination algorithm.

The magnitude data were reduced by simply transforming the intensity in volts to magnitude. There are several general equations to accomplish this. Much of the remainder of this paper is concerned with which of two possible candidates was thought to be best for the SMM, i.e., provided the most usable results.

4. ANALYSIS

Data Collection

Because of the large volume of FHST data collected during the flight of the SMM, data reduction was necessary for presentability. It was assumed that the processes that changed the sensitivity were not instantaneous, and, in practice, it was found that large changes took 6 months to 1 year to occur. In addition, about 3 months of data needed to be collected to provide meaningful results. Thus, the periods of data collection were defined as presented in Table 1. The selection of the exact time periods during the year depended on the availability of data.

Table 1. FHST Data Interval Timespans

FHST 1		**FHST 2**	
October - December	1984	October - December	1984
July - September	1985	July - September	1985
January - March	1986	June - August	1986
July - September	1987	June - August	1987
July - September	1988		
March - May	1989		

The exact nature of the data collected during the timespans was chosen to facilitate the analysis in each of the goals of this work. From the FHST telemetry, each observation was tagged with the observation time, FHST ID, horizontal and vertical coordinates, and observed intensity. From the SKYMAP star catalog[4], the star identification number, predicted magnitude, and star color represented as the difference in blue and visual magnitude were collected for each observation.

Mathematical Definitions

To gauge the accuracy of the transformation from visual magnitude to FHST intensity the ratio of observed, I_o, to predicted intensity, I_p, is used. Thus the sensitivity value, S, is defined to be

$$S \equiv I_o / I_p \quad . \tag{3-1}$$

Observed intensities were provided directly from telemetry as voltages. Predicted intensities were calculated by one of two transformation equations. The more accurate the transformation equation, the closer the sensitivity value would be to 1 for an observation.

To correlate the calculated sensitivity values with position, each FHST FOV was divided into 0.2 × 0.2 deg squares. For each observation, a sensitivity value was calculated and, using the observed H and V coordinates, associated with a 0.2 × 0.2 deg square. All the sensitivity values for a particular timespan and square were then averaged to obtain one sensitivity value for each square.

Evaluation of Two Transformation Equations

Although there are many ways the transformation from magnitude to intensity could be represented, this work will consider two that show considerable promise but for different reasons. The first is a simple conversion from intensity to magnitude based on the logarithmic scale of magnitude. The logarithm of the intensity is multiplied by a calibrated scale factor and added to a calibrated bias. Thus,

$$\mathrm{M} = A * \log_{10}(I_p) + B \quad , \tag{3-2}$$

where M is the visual magnitude from the SKYMAP catalog, and A and B are the calibration coefficients. This equation lacks any term for star color dependence or FOV position. The second equation was recommended for use on SMM before launch[5]. In fact, this analysis is a result of analysis recommended in that report. The equation has the form

$$\mathrm{M} = A_o + A_1 H + A_2 V + A_3 H^2 + A_4 HV + A_5 V^2 - 2.5 \log_{10}(I_p) \quad , \tag{3-3}$$

where H and V are the FHST coordinates and the A_i's are defined as

$$A_i = \alpha_{i0} + \alpha_{i1}(B-V) + \alpha_{i2}(B-V)^2 \quad , \tag{3-4}$$

where $B - V$ represents the star color and the α_{ij}'s are the calibration coefficients.

During the initial FHST data processing, Equation (3-2) was used to reduce the data because it was thought at the time that it would provide adequate results while greatly easing implementation. Thus, this assumption may be tested using the actual data.

To perform the tests, the calibration coefficients for each equation were calculated for each segment of data in Table 1. Since Equation (3-2) was used throughout the mission, its calibration coefficients were available directly from the SMM archives. The coefficients in Equation (3-4) were calculated using a small sample of data before each 3-month data segment. Then the observations in each segment were reduced using the equations, and the results were graphed and assembled in tables.

Fig. 2 through Fig. 5 show typical results for each equation. In the figures, the small boxes are the 0.2 × 0.2 deg data sample squares. If the box is empty, the sensitivity value is between 0.95 and 1.05, i.e., it is within 5 percent of the expected value. If the box is filled, then the sensitivity value is either less than 0.95 or greater than 1.05. Thus, dark areas represent the areas of an FHST where stellar magnitude will be sensed less accurately. Areas of the FOV with no sample boxes were areas with no data during that timespan.

The results in these figures show the superiority of the polynomial transformation. Fig. 2 represents the simple transformation for the July through September 1988 time segment. The large area of poor sensitivity in the -V half of the FOV, especially in the -V,+H quadrant, should be noted. Fig. 3 is the same data fit with the polynomial transformation. The reduction in the number of boxes with poor sensitivity values and the lack of any concentrated area with poor sensitivity should also be noted. Figs. 4 and 5 show a similar improvement in FHST2 between the simple and polynomial transformations.

These improvements are expected because the polynomial transformation fully accounts for all the factors. We now address the question of whether the improvement is worth the effort and computer memory needed to obtain it?

Further analysis of the comparison and an examination of the errors that result from poor FHST sensitivity can resolve this question. The complete comparison results are provided in Tables 2 and 3. The columns entitled Simple and Polynomial provide the important numbers for the comparison study. The column entitled Original contains the data that illustrates FHST aging, which is addressed later. Table 2 presents comparison data for the total FOV of each data segment while Table 3 is confined to the -V,+H quadrant of FHST1. The numbers in the columns represent the percentage of the total number of observed sample boxes that have sensitivity within 10 percent of unity. The data in both tables show that the polynomial transformation yields significantly better results.

The type of error that can occur from a lack of sensitivity is a misidentification of a star. For the SMM, this misidentification mainly affected the attitude control. If a star was misidentified during the attitude determination process, the observation was simply discarded. Since there were usually many observations, discarding one observation did not significantly affect the attitude determination. However, the SMM used guide stars to control the spacecraft. An identification error could take two forms. First, the wrong star could be identified as the guide star. Thus, the spacecraft would correct its attitude by rotating to place the misidentified star in the predicted position in the FHST FOV. This would cause the attitude of the spacecraft to be in error. Adding to this error would be the fact that the OBC would indicate the attitude was correct because it believed it had identified the correct guide star.

The second type of identification error occurred when the guide star could not be located by the search pattern. This happened because the guide star intensity was so poorly predicted; no observed star whose intensity matched the prediction could be found. Thus, the attitude of the spacecraft would not be corrected and an error would accrue.

Because of the high accuracy requirements of the SMM, both of these situations would cause significant errors in the roll attitude of the spacecraft. However, only in several isolated incidents was the roll error higher than the 0.1 deg roll error limit. At these times, the roll attitude was computed by the ground support system, and the spacecraft was

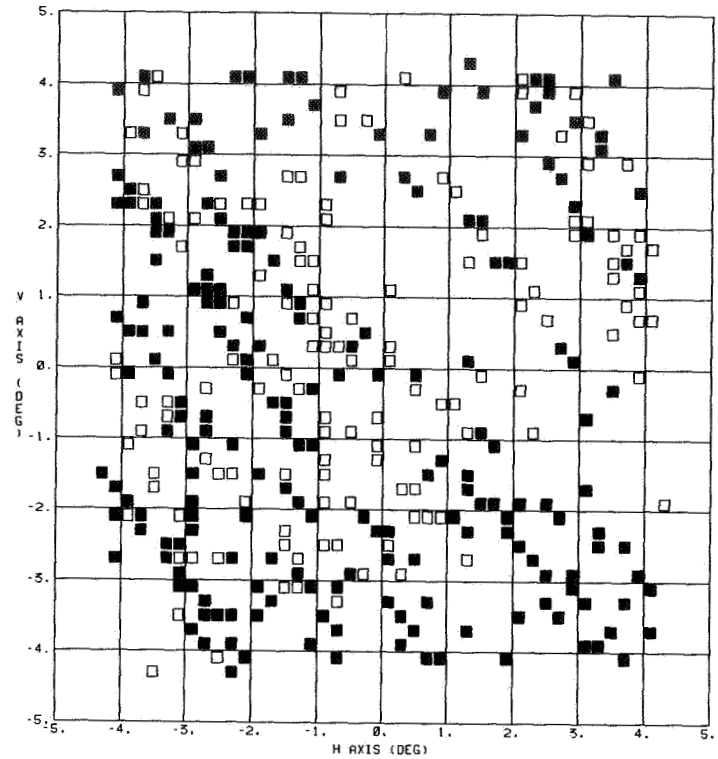


Figure 2. FHST1 Sensitivity Index Plot Using Simple Transformation, 1988 Data

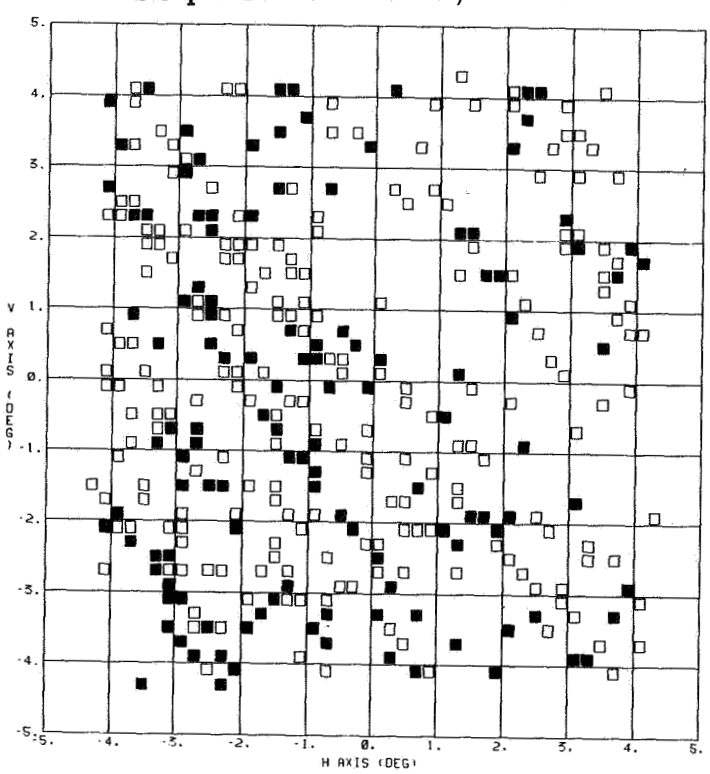


Figure 3. FHST1 Sensitivity Index Plot Using Polynomial Transformation, 1988 Data

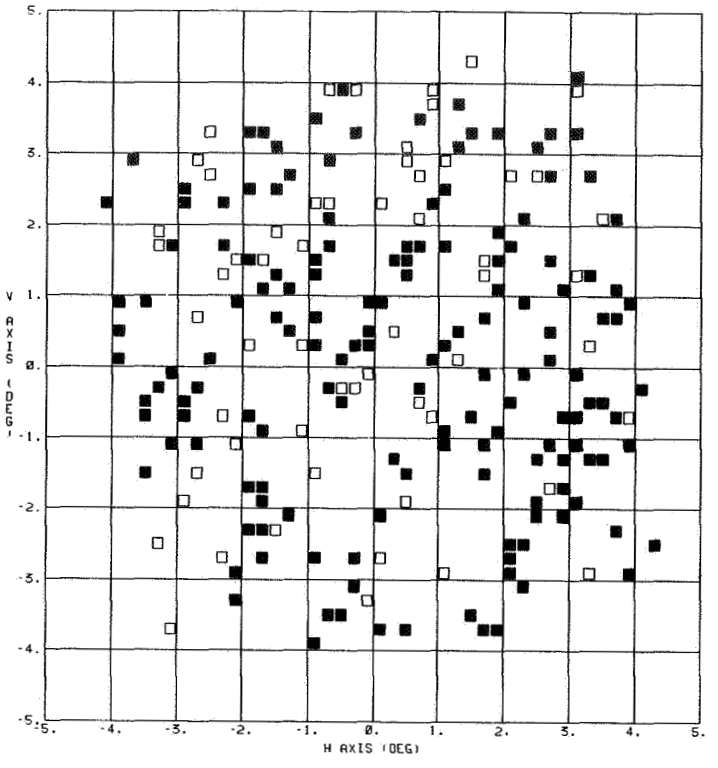


Figure 4. FHST2 Sensitivity Index Plot Using Simple Transformation, 1984 Data

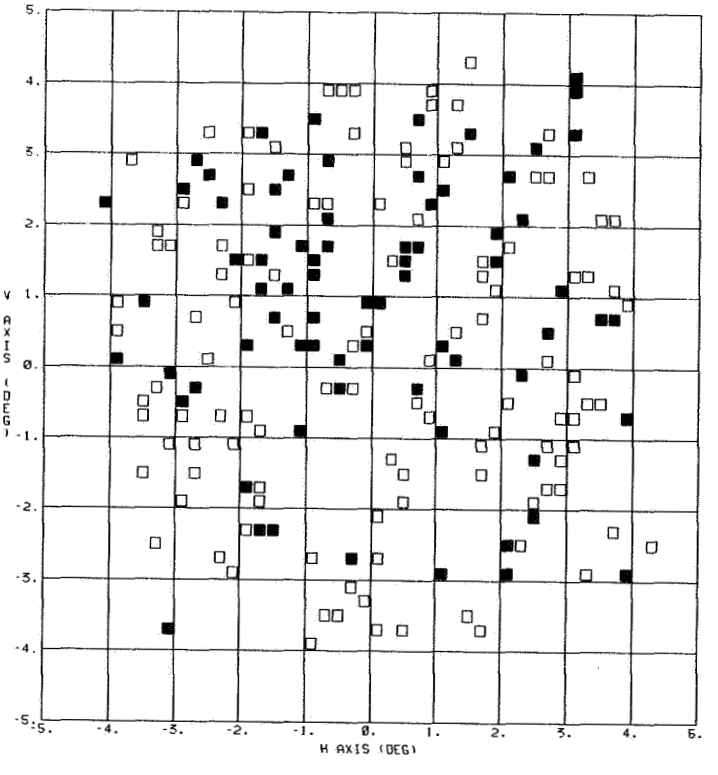


Figure 5. FHST2 Sensitivity Index Plot Using Polynomial Transformation, 1984 Data

Table 2. FHST Sensitivity Values Within 10 percent of Unity

		Original	Simple	Polynomial
FHST1				
October - December	1984	72	72	78
July - September	1985	69	62	73
January - March	1986	57	45	52
July - September	1987	48	64	77
July - September	1988	34	66	90
March - May	1989	29	62	77
FHST2				
October - December	1984	47	45	91
July - September	1985	39	62	85
June - August	1986	27	51	71
June - August	1987	33	65	81

Table 3. FHST1 Sensitivity Values Within 10 Percent of Unity for the H+, V- Quadrant

		Original	Simple	Polynomial
FHST1				
July - September	1985	72	53	81
January - March	1986	18	36	59
July - September	1987	24	41	73
July - September	1988	8	61	91
March - May	1989	22	56	75

manually commanded back to the correct attitude.

Some future missions, most notably the Upper Atmosphere Research Satellite (UARS), will contain an onboard star catalog. Thus, another type of error may occur that did not affect the SMM. Since UARS will be computing onboard attitudes using the star catalog and the intensity-to-magnitude equation, accurate sensing of the correct star magnitude is very important. Thus, misidentification of stars, especially when the FHSTs are observing sparse star fields, will cause an error in the determined attitude.

Thus, it is recommended here that missions with high accuracy requirements should use the polynomial transformation. However, for missions where the requirements are not as stringent, the extra accuracy given by the polynomial fit does not justify the extra effort and memory required.

Affect on Magnitude Response of the Individual Terms

To ascertain why the polynomial transformation is more accurate than the simple transformation, the individual components of the polynomial transformation will be examined. Terms that depend on FOV position and star color are unique to the polynomial

transformation whereas both transformations contain a dependence on the stellar magnitude of the star.

The first term evaluated is the correction for the position coordinates. The analysis in the previous section showed that areas of poor sensitivity were removed by including the position coordinates in the polynomial transformation. However, a quantitative examination of this improvement remains to be performed. Since it contains the most observations, the segment of data from March through May of 1989 will be examined. The polynomial transformation was altered so that only terms accounting for magnitude and star position were included. Table 4 shows that the percentage of good observations improved from 62 percent, using the simple transformation, to 64 percent, using the polynomial transformation with only position correction. However, the complete polynomial transformation improved the results to 77 percent. Thus, the correction due only to position is a small part of the improvement seen using the polynomial transformation.

The polynomial transformation was then adjusted to include only the star color. Table 4 shows that the percentage of good observations improved from 62 percent to 73 percent. Thus, most of the improvement from using the polynomial transformation is due to the star color terms in the equation.

Table 4. FHST1 Sensitivity Values Within 10 Percent of Unity for Star Color or Position Only

			Whole Field
FHST1			
March - May	1989	Star Color only	73
		Position only	64

Fig. 6 shows the dependence of the sensitivity on color. The SMM used a color value known as the blue minus visual magnitude index. This represents the blueness of the star. The more positive the value of the $B - V$, the bluer the star. Fig. 6 shows that the bluer stars, represented by the more positive values of $B - V$, have generally higher values for the sensitivity index. This indicates that the intensity of bluer stars is sensed more accurately by the FHST. Since the simple transformation equation was used to compute the predicted values in the figure, it is expected that more of the bluer star observations will have sensitivity values greater than 1. This is the case as shown in Fig. 6.

The results of the star color analysis have important consequences. The important point to remember is that the NASA standard FHSTs sense the intensity of bluer stars more accurately. The results presented here confirm the analysis done previously by Lorenz[6] and Neste[7]. Lorenz (Ref. 6) first discovered the effect of star color on FHST observations in his analysis of the intensity variations seen by the Landsat-4 FHST observations. Neste (Ref. 7) adapted that work for use on the UARS mission and even presented a quantitative result. The resulting equation computed a correction to the star magnitude based on the star color.

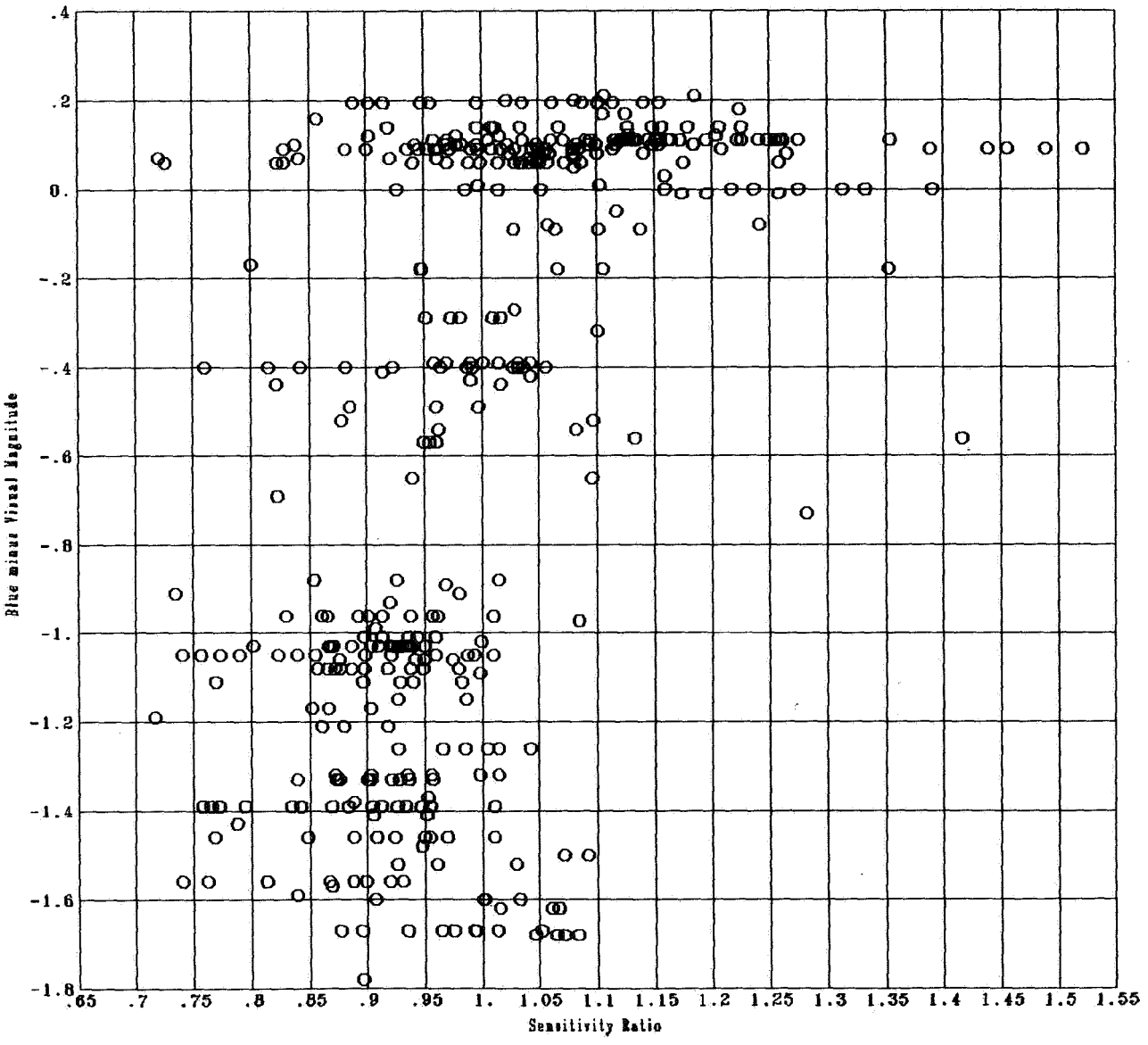


Figure 6. Dependence of Sensitivity Index on Color

5. FHST AGING

As with any electro-mechanical device exposed to the space environment, the SMM FHSTs suffered performance degradation in response to problems that may be associated with aging. Aging causes changes in the sensitivity of the sensor that require calibration to correct the observed intensity. However, if the simple transformation equation is used without recalibration on each segment of data, the resulting measurement degradation of the later segments can provide clues about the aging processes of star cameras in general.

For the evaluation of the aging process, each FHST was calibrated using the October through December 1984 data. With these calibration coefficients, the simple transformation was used to compute the predicted intensities for all the remaining data segments. As expected, the sensitivity response for both FHSTs degraded significantly. As shown by the systematic darkening of Fig. 7 through Fig. 9, the FHST1 response was significantly degraded over the lifetime of the SMM. This is the real degradation of the sensor. In practice, the SMM FHST performance was maintained by calibrating the response every few months.

As the figures illustrate, the SMM FHST1 degradation was very significant. This FHST appeared to age faster in the lower-right quadrant (the +H,-V). Table 2 shows this aging for each segment, and these statistics indicate that the FHST became completely unusable after several years. Table 3 shows that the lower-right quadrant aged faster and became unusable in about 1 year.

The results for FHST2 differ from those of FHST1. Table 2 shows that the FHST2 sensitivity response degraded only slightly over the years. Fig. 10 through Fig. 12 confirm this pattern. While the sensitivity response of FHST2 did degrade, it was not as profound as that of FHST1. This behavior can be partially explained by the alignment of the two FHSTs. The angle between the sensor boresight and the spacecraft to Sun direction was smaller for FHST1 than for FHST2 by approximately 73 deg. Thus, it is possible that more stray sunlight entered FHST1 causing it to age faster.

The causes of aging in star cameras are factors such as temperature variations, space radiation, bright light impingement, space debris collisions, and any other effects native to the space environment. Thus, it may be concluded from this analysis that NASA SSTs age at different rates and, in addition, separate areas of the same FHST may age differently.

CONCLUSIONS

This paper has presented a study of the sensitivity response of the fixed-head star trackers flown on the Solar Maximum Mission. Three studies were performed, and the results may be used by future missions to aid in the performance analysis of any type of star camera. The first study examined two separate intensity-to-magnitude transformation equations and the results indicated that the polynomial transformation provided better results than the simple transformation. On the basis of these results, the polynomial transformation was recommended for use on missions with stringent attitude requirements so that errors due to misidentification of stars may be avoided. The second study was an examination of the individual affects of star color and position on sensitivity. This study showed that star color was the more important effect, and correction for it, either by using

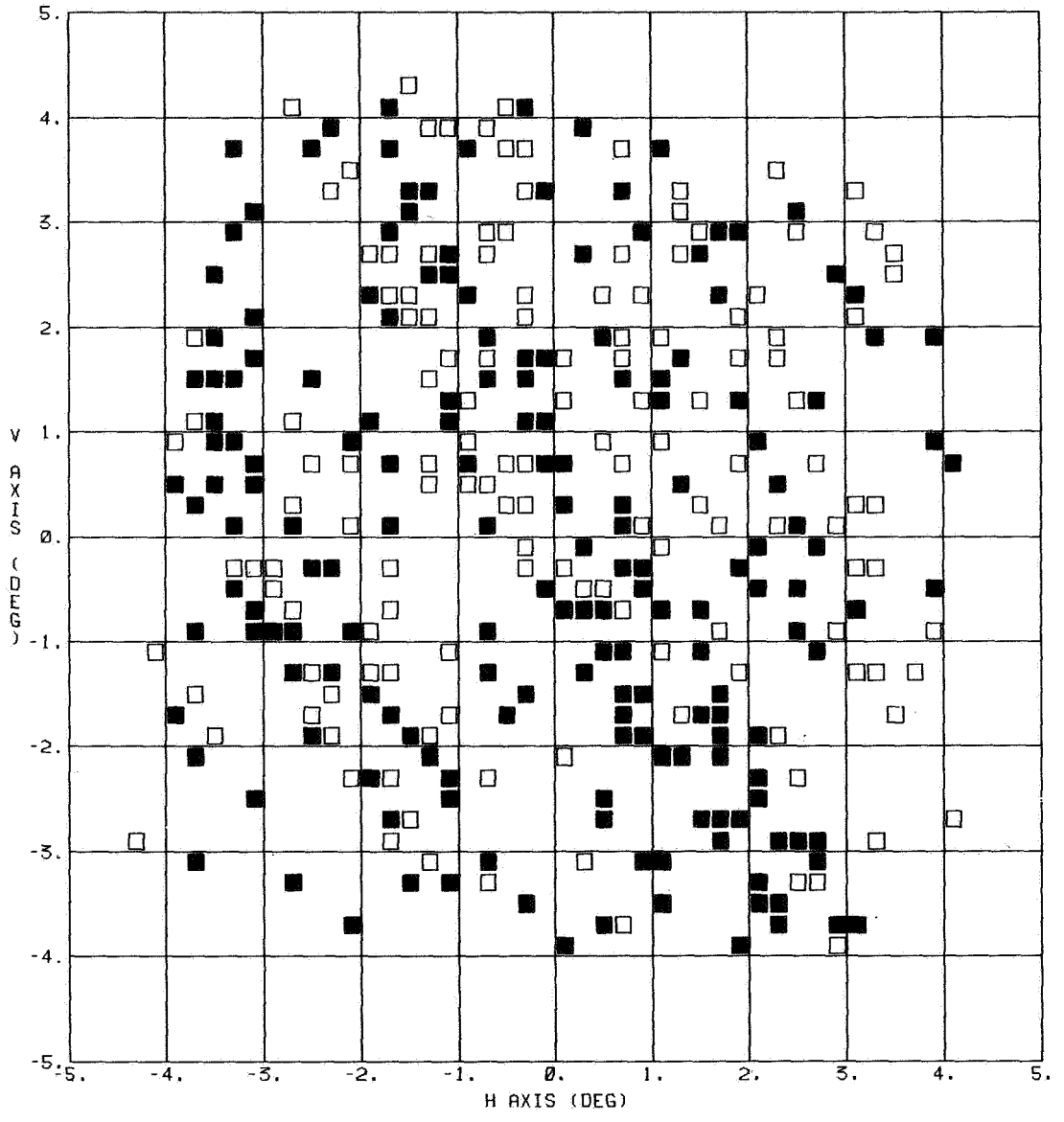


Figure 7. FHST1 Sensitivity Index Plot, Original Transformation Coefficients, 1984

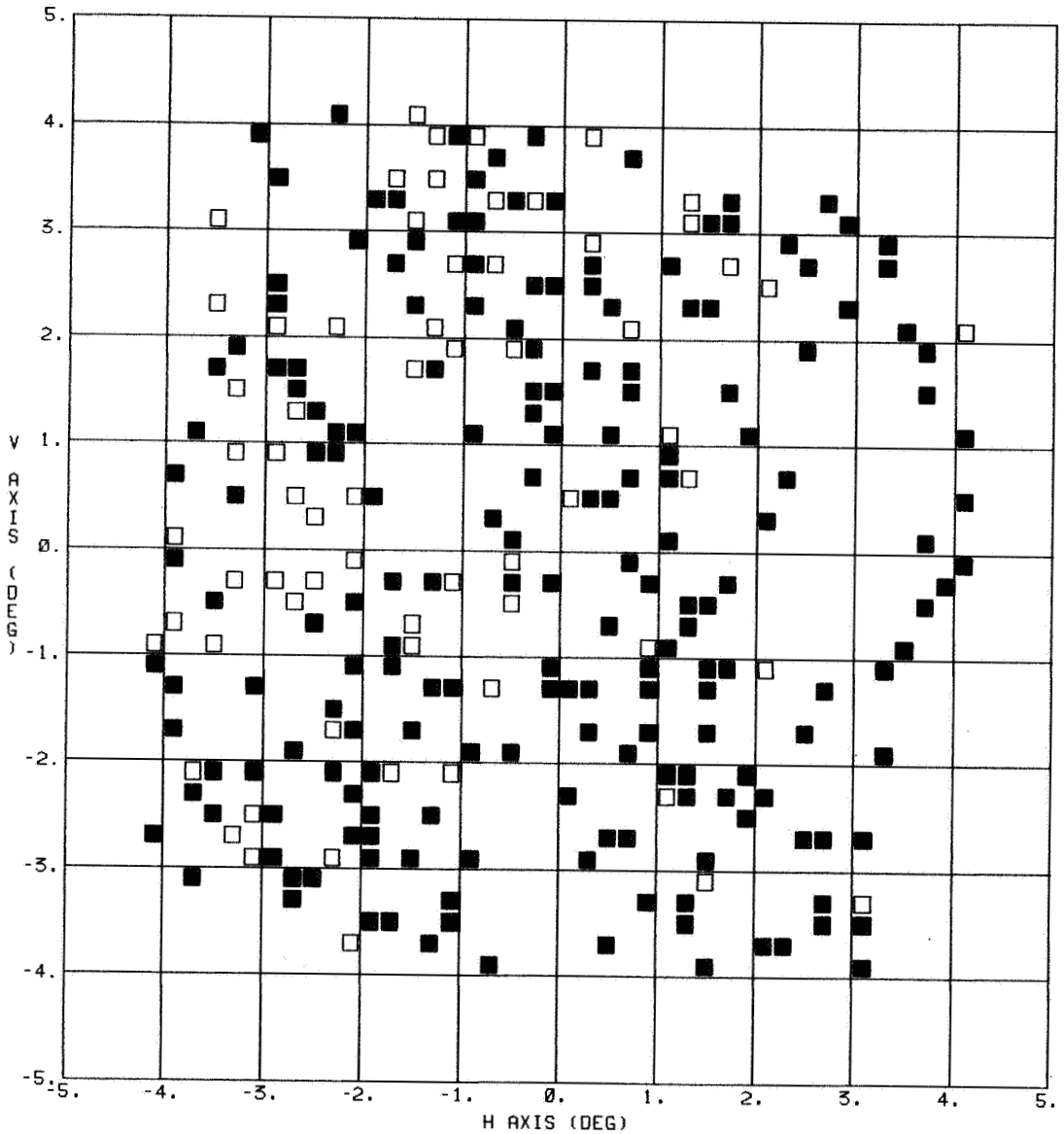


Figure 8. FHST1 Sensitivity Index Plot, Original Transformation Coefficients, 1987

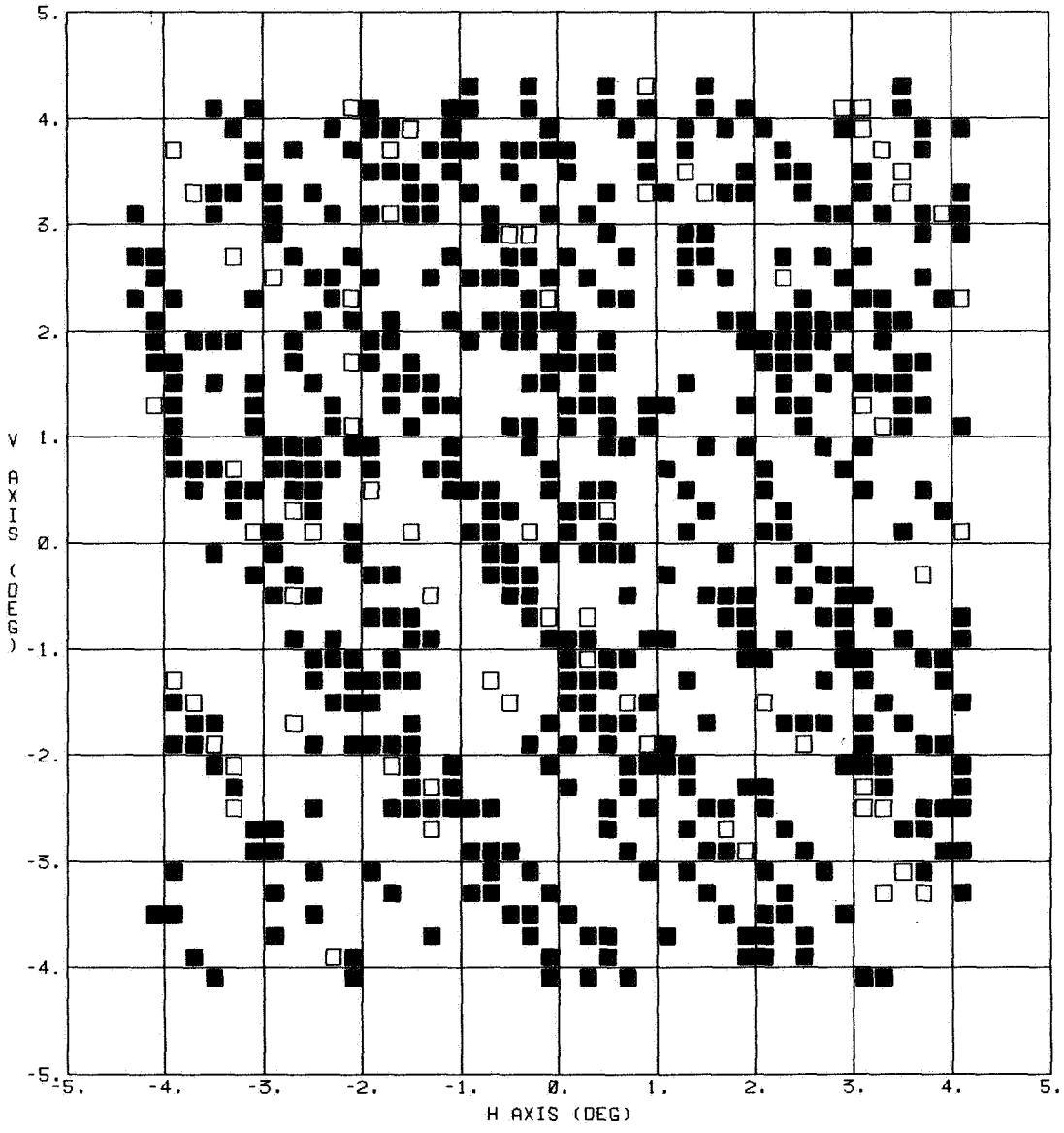


Figure 9. FHST1 Sensitivity Index Plot, Original Transformation Coefficients, 1989

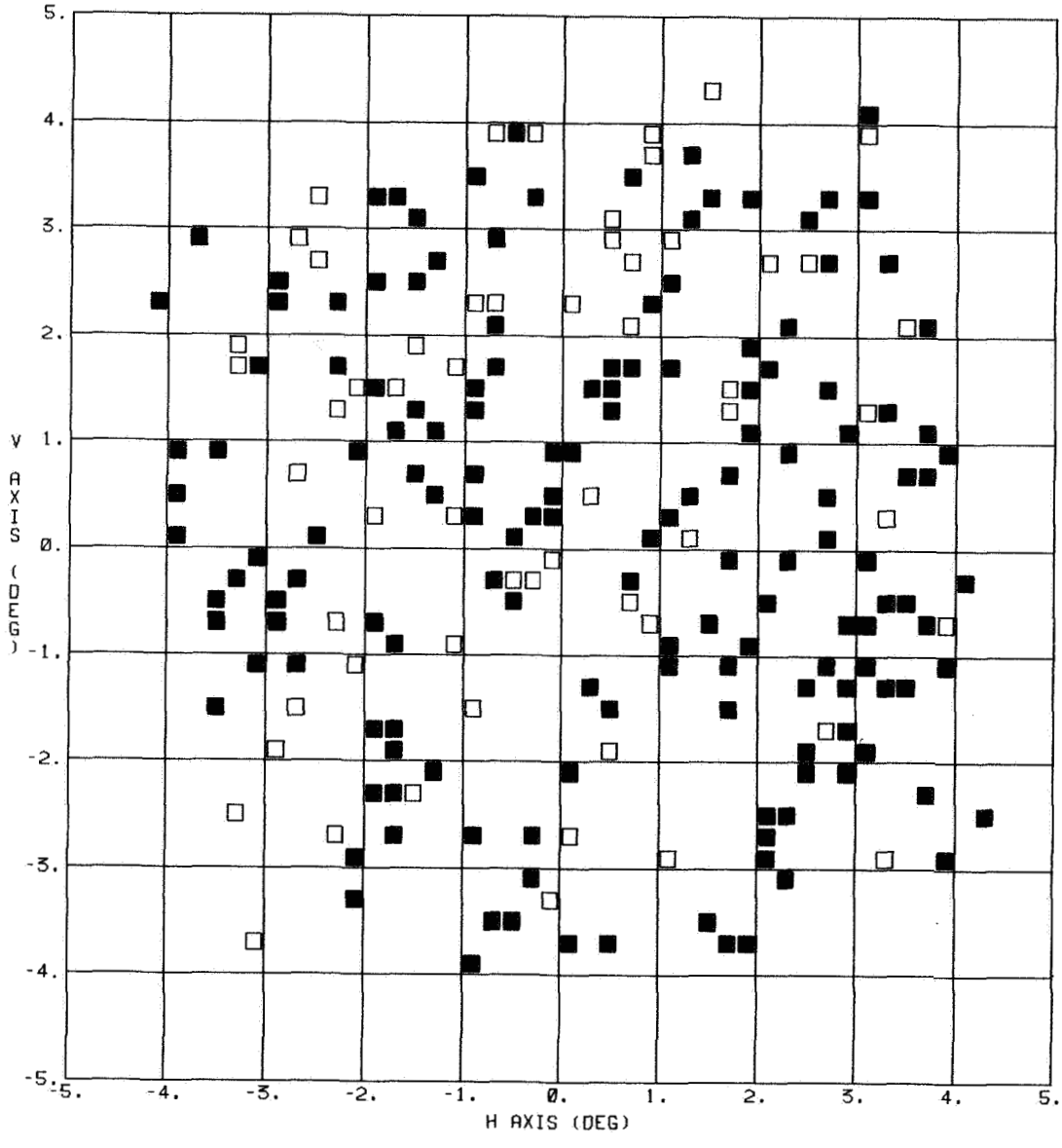


Figure 10. FHST2 Sensitivity Index Plot, Original Transformation Coefficients, 1984

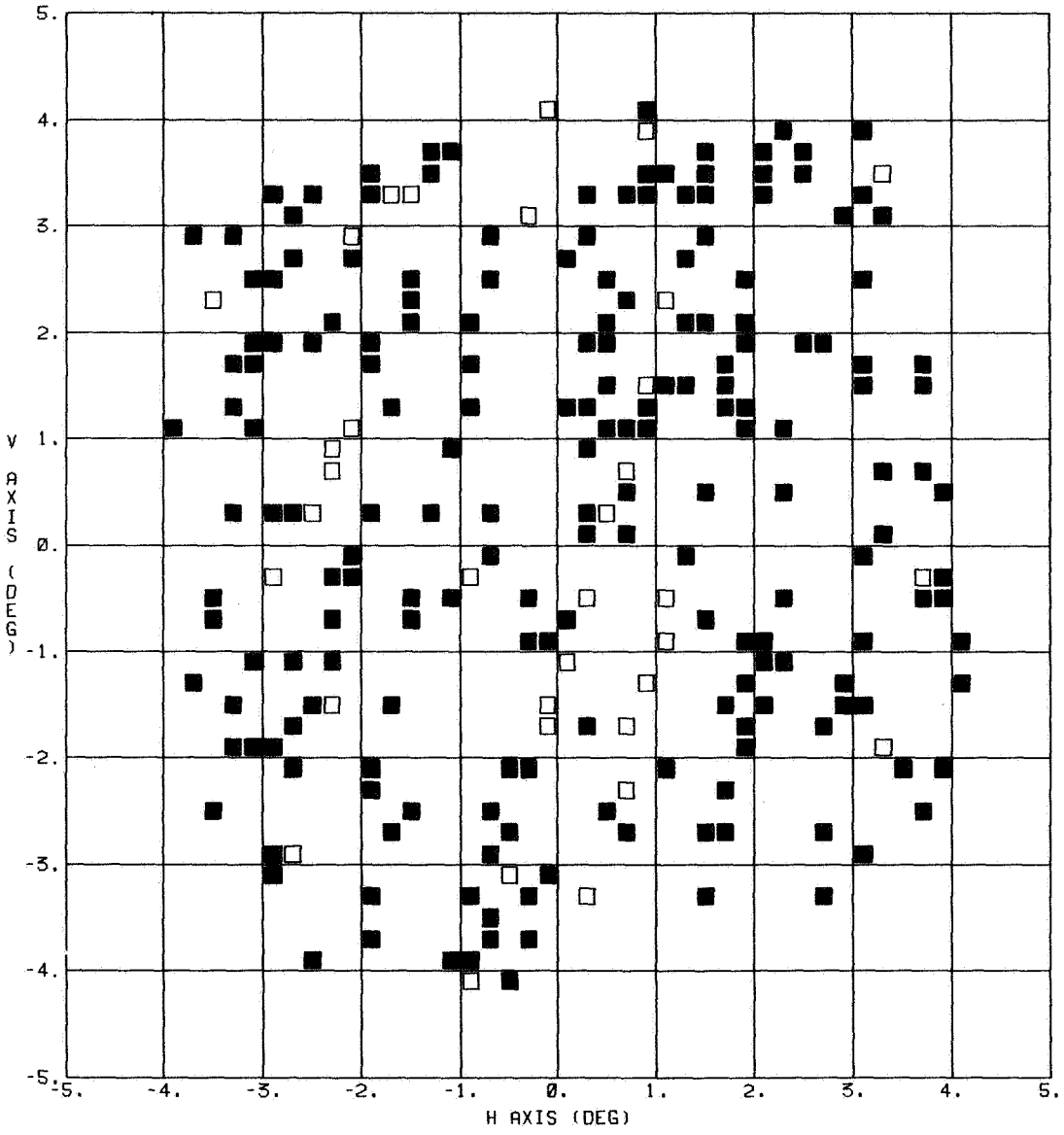


Figure 11. FHST2 Sensitivity Index Plot, Original Transformation Coefficients, 1985

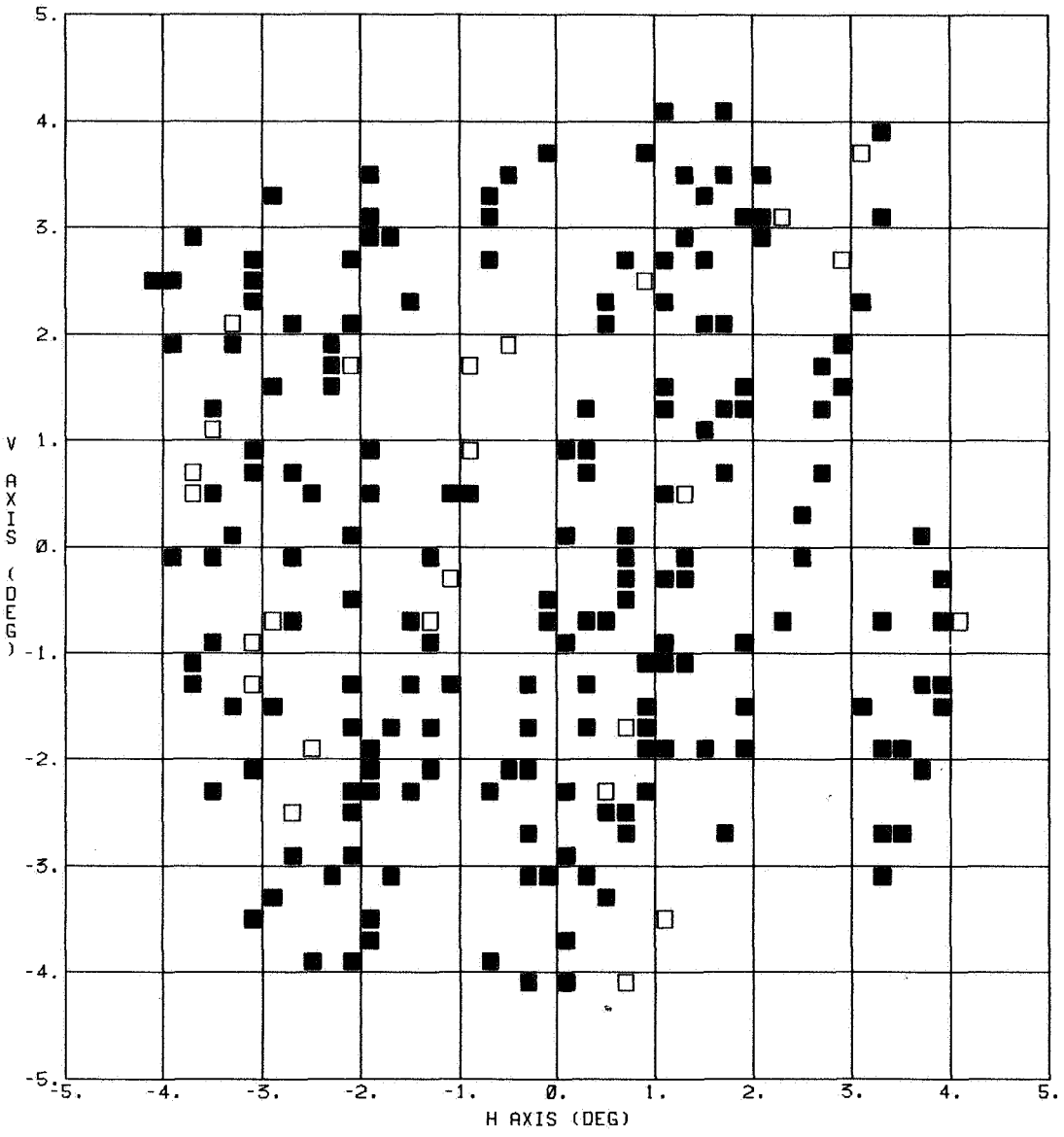


Figure 12. FHST2 Sensitivity Index Plot, Original Transformation Coefficients, 1986

showed that star color was the more important effect, and correction for it, either by using the polynomial transformation equation or a separate equation correcting the final magnitude, is important. The third study examined the aging of star trackers and concluded that aging due to the effects of the space environment was apparent, and that the sensors age at different rates with different characteristics. Therefore, the aging process of each sensor should be studied closely throughout the mission.

ACKNOWLEDGEMENTS

The authors wish to thank the following people for their criticisms and suggestions regarding the material in this paper: Dave Lorenz (GSFC), Mark Slater (CSC), Karen McDaniel (CSC), Milt Phenneger (CSC), Robert Shendock (OAO), and Robert Feiertag (CSC).

REFERENCES

[1] Pitone, D., *Solar Maximum Mission End-of-Mission Document*, Computer Sciences Corporation, CSC/TM-90/6058, March 1990.

[2] *User's Guide for Standard Star Tracker*, Ball Aerospace Systems Division, TM79-04.

[3] Wertz, J.R., ed., *Spacecraft Attitude Determination and Control*, D. Reidel, Dordrecht, the Netherlands, 1978.

[4] McLaughlin, S. and Slater, M., *Skymap System User's Guide, Revision 1*, Computer Sciences Corporation, CSC/SD-89/6076, September 1989.

[5] McCutcheon, R., *SMM Fixed Head Star Tracker Field of View Sensitivity Analysis*, Computer Sciences Corporation, CSC/TM-81/6052, March 1981.

[6] Lorenz, D., *Landsat Attitude Sensor Performance - Evaluation III*, Computer Sciences Corporation, PCA/IM-85/046, September 1985.

[7] Neste, S., *Color Index Computation for the NASA Standard Fixed Head Star Trackers on UARS*, General Electric, PIR NO: U-1k21-UARS-855, June 1987.

COBE ATTITUDE AS SEEN FROM THE FDF*

J. Sedlak, D. Chu, and E. Scheidker
Computer Sciences Corporation

ABSTRACT

The goal of the Flight Dynamics Facility (FDF) attitude support is twofold: to determine spacecraft attitude and to explain deviations from nominal attitude behavior. Attitude determination often requires resolving contradictions in the sensor observations. This may be accomplished by applying calibration corrections or by revising the observation models. After accounting for all known sources of error, solution accuracy should be limited only by observation and propagation noise.

The second half of the goal is to explain why the attitude may not be as originally intended. Reasons for such deviations include sensor or actuator misalignments and control system performance. In these cases, the ability to explain the behavior should, in principle, be limited only by knowledge of the sensor and actuator data and external torques.

This paper documents some results obtained to date in support of the Cosmic Background Explorer (COBE). Advantages and shortcomings of the integrated attitude determination/sensor calibration software are discussed. Some preliminary attitude solutions using data from the Diffuse Infrared Background Experiment (DIRBE) instrument are presented and compared to solutions using Sun and Earth sensors. A dynamical model is constructed to illustrate the relative importance of the various sensor imperfections. This model also shows the connection between the high- and low-frequency attitude oscillations.

*This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, Contract NAS 5-31500.

1. INTRODUCTION

This paper compiles an assortment of results and comments regarding attitude determination and sensor calibration for the Cosmic Background Explorer (COBE). The central problem is to disentangle the true spacecraft motion from the apparent motion reported by imperfect sensors. When these sensors also control the spacecraft, their biases not only produce fictions in the measurements but induce wobbles in the true motion.

First, a brief description of the COBE Attitude Control System (ACS) is given. The Diffuse Infrared Background Experiment (DIRBE) instrument is also described, in the context of using star sightings to increase attitude determination precision. Section 2 discusses the advantages and shortcomings of the integrated attitude determination/sensor calibration software, with emphasis on real-world difficulties. This is followed by a comparison of preliminary attitude solutions using DIRBE data with results using Sun and Earth sensor data. The final section is meant to complement the data reduction techniques of Section 2. Rather than solving for the sensor biases that best match the predicted observations to the real data, the dynamical equations themselves are solved. This allows the separate study of the effects of each bias or misalignment.

1.1 DESCRIPTION OF THE CONTROL SYSTEM

COBE is a spinning, three-axis stabilized spacecraft. Two counter-spinning momentum wheels control the spin rate and leave the spacecraft with zero net angular momentum (except for the 1-rotation per orbit (rpo) pitch rate about the Sun line). Sensors, actuators, and much of the control electronics exist in triplicate on the three control axes (A, B, and C). These axes are symmetrically located in the plane normal to the spin axis. The sensors include a gyroscope, an Earth scanner assembly (ESA), a digital Sun sensor (DSS), and a three-axis magnetometer (TAM) on each control axis. The actuators are reaction wheel assemblies (RWAs) and magnetic torque rods (transverse and X-axis) for momentum unloading. The ACS drives the three reaction wheels independently on each axis in proportion to rate, roll, and pitch error signals (Reference 1).

The DSS most directly facing the Sun measures the Sun elevation and azimuth. A common electronics unit produces the sine and cosine of the Sun azimuth angle. It also produces the sine and cosine of this angle plus or minus 120 degrees (deg) for use by the other axes.

The rate error signal is the gyro output minus an orbit rate-stripping term, normally equal to 1 rpo multiplied by the cosine of the Sun azimuth. Choosing a rate-stripping parameter different from 1 rpo generates an error signal that can be balanced only by a nonzero pitch signal. For this reason, the rate-stripping parameter also serves as the pitchback parameter.

Roll angle is minus the Sun elevation angle. The commanded roll offset is subtracted, and the result is multiplied by the sine of the appropriate Sun azimuth to yield the roll error signal.

The pitch signal is the difference between the instantaneous ESA split-to-index angle and a modulated reference value. The reference is the split-to-index angle sampled when the Sun azimuth is 90 deg (actually 81 deg to avoid Sun interference). At 90 deg, the ESA is directly reading the roll angle. This value is held for a full spin period. It is multiplied by the sine of the Sun azimuth and subtracted from the instantaneous split-to-index. With the roll contribution thus stripped out, the result is the pitch angle multiplied by the cosine of the Sun azimuth.

The rate, roll, and pitch error signals are amplified, filtered, limited, and combined into a torque command signal for the RWAs. The reaction wheel tachometer also provides feedback so that, in the absence of other signals, each wheel will spin down to a commanded speed.

The DSS has a precision of only 0.5 deg, which is an inherent limitation on the ability of the ACS to establish the roll angle and the correct azimuthal phase for modulating the sensor output. Star sightings in the DIRBE instrument, described in the next section, offer data with greater precision. Although these

measurements are not available to the ACS for attitude control, they can be used *a posteriori* for improved ground-based attitude determination.

1.2 DIRBE STAR IDENTIFICATION

DIRBE is a COBE science instrument whose mission is to perform a full-sky survey of diffuse infrared radiation in the wavelength range of 1 to 300 microns. As a bonus, DIRBE also provides star sightings to the Flight Dynamics Facility (FDF) that enhance the FDF's ability to determine attitude. DIRBE has a square field of view, 0.7 by 0.7 deg. The DIRBE boresight (at the center of the field of view) is oriented 30 deg from the spacecraft spin axis. The field of view sweeps out a spiral pattern on the celestial sphere as COBE spins. Two edges of the square field of view are parallel to the scan direction.

As DIRBE scans the sky, pointlike sources, such as stars and planets, pass through the field of view. These sources lead to a sharp increase in the measured infrared intensity. Examples of these "spikes" in intensity are shown in Figure 1. The spike profile can be processed (Reference 2) to determine when the pointlike source passed through the center of the field of view. This "time of passage" (TOP) is then used to predict the boresight direction by interpolating approximate attitudes computed from the attitude determination subsystems.

The predicted boresight direction at the TOP is used for star identification. The observed star is compared to the SKYMAP 2.2 micron (K-band) wavelength reference catalog (Reference 3). For a positive star identification, a reference star must be located within a user-specified angular radius of the predicted DIRBE boresight. Approximately 2000 well separated, sufficiently bright stars are currently in the reference catalog, and about 1.3 valid DIRBE star identifications occur per minute.

Individual star sighting accuracies are expected to be ± 0.35 deg in COBE body elevation and ± 0.1 deg in COBE body azimuth. These star sightings have been used to determine the attitude of COBE independently of Sun and Earth data. In turn, these attitude solutions can be used for sensor calibration.

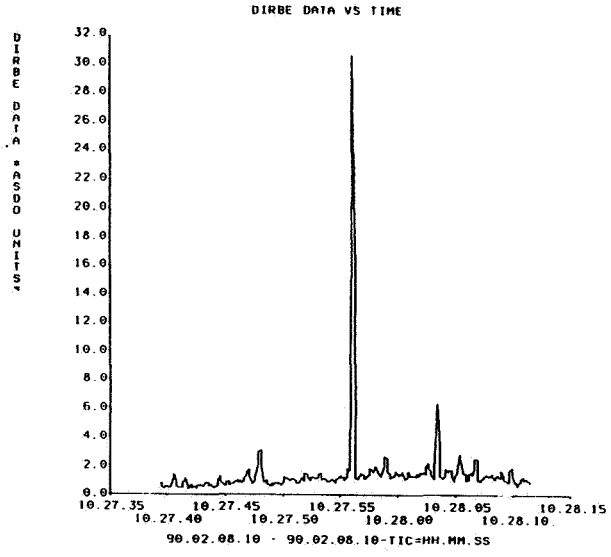


Figure 1. Converted DIRBE Data

2. DETERMINING THE ATTITUDE

Two of the tools used to determine the attitude of COBE are the Coarse Attitude Determination Subsystem (CADS) and the Fine Attitude Determination Subsystem (FADS).

CADS computes single-frame attitudes from "simultaneous" sensor measurements that have been interpolated to a common time. As the content of ESA and DSS data is partly redundant, the user can choose how heavily to weight each. This choice results in a range of solutions reflecting the different sensitivities and misalignments of the various sensors.

FADS estimates these misalignments, the gyro biases, and an epoch attitude by performing a least-squares fit to the data over a given timespan. The bias-corrected gyro data are used to propagate the epoch attitude over the timespan.

2.1 INTEGRATED ESTIMATION

2.1.1 BACKGROUND

COBE is the first spacecraft supported by the FDF to estimate attitude, gyro, and sensor calibrations in a single ground support subsystem, FADS (Reference 4). This integrated capability is the next logical step from the Earth Radiation Budget Satellite (ERBS) FADS (Reference 5), which combined attitude and gyro calibration, and follows in the spirit of the Multisatellite Attitude Determination/Optical Aspect Bias Determination (MSAD/OABIAS) System (Reference 6), which is the standard estimator for spinning spacecraft.

The COBE FADS is a batch-weighted least squares estimator that can use any combination of the 23 observation types to solve for almost any combination of the 49 possible state parameters. The three attitude parameters must be solved for as a group. This approach is attractive because it obviates the need for separate calibration utilities and the need to go back and forth between them to obtain a complete set of self-consistent calibrations. Instead, all parameters are solved for in the same subsystem and, in principle, at the same time. Although the first benefit is definitely realized, the second benefit, simultaneous solution, is problematic for several reasons.

First, solving for many parameters at once often causes the solution to diverge or to converge to an answer worse than the *a priori* solution. Only when the parameters being solved for are very close to their correct values is there a benefit to solving for many parameters at once. Then, the solution usually improves. This problem may be solved, in whole or part, by changing the method of solving the batch least-squares problem. Marquardt's algorithm and others offer alternatives to the standard Newton-Raphson method used here.

Second, a byproduct of this great flexibility is confusion. With so many parameters, it becomes hard to decide what to include in the state vector. At present, choosing parameters depends on the analyst's ability to recognize patterns in the residuals and to identify the parameter causing the pattern. Because many parameters have similar effects and may combine to produce unrecognizable residual patterns, it may take longer than the 2 months usually allotted for launch support to decide exactly what is wrong.

Third, whether it is a simple mistake in the specifications or a more fundamental problem, gyro calibration takes a very long time to converge. As COBE convergence criteria are maximum values arbitrarily set for the change to the state, a more meaningful comment might be that the gyro calibrations approach their final values slowly and monotonically. They never overshoot their true values, as do the other parameters. Moreover, they approach slowly, typically by halves. Because FADS is already slow, the need for so many iterations increases the central processing unit (CPU) time demands and restricts calibration work to off-hours when the mainframe is not being heavily used.

Thus, although it is possible to solve for all 49 parameters at once, that has not been a real benefit of the integrated calibration approach. The main benefits have been in centralizing the solution process and in

providing a better, more complete view of the calibration problem, as reflected in part by the following statistics.

2.1.2 STATISTICS

Statistics are useful for assessing solution accuracy and selecting parameters for the state vector. The following statistics are being tried for COBE. Knowing the solution accuracy can help in recognizing when a solution is diverging and in choosing between competing solutions. To evaluate the quality of a batch solution, the FADS provides the weighted root mean square (RMS) residual (σ_ϱ). This is a single number whose magnitude is the size of a typical residual (Δy_i) at time t_i weighted by the weight matrix (W). The set of calibrations that gives the smaller RMS for the same data is probably the better set. To first order, the RMS residual is independent of the number of data points and the choice of observation weights.

$$\sigma_\varrho^2 = \left(\sum_{i=1}^{N} \Delta \underline{y}_i^T \, W \, \Delta \underline{y}_i \right) \Big/ N \text{ trace } (W) \tag{1}$$

To know how accurate each parameter of the solution really is, FADS provides a residuals-based estimate of the covariance matrix. Rather than assume that the observation residuals are zero mean and white, with the variance assumed for weighting purposes, the covariance estimate is computed from its definition:

$$\text{cov } (x) = E \, [\Delta \underline{x} \, \Delta \underline{x}^T] - E \, [\Delta \underline{x}] \, E \, [\Delta \underline{x}]^T \tag{2}$$

The two terms are computed from the expressions for the change to the state. Here, F_i represents the partial derivative of the observation types with respect to the state variables at time i, based on the same currently estimated state used to predict the observations:

$$E \, [\Delta \underline{x} \, \Delta \underline{x}^T] = \left(\sum_{i=1}^{N} F_i^T \, W \, F_i \right)^{-1} \left(\sum_{i=1}^{N} F_i^T \, W \, \Delta \underline{y}_i \, \Delta \underline{y}_i^T \, W \, F_i \right) \left(\sum_{i=1}^{N} F_i^T \, W \, F_i \right)^{-1} \tag{3}$$

$$E \, [\Delta \underline{x}] = \left(\sum_{i=1}^{N} F_i^T \, W \, F_i \right)^{-1} \left(\sum_{i=1}^{N} F_i^T \, W \, \Delta \underline{y}_i \right) \tag{4}$$

One possible explanation for the divergence of solutions and the slow convergence of the gyro calibration is that the observabilities of the state parameters are too different. For a state vector of two elements, this means that the level curves of the loss function being minimized are elliptical rather than circular. If the minimization algorithm moves perpendicular to these level curves, the less observable parameter can be slow to converge (Figure 2).

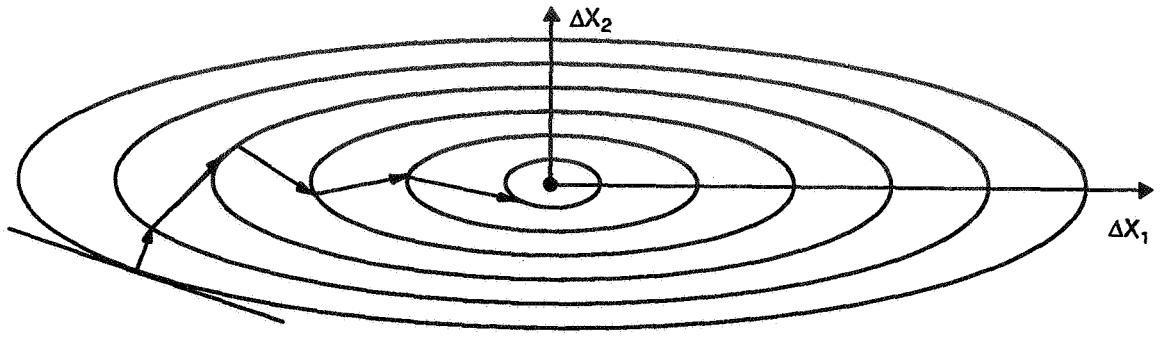


Figure 2. Convergence With Very Different Observabilities

The diagonal elements of the normal matrix ($\Sigma\ F^T\ W\ F$) provide estimates of the relative observability of the different parameters in the state vector. Due to differences in the units used for computation and display, however, the standard deviations in the iteration summaries may not reflect the true conditioning of the normal matrix. A more direct indication of the unbalanced observabilities and singularity is the condition number (K), which is the product of the row norm of the normal matrix and that of its inverse (Reference 7):

$$K(A) = \| A \| \ \| A^{-1} \|$$

When K is large, the parameters solved for have very different observabilities, and the matrix is nearly singular. As a further check, the maximum and minimum eigenvalues of the normal matrix are computed along with their eigenvectors. The associated eigenvectors indicate which parameters are the most or least observable. This tool has not proven very useful because condition numbers routinely reach 1 million without the solution diverging, and convergence can be slow without there being a very large condition number.

Knowing what is most observable and least observable, however, does not dictate the parameters to include in the state vector. The most observable parameter may be completely accurate, and the error may be caused by the least observable one. An enhancement, which has not yet been implemented, takes the inner product of the residual histories over the batch timespan with the predicted residual histories due to each parameter error. The computation entails normalizing the derivative of the residual vector ($\Delta \underline{y}$) at time i with respect to each parameter (x_j):

$$\hat{\underline{f}}_{ji} = \underline{f}_{ji} \Big/ \left(\sum_{i=1}^{N} \underline{f}_{ji}^T \underline{f}_{ji} \right)^{1/2} \qquad (5)$$

where

$$\underline{f}_{ji} = \frac{\partial \Delta \underline{y}_i}{\partial x_j}$$

and accumulating this dot product to indicate which parameters best fit the residuals.

$$(\underline{c})_j = \frac{1}{N} \sum_{i=1}^{N} \hat{\underline{f}}_{ji}^T \, \Delta \underline{y}_i \qquad (6)$$

2.1.3 COBE CALIBRATION HISTORY

Before launch, calibration looked straightforward. FADS was working. Signatures of the individual calibration errors had been identified, and a "foolproof" procedure had been set down whereby anyone could determine the calibrations. The only thing that would prevent calibration from being completed the day of launch was that DIRBE measurements would not yet be available.

Important results were obtained in the first few days. Earth sensor-A appeared to have a twist misalignment (split-to-index bias) of 0.5 deg, and the X-gyro scale factor was 0.1 percent high. After that, however, progress was slow. Finding scale factor corrections for the transverse axis gyros proved treacherous, and when transverse gyro biases were solved for, the results were not repeatable. Localized sources of error, such as horizon radiance, also turned out to be more significant than expected.

One of the early "casualties" in this struggle was the original calibration plan. Reasonably, it had seemed, sensor alignments should be solved first, over short timespans, before gyro errors could become significant. Gyros would then be calibrated. In practice, however, the X-gyro scale factor error could not be ignored even over a single 75-second spin period, and the procedure was changed to make it the first parameter solved.

With the Earth sensor and X-gyro scale factor corrections, it was possible to increase the length of the batches to 20 minutes without the observation residuals exceeding 0.5 deg. Prelaunch information had also suggested that batches longer than this would have propagation errors of more than 0.1 deg due to gyro noise. As a result, 20 minutes became the standard batch length. The next significant problem was the growing Earth sensor split-to-index residuals. These residuals could be reduced by assuming large transverse scale factor errors. This solution was considered valid for several weeks until it was noted that attitude histories computed in this way had pitch discontinuities at the batch boundaries. The large scale factors had to be abandoned.

After a few weeks, it was realized that without the large scale factors the roll drifted off, and solving for sensor alignments along with the X-gyro scale factor made the residuals and discontinuity much smaller. When a full orbit was processed, the roll came back over the second half of the batch. This identified the problem as a spin phase error similar to that caused by incorrect telemetry time tags or single-step attitude propagation (Reference 8). It turned out that the Sun sensors that measure yaw were slightly misaligned with respect to the gyros. By correcting the Sun sensor azimuth alignments by approximately 0.1 degree, the problem was resolved.

Recently, attention has returned to the question of transverse gyro scale factor errors, because single-frame and FADS pitch solutions diverged slightly over long times. When 0.002 scale factor corrections were made, the Earth sensor residuals were halved, and the gyro-propagated attitude repeated from orbit to orbit. These results corroborate earlier ones obtained without examination of pitch repeatability, the full importance of which was not appreciated at the time.

One of the first calibration objectives following launch was to determine transverse gyro biases. Even today, these remain a puzzle. The problem is that Earth sensor twist misalignments and transverse gyro biases have the same effect on the split-to-index residuals. If zero biases are assumed, Earth sensor-A twist appears to be 0.5 deg. If the twist is fixed at its prelaunch value of 0.3 deg, the biases on gyro-A and -C are on the order of 30 deg per hour. It has not been determined what the real answer is.

The small coning motion implied by such a twist or bias is not readily observable from the Sun sensor elevation measurements because of their coarse 0.5 deg quantization and the attitude being controlled to maintain constant Sun elevation. DIRBE elevation measurements, with their 0.7 deg field of view and approximate 1-minute spacing, are similarly crude. Nor does the control system provide any help. As both the Earth sensors and the gyros are in the control loop, various mixtures of the two errors also cause coning. This may be a situation in which if it is so hard to decide, it probably does not matter.

Finally, although observation residuals are now about as small for a full orbit as they are for one spin cycle, old data must be reexamined to see if the calibrations that work well now would have been as successful then. Horizon radiance effects, which peak at the solstices and vanish at the equinoxes, might have contributed to those large early residuals. Other localized errors that appear at the orbit rate and its harmonics are also under study.

The COBE calibration effort has not been quick or smooth; each advance has been hard won. Moreover, these solutions could probably not have been foreseen before launch, even with twice the preparation time. The combined spinning and pitching made COBE calibration more confusing, but each mission is certain to have its own challenges, and so calibration may always be a puzzle. In solving the puzzle, ideas are like pieces, the more the better, but each must be evaluated and recorded to avoid confusion and needless duplication of effort. The integrated estimation software was able to reduce some of this confusion by providing a broad view of all the observations and calibrations at once, but it was not a panacea.

2.2 DIRBE ATTITUDE RESULTS

Preliminary results with DIRBE indicate good agreement with Sun and Earth attitude solutions. However, DIRBE misalignments have not yet been computed, and the solutions seem sensitive to these misalignments. A comparison was made between FADS runs using data from January 29, 1990. One (the nominal case) assumed the DIRBE boresight to be aligned with the nominal body azimuth (240 deg in the body). The other (the misaligned case) assumed a -0.3 deg misalignment of the boresight with respect to the nominal azimuth. The computed roll in the nominal case, when averaged over a spin cycle to eliminate the effects of short-term oscillations, experienced a long-term drift of 0.3 deg over 25 minutes (Figure 3). This agreed with FADS solutions, using Sun and Earth data, to better than 0.06 deg in the roll, pitch, and yaw. In the misaligned case, however, roll experienced a long-term drift of only 0.1 deg in 25 minutes (Figure 4). Sun elevation data drift by roughly 0.1 deg over the 25-minute time interval. Based on Sun elevation information alone, the misaligned DIRBE solution is more believable. However, the half-degree digitization of the DSS makes such a comparison suspect. Adding to the uncertainty is the possibility that the Earth scanner information is corrupted by unmodeled horizon radiance effects.

3. DYNAMICAL MODEL

This section approaches the attitude problem from a different direction. The intent is to understand what is driving both the high-frequency motion (time scale = spin period) and the low-frequency motion (time scale = orbital period), which have been observed since the failure of the B-axis control gyro. The starting point is the system of Euler equations and the attitude control laws. Sensor misalignments, biases, and failures can be added one at a time to see how each contributes to the attitude motion.

During one spin period, the COBE body X_C-axis traces out two unequal cones; the pitch and roll angles oscillate with dominant frequency components equal to the spin rate and twice the spin rate. The amplitudes are a few tenths of one degree and are dependent on the spin rate.

During one orbit period, the spin-averaged pitch angle varies by a few degrees, with the amplitude depending on the time of day. Two unequal peaks appear each orbit, so the dominant frequency components are ω and 2ω, where ω is the orbital frequency.

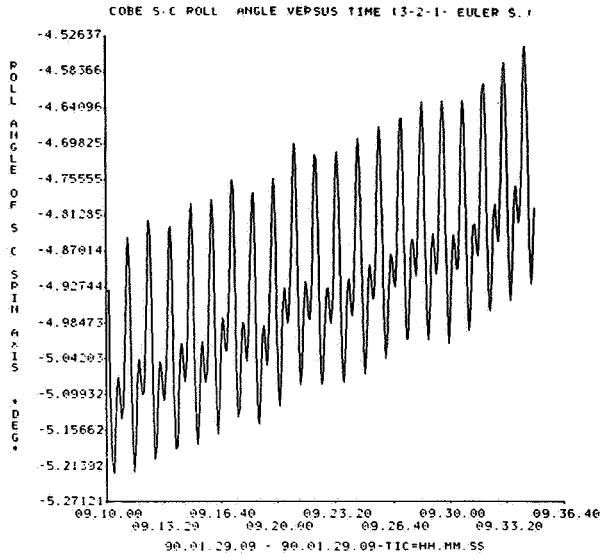


Figure 3. Roll Angle Solution Using Nominal DIRBE Alignment

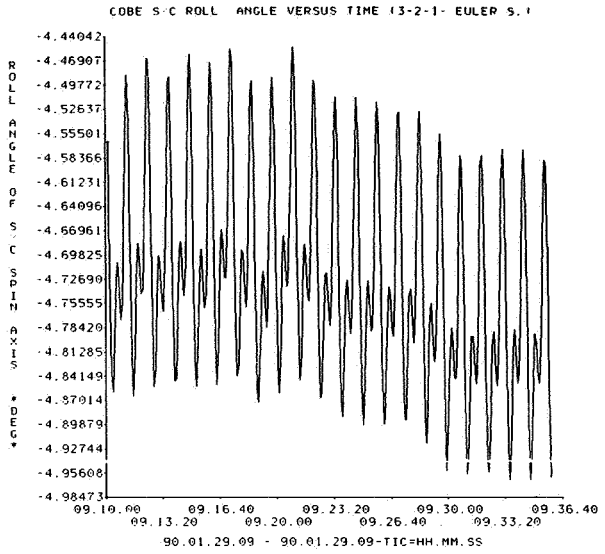


Figure 4. Roll Angle Solution Using -0.3 Deg Azimuthal DIRBE Misalignment

3.1 MODEL EQUATIONS

The derivation of the equations of motion and control laws is outlined in Appendix F of Reference 1. The following is a brief summary of those equations and definitions. (Reference 9 also derives similar equations and applies them to a stability analysis of the nominal ACS configuration.)

A convenient reference system is the despun frame, which is the COBE body frame rotated to zero deg yaw angle. The Euler equations in this frame can be written

$$\dot{L}_x = \Omega \cos\theta\, h_r - \dot{\theta}\, h_p + T_x$$

$$I\,\ddot{\theta} = -\Omega \cos\theta\, L_x - I\,\Omega^2 \sin\theta \cos\theta - \Omega \sin\theta\, h_p - \dot{h}_r + T_r \qquad (7)$$

$$I\,(\ddot{\phi} - \ddot{\eta}) \cos\theta = 2\,I\,\Omega\,\dot{\theta} \sin\theta + \dot{\theta}\, L_x + \Omega \sin\theta\, h_r - \dot{h}_p + T_p$$

where

ψ, θ, ϕ = yaw, roll, pitch

η = ideal pitch anomaly

Ω = $\dot{\phi} - \dot{\eta} + \omega$

ω = orbital rate

I = transverse moment of inertia = 2700 ft-lb-sec^2

L_x = X-axis angular momentum

h_r, h_p = net angular momentum of the three reaction wheels, measured in the body frame and transformed to the despun frame

T_x, T_r, T_p = external torques (e.g., gravity gradient and torquer bars)

The ACS control torque commands to the three wheels are projected on the body Y_C- and Z_C-axes and then transformed to the despun frame:

$$\begin{aligned} \dot{h}_r &= -K_t h_r + 1.5 (K_p + K_r F_{20}) (\theta - \theta_o) + 1.5 K_g \dot{\theta} - \dot{\psi} h_p \\ \dot{h}_p &= -K_t h_p + 1.5 K_p \phi + 1.5 K_g (\Omega \cos \theta - \Omega_{cmd}) + \dot{\psi} h_r \end{aligned} \tag{8}$$

where θ_o is the roll offset and Ω_{cmd} is the orbit rate-stripping parameter. The feedback loop gains are

$$K_p = 1.07 \text{ ft-lb/rad}$$

$$K_r = 6.59 \text{ ft-lb/rad}$$

$$K_g = 189.6 \text{ ft-lb/(rad/sec)}$$

$$K_t = 0.0168 \text{ sec}^{-1}$$

The yaw rate is related to the X-angular momentum through the following equation:

$$\dot{\psi} = \dot{\psi}_o + L_x/I_x + \Omega \sin \theta \tag{9}$$

where $\dot{\psi}_o$ is the nominal yaw rate.

The ACS roll loop has a 20-second filter to smooth out the 0.5-deg digitization of the DSSs. Although the digitization is not included in the model, the 20-second filter has been retained. The filter operator F_{20} converts θ into θ_F where

$$(20 \text{ sec}) \dot{\theta}_F + \theta_F = \theta - \theta_o \tag{10}$$

To a good approximation, the pitch anomaly is

$$\eta \approx - \frac{360^\circ}{2\pi} \tan^2 (\varphi/2) \sin (2\omega t) \tag{11}$$

where φ is the difference between the Sun's declination and that of orbit normal.

Linearizing the dynamical system in the roll, pitch, L_x, and control torques, and taking the Laplace transform, yields

$$\begin{pmatrix} I(s^2 + \omega^2) & 0 & (s^2 + \omega^2)/s & 0 \\ 0 & I\,s^2 & 0 & s \\ -B & 0 & s + K_t & \dot{\psi}_0 \\ 0 & -A & -\dot{\psi}_0 & s + K_t \end{pmatrix} \begin{pmatrix} \bar{\theta} \\ \Delta\bar{\phi} \\ \bar{h}_r \\ \bar{h}_p \end{pmatrix} = \begin{pmatrix} \tau_r \\ \tau_p \\ H_r \\ H_p + P + E \end{pmatrix} \tag{12}$$

where s is the complex Laplace transform variable, $\Delta\bar{\phi}$ is the transform of $\phi - \eta$, and

$$\begin{aligned} A &= \frac{3}{2}\,(K_p + s\,K_g) \\ B &= A + \frac{3}{2}\,K_r/(1 + 20s) \\ E &= -\frac{3}{2}\,K_p \tan^2\left(\frac{\varphi}{2}\right) \frac{2\omega}{4\omega^2 + s^2} \\ P &= \frac{3}{2}\,K_g\,(\Omega - \Omega_{cmd})/s \end{aligned} \tag{13}$$

The initial conditions have all been taken to be zero. P is the pitchback command, and E is the pitch anomaly. H_r, H_p are errors caused by the failed gyro and ESA twist misalignment discussed in the next subsections. T_x is assumed to be zero.

The characteristic equation is obtained by setting the determinant of the matrix in Equation (12) to zero. The determinant can be written

$$\det = I^2\,(s^2 + \omega^2)\,D \tag{14}$$

$$D = \frac{1}{0.05 + s} \prod_{k=1}^{5} (s - s_k) \tag{15}$$

With $\dot{\psi}_o = -0.815$ revolutions per minute (rpm) and two working gyros, the roots are

$$\begin{aligned} s_1 &= -0.009715 \\ s_2 &= -0.02740 + i\,0.01604 \\ s_3 &= -0.07978 + i\,0.08095 \end{aligned} \tag{16}$$

and the complex conjugates, all having units of sec^{-1}. These are the time constants for the decay of transients. They all have negative real parts, so the system is linearly stable.

One can solve the algebraic equations for $\Delta\bar{\phi}$ and then obtain $\Delta\phi(t)$ by inverting the Laplace transform. The inverse is obtained by inspection from the singularities in $\Delta\phi$, discarding the transient solutions. The residues of the simple poles give the amplitudes. Double poles lead to oscillations plus secular terms; that is, terms that grow linearly with time. These occur when an external roll torque oscillates at the orbital rate, always pointing the same direction in inertial space. Such secular terms (for example, the gravity gradient torque) must be canceled by the secular terms from the momentum management assembly (MMA) torque. On average, the spacecraft is not accumulating angular momentum. These terms are dropped in the following solutions.

In this manner, the equations have been solved for the commanded pitchback (PB), the pitch anomaly (PA), and the gravity gradient (GG) source terms:

$$\begin{aligned} \phi(t) = &-6.883 \quad \text{(PB)} \\ &-0.1833 \sin(2\omega t + 76.75 \text{ deg}) \quad \text{(PA)} \\ &-0.3721 \sin(\omega t - 34.14 \text{ deg}) \quad \text{(GG)} \end{aligned} \tag{17}$$

in units of degrees. These solutions assume that the yaw rate is −0.815 rpm, the Sun declination is −21 deg, the roll offset is −4 deg, and the gain in the gyro feedback loop is appropriate for two working gyros. The pitch anomaly has been added back in, so Equation (17) represents the actual pitch angle and not $\Delta\phi$.

3.2 B-AXIS GYRO FAILURE

With the failure of the B-gyro, the ACS torque command to the B-axis RWA became

$$\begin{aligned} \dot{h}_B &= \dot{h}_B^{(nom)} - K_g\left[\dot{\theta}\sin\psi_B + (\dot{\phi} - \dot{\eta} + \Delta\Omega)\cos\psi_B\right] \\ \Delta\Omega &= \omega - \Omega_{cmd} + \Omega_B \end{aligned} \tag{18}$$

where $\dot{h}_B^{(nom)}$ is the nominal torque command, $\psi_B = -(\psi + 120 \text{ deg})$ is the Sun azimuth measured from the B-axis (note that azimuth and yaw angles are opposite in sign), Ω_{cmd} is the rate-stripping parameter on the A- and C-axes, and Ω_B is the rate-stripping parameter for the B-axis. Since the gyro failure, rate-stripping has been set to zero on all three axes, but it may be desirable to command a different value for the B-axis now that the symmetry has been broken. To allow for this, the nominal

Ω_{cmd} is subtracted, and Ω_B is explicitly included. The sine and cosine factors in Equation (18) represent the projection of roll and pitch rates on the B-axis.

After projecting the new B-axis torque command on the body Y_C- and Z_C-axes, transforming to the despun frame, and taking the Laplace transform, the resulting source terms become

$$H_r = -\frac{K_g}{2}\left\{s\bar{\theta} + \frac{\Delta\Omega}{2}\left(\frac{3^{1/2}\,s + 2\dot{\psi}}{s^2 + 4\dot{\psi}^2}\right) + L\left[\dot{\theta}\cos(2\psi + 60°) + \Delta\dot{\phi}\sin(2\psi + 60°)\right]\right\} \quad (19)$$

$$H_p + P = -\frac{K_g}{2}\left\{s\Delta\bar{\phi} - \frac{\Delta\Omega}{2}\left(\frac{s - 3^{1/2}\,2\dot{\psi}}{s^2 + 4\dot{\psi}^2}\right) - \frac{2\omega - 2\Omega_{cmd} - \Omega_B}{s} + L\left[\dot{\theta}\sin(2\psi + 60°) - \Delta\dot{\phi}\cos(2\psi + 60°)\right]\right\} \quad (20)$$

The first term on the right-hand side of Equations (19) and (20) can be moved to the left-hand side of Equation (12). This effectively reduces the rate gain from $3K_g/2$ to K_g in the matrix elements A and B. The second term in both equations drives the system at twice the spin rate.

The third term in Equation (20) determines the pitchback angle when it is balanced by the pitch feedback term from Equation (8). The nominal pitchback term, P, is included in Equation (20) to make explicit the net pitchback. Writing $2\Omega_{cmd}$ as $\Omega_A + \Omega_C$, the pitchback angle for the ACS with N_g working control gyros is

$$\phi_o = -\frac{K_g}{K_p}\left(\frac{N_g\,\omega}{3} - \frac{\Omega_A + \Omega_B + \Omega_C}{3}\right) \quad (21)$$

With the rate-stripping parameters set to zero and $N_g = 2$, the pitchback is −6.883 deg.

In the final terms of Equations (19) and (20), the L is the Laplace transform operator. These terms are dropped in first order. They are added back in iteratively, using the rates obtained in the first-order solution. This produces terms with zero frequency (constant 0.1-deg roll and 0.75-deg pitch offsets) and small oscillatory terms with a frequency four times the spin rate (amplitude 10^{-5} deg).

In solving for roll and pitch as functions of time, it is found that the failed gyro forces the spacecraft to move on an almost circular cone, twice per spin:

$$\theta = 0.100 - 0.103\sin(2\psi + 14.46°) \quad (22)$$

$$\phi = -6.883 + 0.745 - 0.102\sin(2\psi - 76.60°) \quad (23)$$

where $\psi = \dot{\psi}_o t$ and $\dot{\psi}_o = -0.815$ rpm. Angles are given in degrees. The −6.883 deg term is the normal pitchback, and the other two constants are the second-order corrections. The cone is larger and less circular for lower spin rates.

The part of the high-frequency motion that is driven by the failed gyro can be prevented by choosing rate-stripping parameters that make the oscillatory source terms vanish; that is, set $\Delta\Omega = 0$. This can be done for any desired pitchback angle, ϕ_o. The required parameters are

$$\Omega_A = \Omega_C = \Omega_{cmd}$$

$$\Omega_B = \Omega_{cmd} - \omega \qquad (24)$$

$$\Omega_{cmd} = \omega + \phi_o\,(K_p/K_g)$$

The choice of Ω_A and Ω_C is not affected by the failed gyro, and the value of Ω_B is reduced by ω.

3.3 ESA TWIST MISALIGNMENT

The sensor calibration feature in FADS has established the ESA twist misalignment angles to be T_A = 0.54 deg and T_C = −0.15 deg. These twists alter the index angle in measurements of the split-to-index. (This FADS solution assumes that all the error is due to ESA twist and none to transverse gyro drift bias.)

Analysis similar to that of the previous section yields the attitude solutions (in deg):

$$\theta = -0.0182 - 0.0394 \sin(\psi - 85.38°) + 0.0063 \sin(2\psi + 32.54°) \qquad (25)$$

$$\phi = +0.0355 \sin(\psi + 0.72°) + 0.0062 \sin(2\psi - 58.53°) \qquad (26)$$

The error in the instantaneous split-to-index drives the system at the spin rate; the error in the sample-and-hold split-to-index drives the system at twice the spin rate and generates the offset.

The predicted high-frequency attitude motion is the sum of these expressions and the failed gyro solutions given in Equations (22) and (23). This motion is plotted in Figure 5. The figure shows two spin periods, starting from $\psi = 0$ at t = 0. Figure 6 shows two spin periods of spacecraft data as determined by FADS. In that figure, the yaw angle is zero at time 19.31.11. The phase and overall shape of the FADS solution are well reproduced by the model equations driven only by gyro and ESA errors. The average values disagree because neither the commanded roll offset nor the magnetic torques are included in Figure 5. The predicted amplitude falls below the actual value by approximately 0.05 deg. These calculations have been repeated for ψ_o = −0.225, −0.4, and −0.6 rpm. The qualitative agreement with FADS is good at all spin rates, but the amplitude is consistently low. This is possibly caused by an unmodeled momentum wheel misalignment, which also shows up as a larger-than-predicted, spin-rate-dependent offset of the reaction wheel speeds. Gyro drift biases of roughly 30 deg per hour also could possibly make up the difference.

3.4 MAGNETIC TORQUES

Under ideal conditions, the RWA holds the spacecraft at the desired attitude with the wheels absorbing all the accumulated angular momentum. The MMA takes the tachometer signals, filters out any constant offset, and energizes the torquer bars to dump this excess angular momentum. However, imperfect sensors cause the ACS to spin the wheels up and down, even when the attitude is nominal. The MMA receives the oscillating wheel speed signals and attempts to dump excess momentum when there is none.

The model assumes a dipole Earth field, oriented so that COBE's orbit passes over the dipole axis at the northernmost point at t = 0. The field components in the body frame depend primarily on the yaw angle. For simplicity, they are calculated with pitch and roll assumed as zero. The reaction wheel speeds are

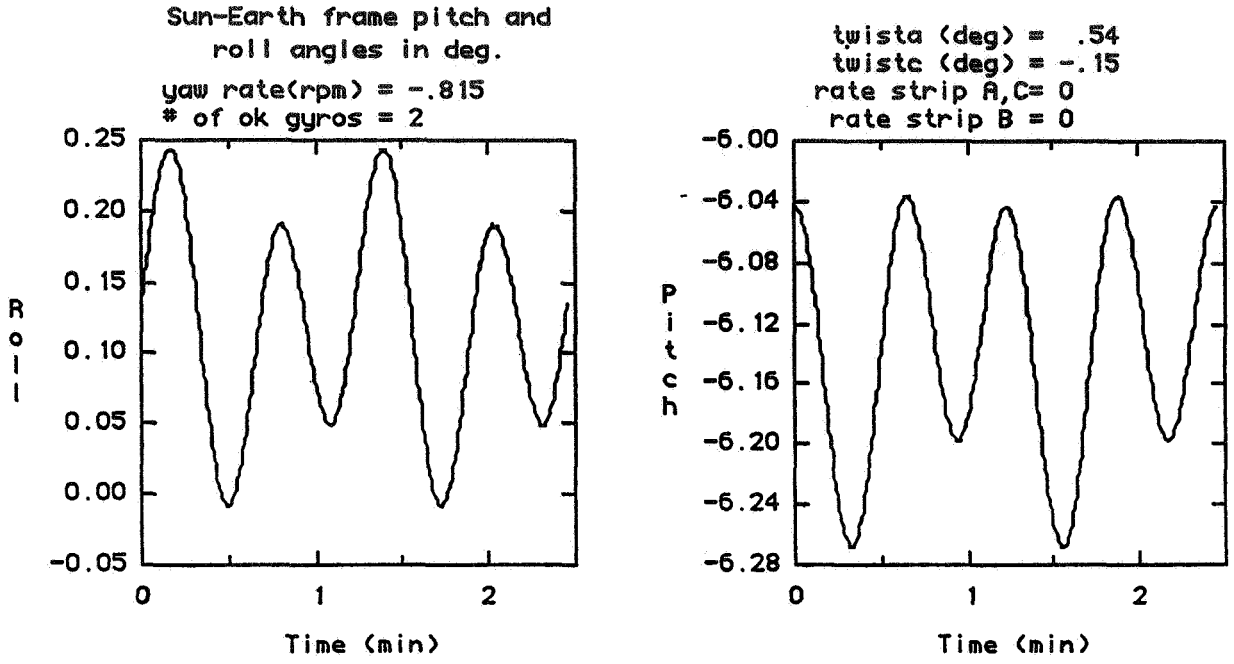


Figure 5. Roll and Pitch Angles Driven by ESA Misalignment and Failed B-Gyro

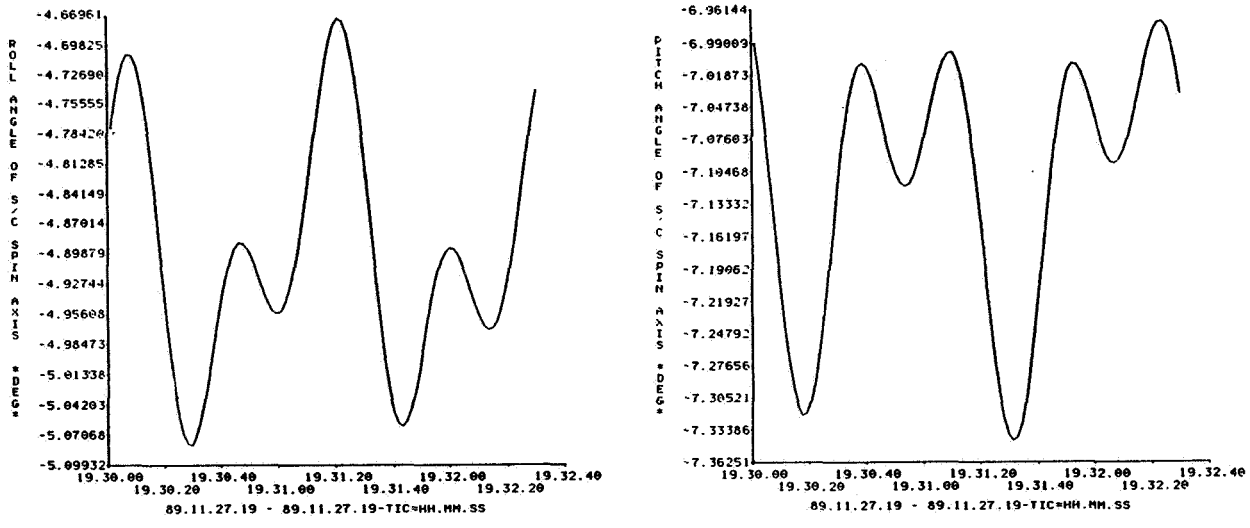


Figure 6. FADS Roll and Pitch Solutions

found by integrating the ACS control laws, using the high-frequency attitude solutions already obtained. In agreement with spacecraft data, the A-, B-, AX-, and BX-torquers are driven to their limits by the large oscillations of the A- and B-wheels.

Transforming the A-, B-, C-dipoles to the despun frame multiplies them by sine and cosine of the yaw angle. This beats with the reaction wheel frequencies (predominantly the spin rate). As a result, the pitch and roll components of the MMA dipole have both an offset and an oscillation at twice the spin rate. The MMA flips the sign in the Southern hemisphere. Multiplying by the X_C-component of the Earth's magnetic field and spin-averaging yields torques that vary as the absolute value of cos (ωt).

The final step is to put the MMA torques into Equation (12) and solve for the attitude. The spin-averaged torques are crudely approximated to be

$$T_r = -0.9 \times 10^{-3} |\cos \omega t| \ (\text{ft} - \text{lb}) \tag{27}$$

$$T_p = -0.2 \times 10^{-3} - 1.1 \times 10^{-3} |\cos \omega t| \ (\text{ft} - \text{lb}) \tag{28}$$

Solving for $\Delta\bar{\phi}$, one finds terms with single and double poles at s = 0, and single poles at s = $\pm i\omega$ and s = $\pm i2n\omega$ (n running from one to infinity). The amplitudes drop off rapidly with n; only the n = 1 terms are kept. They drive the system at twice the orbital frequency.

Figure 7 shows two orbits of the pitch motion driven by pitch anomaly, gravity gradient, and MMA torques. The spacecraft data are shown in Figure 8. The agreement is not good, although the model does give the large amplitude and two unequal peaks. The shape and amplitude are fairly sensitive to the parameter used in Equation (27), which in turn is sensitive to the details of the model (for example, inclusion of ESA misalignments and second-order gyro effects).

The solution could be improved by using a more realistic magnetic field model and by taking proper account of the attitude when calculating the field components in the body frame. Using the true attitude rather than the nominal values would also improve the gravity gradient results. This could be important because the gravity gradient natural period is approximately 93 minutes, comparable to the orbital period.

With the absence of any other large perturbations, it is concluded that the MMA is the immediate cause of the low-frequency pitch motion. In turn, the MMA is responding primarily to the inconsistent pitchback signals due to the failed B-gyro. It is suggested that commanding the same pitchback angle on all three axes, as in Equation (24), could significantly improve spacecraft performance.

ACKNOWLEDGMENT

The authors would like to thank Dr. Ming-San Chen for his many contributions to improving the performance of the COBE software.

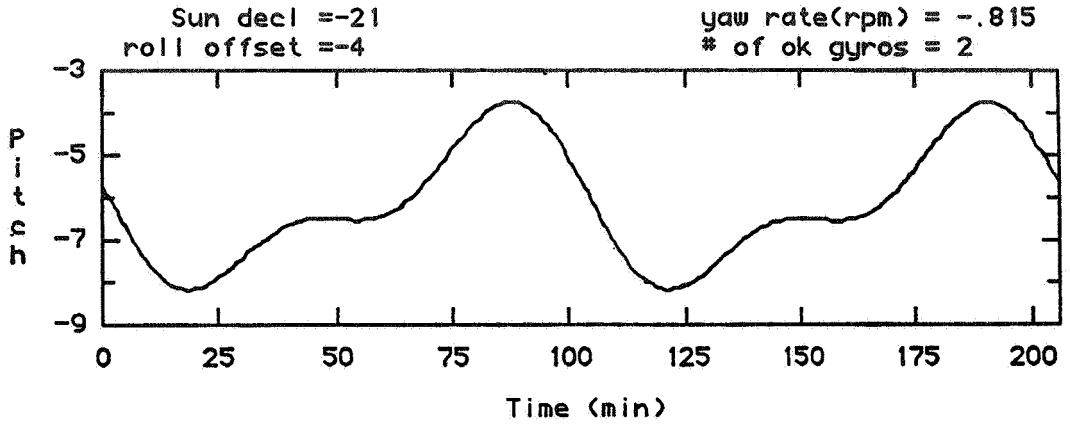


Figure 7. Pitch Motion Driven by Pitch Anomaly, Gravity Gradient, and MMA Torques

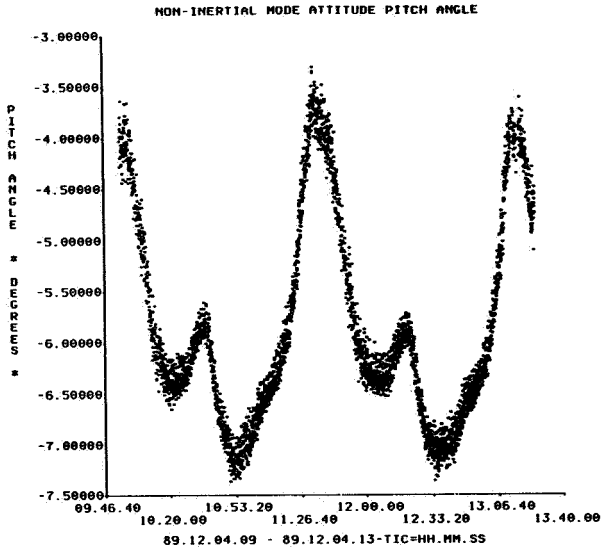


Figure 8. Two-Orbit CADS Pitch Solution

REFERENCES

1. Computer Sciences Corporation (CSC), CSC/TM-89/6007, *Cosmic Background Explorer (COBE) Attitude Control System (ACS)*, J. Sedlak, E. Scheidker, and K. Hall, September 1989

2. --, CSC/TM-86/6033, Mission Report 86002, *Analysis of the Diffuse Infrared Background Experiment (DIRBE) for COBE Attitude Determination*, E. Scheidker and K. Hall, 1986

3. --, CSC/TM-87/6032, *Star Catalogs Generated for Support of the Diffuse Infrared Background Experiment (DIRBE) on the Cosmic Background Explorer (COBE)*, S. McLaughlin and J. Seebach, May 1987

4. --, CSC/TR-87/6004, *Cosmic Background Explorer (COBE) Flight Dynamics Support System (FDSS) Specifications, II. Functional Specifications*, K. Hall, June 1987

5. --, CSC/SD-82/6013, *Earth Radiation Budget Satellite (ERBS) Attitude Ground Support System (AGSS) Functional Specifications and Requirements*, G. Nair, September 1982

6. --, CSC/TR-75, 6001, *Multisatellite Attitude Determination/Optical Aspect Bias Determination (MSAD/OABIAS) System Description and Operating Guide, Volume 1, Introduction and Analysis*, M. Joseph et al., April 1975

7. S. D. Conte and C. deBoor, *Elementary Numerical Analysis*, New York: McGraw-Hill Book Company, 1980

8. D. Chu, "COBE Nonspinning Attitude Propagation" (paper presented at the 1989 Flight Mechanics/Estimation Theory Symposium, Goddard Space Flight Center, Greenbelt, Maryland, May 23, 1989)

9. B. Bromberg and D. Croft, "On Orbit Attitude Control of the Cosmic Background Explorer (COBE)" (paper presented at Annual Rocky Mountain Guidance and Control Conference, American Astronautical Society (AAS), Keystone, Colorado, February 2-6, 1985)

ACCURACY OF THE ERBS DEFINITIVE ATTITUDE DETERMINATION SYSTEM IN THE PRESENCE OF PROPAGATION NOISE*

D. Chu and E. Harvie
Computer Sciences Corporation

ABSTRACT

Definitive attitude solutions are supposed to be the most accurate possible. For the Earth Radiation Budget Satellite (ERBS), this has been accomplished by using gyro rates to transform many nonsimultaneous observations to a common time point and then averaging to reduce the effects of observation noise. Rate quality is critical to realizing improved accuracy with this method. Gyro deterioration, which shows up as large observation residuals and discontinuities between contiguous batch solutions, now discourages using the batch approach for ERBS. To address this problem, a simple Kalman filter is tried in place of the batch estimator. The filter works well as long as the attitude is completely observable. During periods without Sun coverage, however, the extrapolated yaw may diverge and then change abruptly when the Sun returns to the sensor field of view. Causes of this behavior are discussed, and some solutions are tried that address the observability aspect of the problem.

*This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, Contract NAS 5-31500.

1. INTRODUCTION

Definitive attitude solutions are to be the most accurate possible and are to be continuous over the entire mission. Under the best conditions, this can be difficult both technically and operationally (Reference 1). In this paper, we consider one of the technical problems—that caused by attitude propagation error.

As a spacecraft gets old, its gyros may become noisy. This has been the case for the Earth Radiation Budget Satellite (ERBS). Without accurate angular velocity measurements, its definitive attitude determination system has been unable to provide continuous solutions or fit all observations.

To remedy this, the batch estimator is replaced by a Kalman filter that accounts for the gyro noise. The results obtained are described, and recommendations are made for further work to improve the approach.

2. ERBS BACKGROUND

ERBS is an Earth-pointing spacecraft in a 57-degree (deg) inclination, nearly circular orbit with a period of 97 minutes. It has two horizon sensors for measuring geodetic pitch and roll plus two Sun sensors whose alpha angles measure yaw. In addition, ERBS has a three-axis magnetometer. The horizon sensors are accurate to 0.5 deg, due primarily to horizon radiance effects. The Sun sensor is accurate to 0.05 deg, and the magnetometer is accurate to 3 deg. Because of its lower accuracy, the magnetometer is not used for definitive attitude determination. ERBS also has two redundant three-axis gyros for measuring its angular velocity. Both gyros have now partially failed. The first failure occurred on a pitch-axis gyro in August 1986, 22 months after launch. The second gyro failed on a pitch axis as well in July 1988. At present, only roll and yaw rates are available.

While the horizon sensors always see the Earth, the Sun sensors see the Sun for only part of each orbit. This means that attitude is not completely observable at all times, and good angular velocity measurements are needed to provide a complete history of the attitude. This incomplete observability also necessitates special care in dividing up each day's data for computing attitude.

3. DEFINITIVE ATTITUDE DETERMINATION

The ERBS definitive attitude system has three parts that determine the effect of propagation error—the estimator, the segmenter, and the smoother (Reference 2). The current batch least squares estimator updates its previous estimate of the state (epoch attitude, gyro bias, and scale factor errors) using all the observations at once. Sensor observations from different times are in effect propagated back to one epoch time, where they are averaged to reduce the effect of measurement noise.

If propagation were perfect, any number of observations could be so transformed, giving a solution of unlimited accuracy. In practice, however, propagation does add uncertainty,

which increases with the time between the epoch and the observation times. Thus, the residuals, or differences between the observed and predicted values, reflect both sensor noise, which is stationary, and propagation error, whose uncertainty grows with time.

If these residuals grow linearly with time from epoch or in proportion to the gyro output rates, they can be reduced by solving for rate bias and scale factor corrections along with the epoch attitude. Since bias errors grow with time, knowing the attitude at both ends of the batch affords the best opportunity for estimating biases. For this reason, the segmenter divides the day's data so that there is Sun sensor coverage at the start and end of every batch. This ensures that the rate biases can be accurately determined.

The smoother also serves to reduce the evidence, if not the effect, of propagation error. Because not all gyro errors are eliminated by bias and scale factor corrections, the time dependence of the attitude history may still be incorrect. No matter what the choice of epoch attitude, the solution cannot be correct over the entire batch. This shows up in the definitive attitude as discontinuities between batches. The smoother discards points around the junction and replaces them with values obtained by linearly interpolating between the new end points as in Figure 1. This is done without regard to sensor observations, and may not significantly improve the solution.

The batch estimator, by ignoring this random part of propagation error, creates a problem that the segmenter cannot fix and the smoother simply covers up.

4. MODELING PROPAGATION ERROR

Although propagation error can cause batch junction discontinuities, it remains to be shown that the ERBS gyro deterioration actually produces such discontinuities. To make this connection, a simplified Farrenkopf gyro noise model is used to predict discontinuity size based on the observations of gyro noise. These predictions are then shown to track the daily average pitch, roll, and yaw junction discontinuities.

Unlike the white noise seen in the sensor observations, errors in the propagated attitude tend to change continuously. For this reason, the propagated angle (θ) and its error ($\Delta\theta$) are modeled as a random process. The model in Figure 2 is a linear second-order system with four zero mean value sources of error:

- Float torque derivative noise (n_u)
- Float torque noise (n_v)
- Initial attitude error ($\Delta\theta(0)$)
- Drift rate bias (b)

The float torque derivative noise is integrated to give a random walk that is added to the true angular rate (ω). A constant, but imprecisely known bias and the float torque noise are also added at that point. At the output, another constant error is added representing the initial or epoch attitude error.

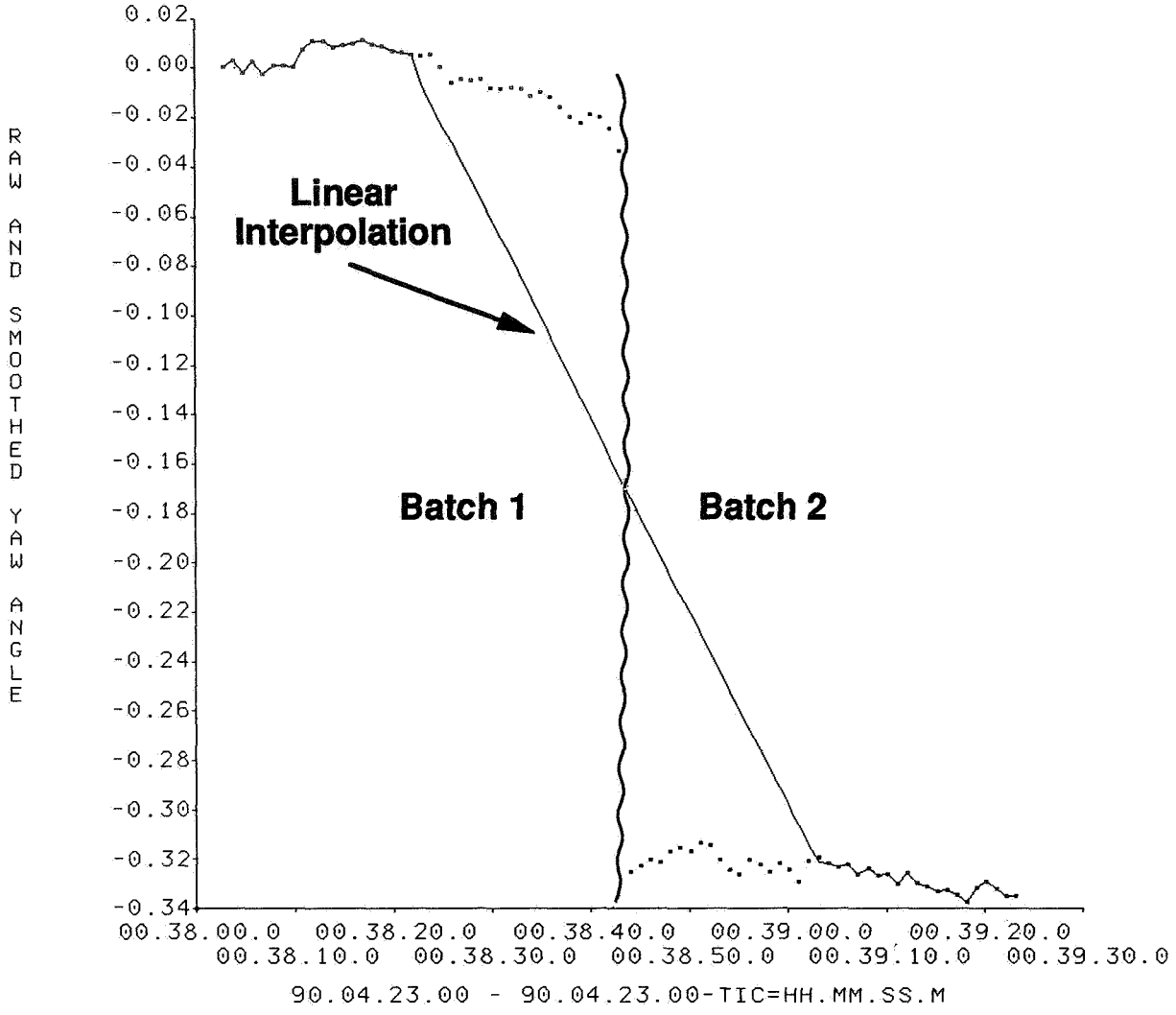


Figure 1. Smoothed Junction Discontinuity

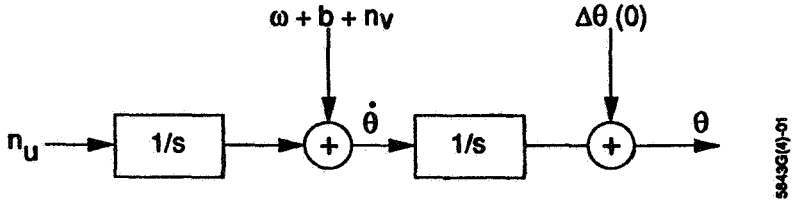


Figure 2. Gyro Noise Model

Because the four error sources have zero mean value, the expected propagation error is zero mean as well. The variance of the error, however, grows with propagation time (t) according to the expression

$$\sigma^2_{\Delta\theta}(t) = \sigma^2_{\Delta\theta}(0) + \sigma^2_v t + \sigma^2_b t^2 + \sigma^2_u t^3/3 \quad (1)$$

The individual error source variances are defined as follows where E is the expectation operator and δ is the Dirac delta function.

$$E[b^2] = \sigma^2_b \quad (2)$$

$$E[\Delta\theta^2(0)] = \sigma^2_{\Delta\theta}(0) \quad (3)$$

$$E[n_v(t)\, n_v(\tau)] = \sigma^2_v\, \delta(t - \tau) \quad (4)$$

$$E[n_u(t)\, n_u(\tau)] = \sigma^2_u\, \delta(t - \tau) \quad (5)$$

Standard deviations of the ERBS gyro rates were compiled over the course of the mission and provide an estimate of the gyro noise. Figure 3 shows this history of increasing noise.

These standard deviations for the sampled angular velocity ($\sigma_{\dot{\Delta\theta}}$) give an estimate of the float torque noise standard deviation (σ_v) that can be used to predict a typical batch junction discontinuity. Because the sampled rates are actually averaged over the sampling interval (T), their variance is smaller than the float torque noise variance by a factor of 1/T.

$$\sigma^2_{\dot{\Delta\theta}} = \sigma^2_v/T \quad (6)$$

Since T is 1 second, the numerical values of $\sigma_{\dot{\Delta\theta}}$ and σ_v are equal, but their units differ as they should.

5. EVIDENCE OF GYRO NOISE

Figure 4 plots daily average junction discontinuities and observation residuals against corresponding gyro rate standard deviations. In all cases, the discontinuities and residuals increase with increasing noise.

One thing that requires explanation here is the very steep slope of the yaw and Sun A curves. This is due to the low noise on the yaw gyro and the coupling of propagation

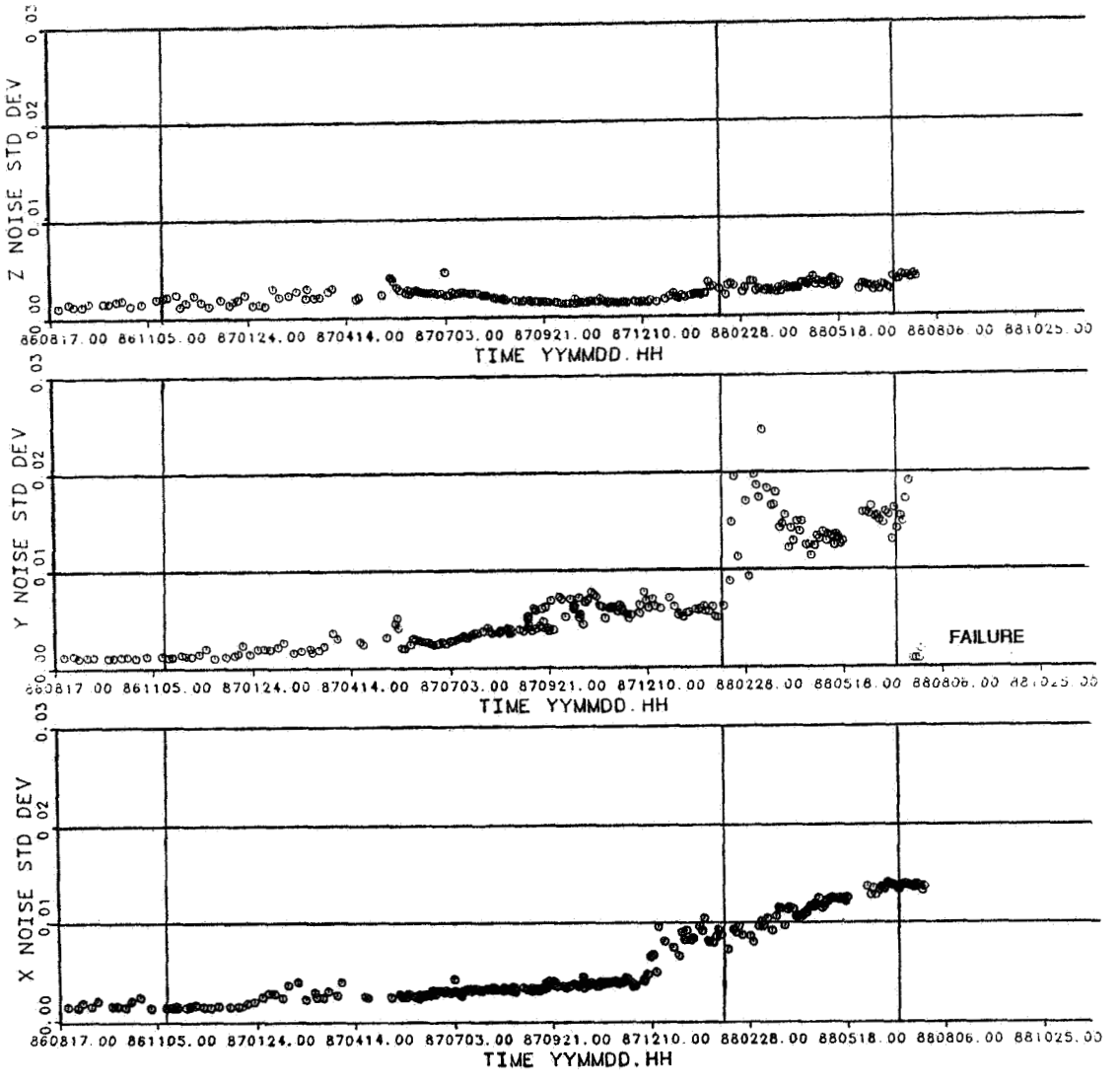
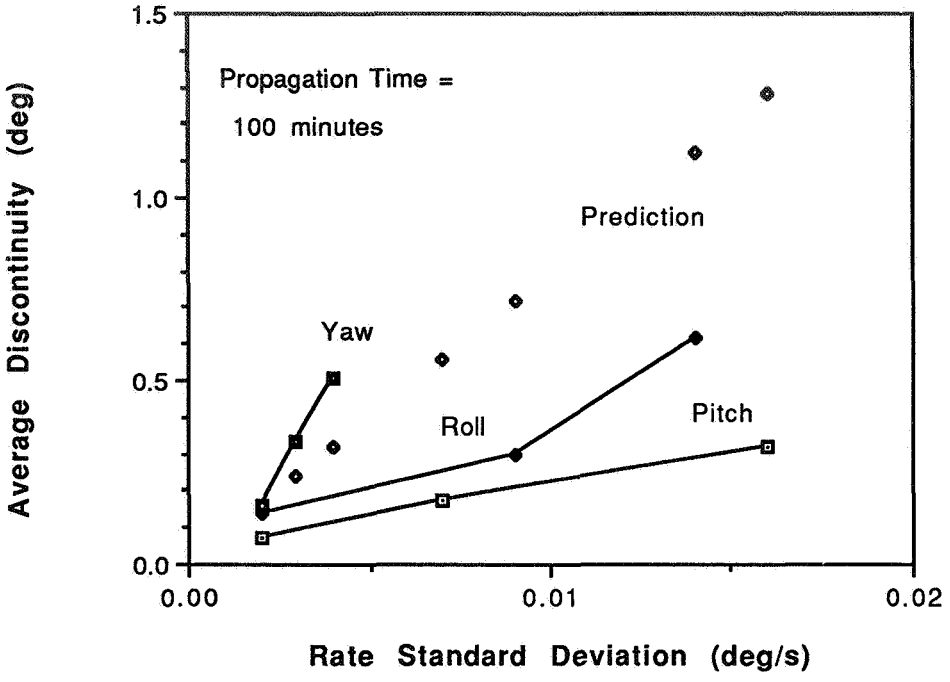


Figure 3. Gyro Rate Standard Deviations

errors on the roll and yaw axes. The large roll gyro noise causes yaw as well as roll propagation errors. Plotting yaw error against yaw gyro noise alone is therefore misleading. A better way to show the dependence might be to combine the roll and yaw into a single curve.

The predicted values of junction discontinuity are based on Equation 1 and use the gyro rate standard deviation for σ_v. Although inaccurate, the predictions are of the same order of magnitude as the observed discontinuities. Also, the rate standard deviation is an overestimate of the gyro noise since it includes the actual angular acceleration of the spacecraft. With a smaller assumed value for σ_v, the predictions would be closer to the observed discontinuities.



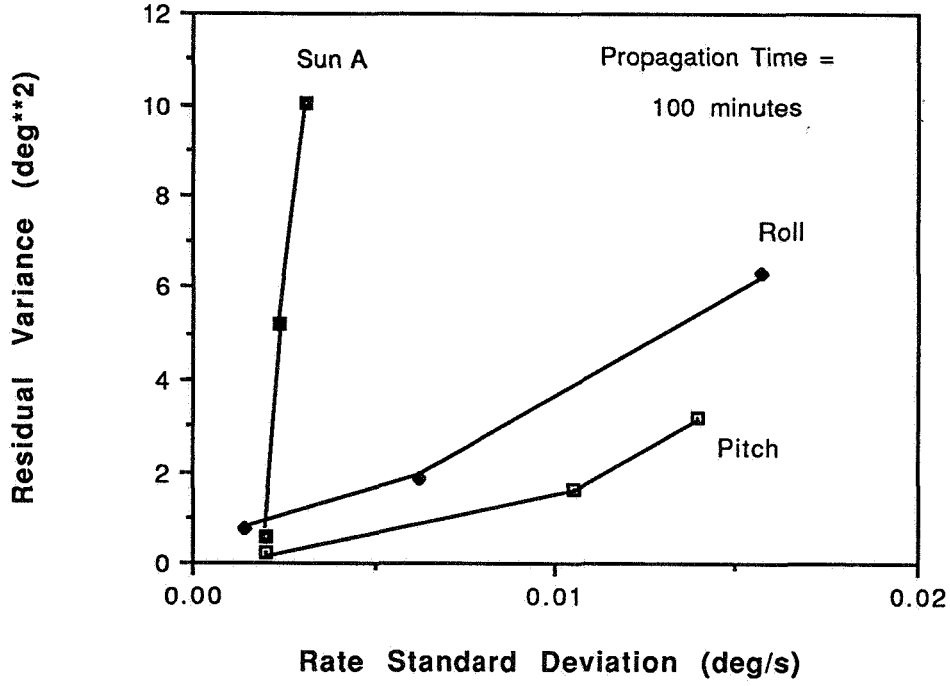


Figure 4. Discontinuities and Residuals

6. THE SIMPLE KALMAN FILTER

The fact that pitch should be so hard to estimate illustrates one weakness of the batch estimator. The reason it is so susceptible to propagation error is that it accepts the angular rates almost at face value. After correcting for biases and scale factors, the propagation is assumed to be perfect and observations far from the epoch are given equal weight with those close to the epoch.

The Kalman filter differs from the batch estimator in that it estimates the attitude at each time point rather than at an epoch time (Reference 7). It also gives greater weight to observations made close to the solution time. This seemed to be the answer to the propagation error/discontinuity problem. Old observations would be forgotten, giving the solution more freedom to follow recent observations. Even more, solutions could be made continuous at batch boundaries by starting with the previous batch solution and covariance.

Although different in approach from the batch estimator, the filter makes many of the same computations. For a very simple filter estimating attitude and constant bias, only three major modifications are required to the ERBS batch estimator.

1. Change the partial derivatives

   a. In the batch estimator, the partial derivatives of the current observation residual vector $\Delta\underline{y}_j$

$$\Delta\underline{y}_j = \underline{y}_j \text{ (observed)} - \underline{y}_j \text{ (computed)} \tag{7}$$

   with respect to the current attitude error vector $\underline{a}_j$ are postmultiplied by the epoch-to-current-time propagation matrix $\phi(t_j, t_0)$. This multiplication is omitted in the filter

$$\frac{\partial \underline{y}_j}{\partial \underline{a}_0} = \frac{\partial \underline{y}_j}{\partial \underline{a}_j}\,\phi(t_j, t_0) \rightarrow \frac{\partial \underline{y}_j}{\partial \underline{a}_j} \tag{8}$$

   b. Filtered gyro biases $\underline{b}_j$ act only over individual time steps before being updated with the bias error vector $\underline{\alpha}_j$. Rather than accumulate derivatives of the current attitude error with respect to gyro biases from epoch to the current time, as in the batch estimator, the filter computes them only for the most recent time step.

$$\frac{\partial \underline{a}_j}{\partial \alpha_j} = \psi(t_j, t_{j-1}) + \phi(t_j, t_{j-1})\,\frac{\partial \underline{a}_{j-1}}{\partial \alpha_{j-1}} \rightarrow \psi(t_j, t_{j-1}) \tag{9}$$

The matrix $\psi(t_j, t_{j-1})$ comes from integrating the propagation matrix and assumes the bias is constant over this interval.

$$\psi(t_j, t_{j-1}) = \int_{t_{j-1}}^{t_j} \phi(t_j, \tau)\, d\tau \tag{10}$$

c. To simplify notation, the combination of the attitude and gyro bias error vectors are referred to as the state $\underline{x}_j$

$$\underline{x}_j^T = [\underline{a}_j \vdots \underline{\alpha}_j] \tag{11}$$

and the partial derivatives of the observation residuals with respect to the state is called F_j

$$F_j = \left(\frac{\partial \underline{y}}{\partial \underline{a}} \vdots \frac{\partial \underline{y}}{\partial \underline{\alpha}}\right)_j = \left(\frac{\partial \underline{y}}{\partial \underline{x}}\right)_j \tag{12}$$

2. Update the state at each time step—The batch estimator weights, transforms, and accumulates the observation residuals over the entire batch to give

$$FTWY \equiv \sum_{i=1}^{N} F_i^T W \Delta y_i \tag{13}$$

The observation weight matrix W is diagonal with elements equal to the reciprocal of the observation variances

$$W = \begin{pmatrix} 1/\sigma_{y1}^2 & & 0 \\ & \ddots & \\ 0 & & 1/\sigma_{ym}^2 \end{pmatrix} \tag{14}$$

The batch estimator also accumulates the normal matrix, which is the inverse of the covariance matrix P.

$$\sum_{i=1}^{N} F_i^T W F_i + W_0 \equiv P_N^{-1} \tag{15}$$

The a priori estimate weight matrix W_0 is also diagonal and has as its elements the reciprocal of the a priori state variable variances

$$W_0 = \begin{pmatrix} 1/\sigma_{y1}^2 & & 0 \\ & \ddots & \\ 0 & & 1/\sigma_{yn}^2 \end{pmatrix} \tag{16}$$

In batch estimation, the epoch state estimate is updated after all the observation residuals have been accumulated.

$$\Delta \underline{x}_0 = (FTWF)^{-1} (FTWY) = (P_N) (FTWY) \tag{17}$$

The filter makes this correction at each time step using the weighted transformed observation residuals for that instant and the accumulated covariance matrix.

$$\Delta \underline{x}_j = (P_j) (F_j^T W \Delta \underline{y}_j) \tag{18}$$

3. Propagate the covariance matrix—The batch estimator propagates the attitude and predicts observations at each time step. The filter propagates the covariance matrix as well. This is done by pre- and postmultiplying the covariance matrix by the incremental state propagation matrix $\Phi(t_j, t_{j-1})$ and its transpose.

$$P_j = \Phi(t_j, t_{j-1}) P_{j-1} \Phi^T(t_j, t_{j-1}) + Q \tag{19}$$

A contribution Q due to the propagation error is also added at this point (Reference 3).

$$Q = \begin{bmatrix} \sigma_v^2 (t_j - t_{j-1}) + \frac{1}{3} \sigma_u^2 (t_j - t_{j-1})^3 & -\frac{1}{2} \sigma_u^2 (t_j - t_{j-1})^2 \\ -\frac{1}{2} \sigma_u^2 (t_j - t_{j-1})^2 & \sigma_u^2 (t_j - t_{j-1}) \end{bmatrix} \tag{20}$$

The state propagation matrix includes the attitude propagation matrix as its upper left block

$$\Phi(t_j, t_{j-1}) = \left(\begin{array}{c|c} \phi(t_j, t_{j-1}) & \psi(t_j, t_{j-1}) \\ \hline 0 & I_{3\times 3} \end{array} \right) \tag{21}$$

The attitude parameterization chosen for the filter is the attitude error vector. This representation fits in with the existing batch differential correction and avoids complications from unit length and orthonormality constraints. To update the estimated geocentric inertial-to-body attitude matrix $A_{B/I}$, the error vector is converted to a rotation matrix $R(\underline{a}_j)$ which is used to premultiply the old value of $A_{B/I}$.

$$A_{B/I}(t_j) = R(\underline{a}_j)\, A_{B/I}(t_{j-1}) \tag{22}$$

The gyro bias corrections $\underline{a}_j$ are simply added to the previous bias values

$$\underline{b}_j = \underline{b}_{j-1} + \underline{a}_j \tag{23}$$

7. COMPARING THE BATCH ESTIMATOR AND THE FILTER

After being implemented, the filter was used with low and high noise gyro rate and was compared to the results from the batch estimator. For timespans having Sun sensor coverage, the attitude is completely observable, and the filter provides smaller root-mean-square (rms) observation residuals. This is shown in Table 1. There is no chance of junction discontinuities because the filter starts with the last attitude of the preceding timespan.

Table 1. RMS Residuals—Complete Observability

Observation Residual Variance (deg ^{2})

Low Noise	Roll	Pitch	Sun A	Overall
Filter	0.16	0.30	0.03	0.87
Batch	0.10	0.31	0.01	0.87
High Noise	**Roll**	**Pitch**	**Sun A**	**Overall**
Filter	0.05	0.13	0.00	0.17
Batch	0.85	0.73	0.28	1.07

Over timespans without continuous Sun coverage, the filter can diverge and cause jumps even greater than those between batches. An example of such a filter discontinuity is provided in Figure 5. There, the yaw diverges until the Sun comes back into view. The size of the jump depends strongly on the actual gyro noise and the assumed gyro noise, i.e., tuning of the filter.

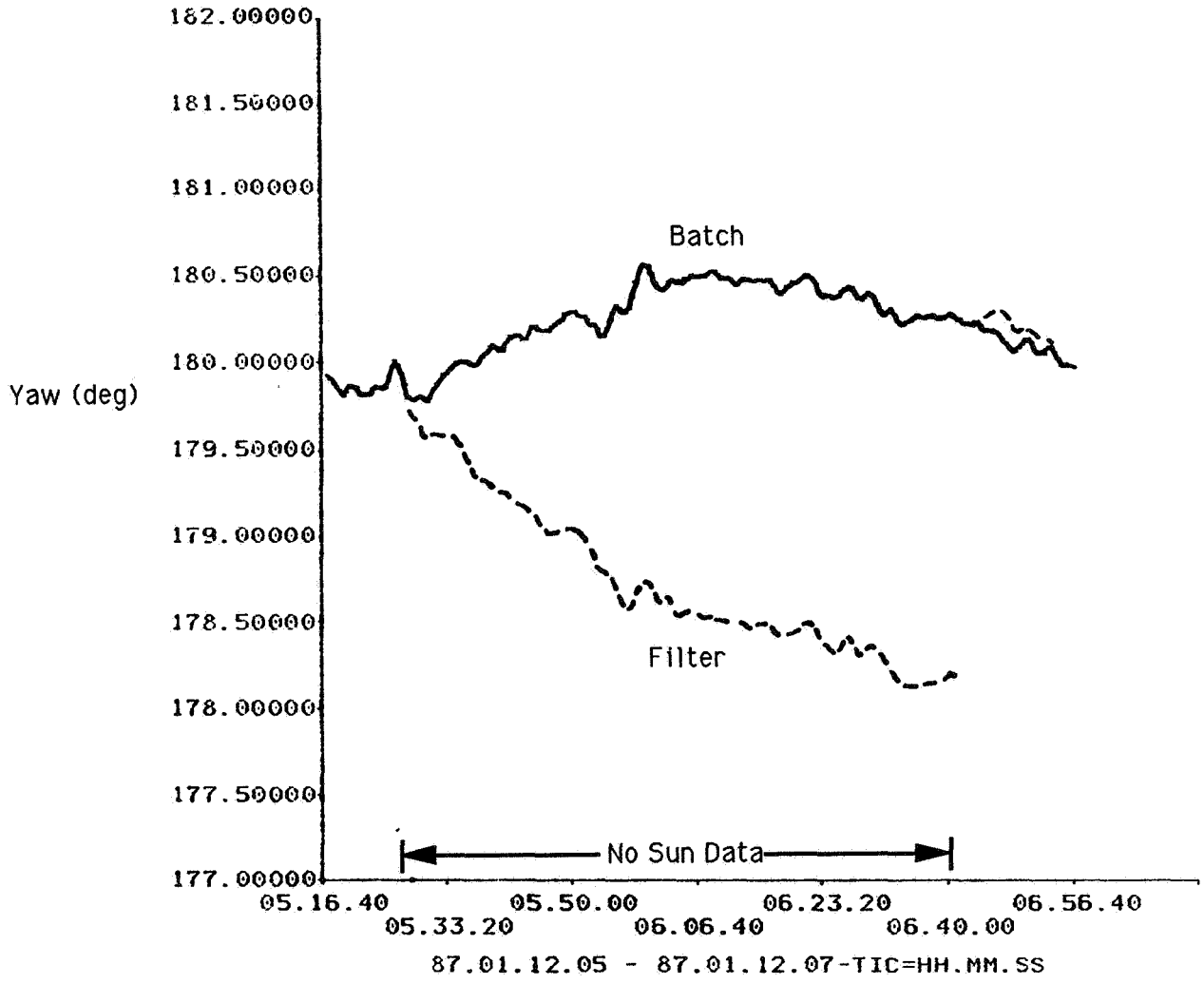


Figure 5. Drifting of Filtered Yaw

As shown in Table 2, the residuals are still much smaller for the filter than for the batch estimator. This is in spite of the discontinuity. The reason is that the filter fits all the observations better except for the first few following the start of Sun coverage.

Table 2. RMS Residuals—Incomplete Observability

Observation Residual Variance (deg ^{2})

Low Noise	Roll	Pitch	Sun A	Overall
Filter	0.09	0.46	1.26	0.40
Batch	0.31	2.58	0.14	1.50
High Noise	**Roll**	**Pitch**	**Sun A**	**Overall**
Filter	0.04	0.08	1.45	0.19
Batch	29.66	9.67	0.66	17.37

8. DIVERGENCE REMEDIES

The batch estimator, despite its problems with junction discontinuities and observation residuals, does not drift during periods of incomplete attitude observability. This is because it has access to data from the entire batch at once. Thus, the batch estimator can choose a bias satisfying the observations at both the beginning and the end of the batch. The filter, on the other hand, knows only what has been seen up to the current time and does not have this advantage.

To improve filter performance during periods without Sun, the magnetometer can be used along with the Sun and Earth sensors. Although it is the least accurate sensor, the magnetometer provides some measure of the yaw at all times. As shown in Figure 6, this prevents divergence but does not provide a very accurate definitive attitude solution.

An alternative to improving the observability is to improve the propagation. To do so, the batch estimator can be used to compute a gyro bias over the period lacking Sun coverage. This bias can then be used in the filter to reduce the divergence, as shown in Figure 7.

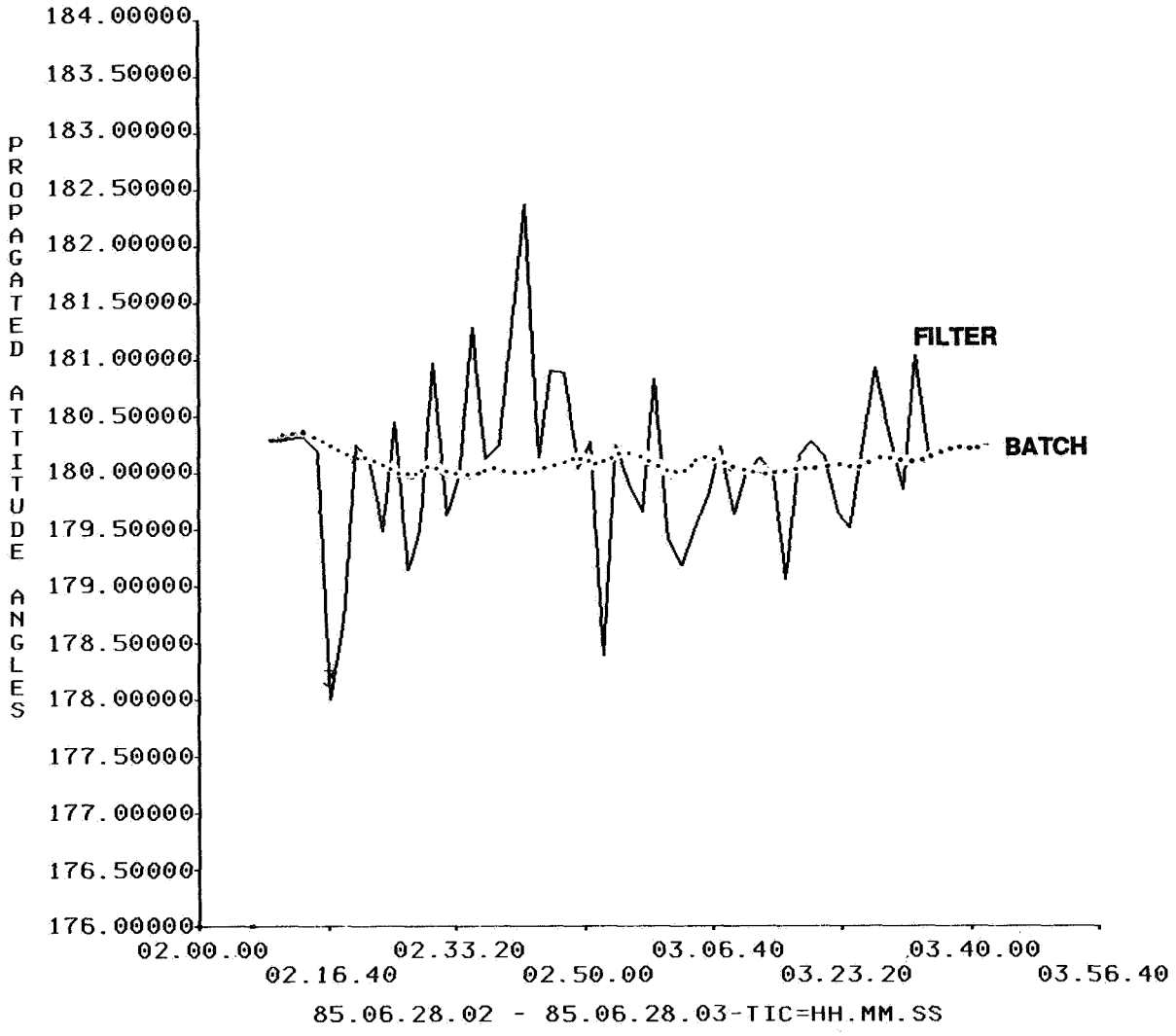


Figure 6. Yaw Using Magnetometer

A variant on this "two-pass" processing is to average the forward-filtered solution with a backward-filtered solution obtained by starting from the end of the timespan and letting time run backward. This "filter-smoother" can still diverge toward the middle of the Sun coverage gap, but does not change abruptly the way a simple filter can.

The approach that should, perhaps, have been tried first is to "tune" the gyro noise parameter values to reduce the divergence. The values used here for the Q-matrix were based on the noise estimated from the raw gyro rates. It was noted that these assumed values strongly affect the filter divergence, but no effort was made to find values that worked well in all cases. Judging from the satisfactory performance of the ERBS gyro-compass, which uses only the Earth sensors and gyro, it should be possible to find such values. Perhaps the gyro noise levels used here were too large and so caused the filter to forget old observations too quickly.

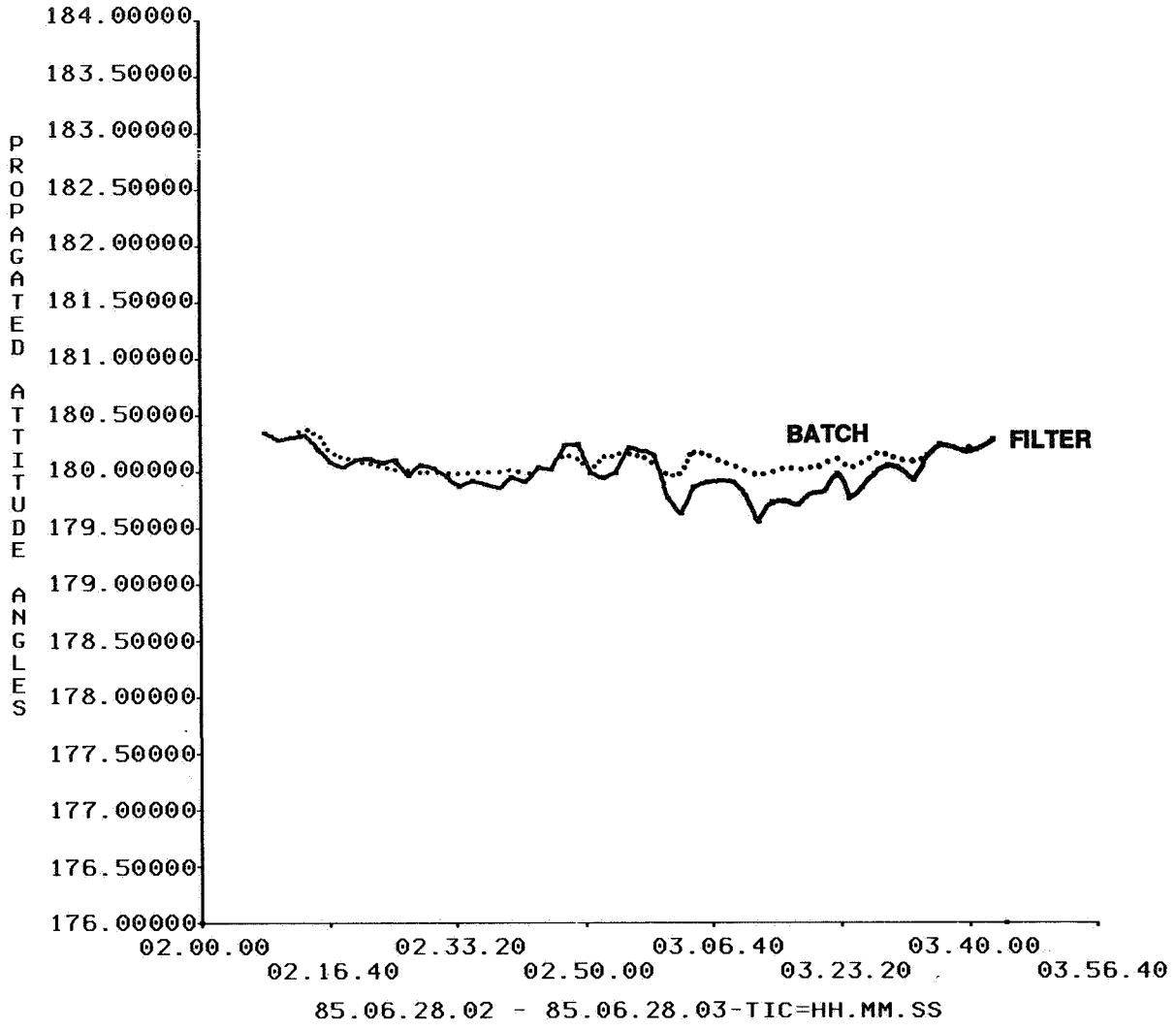


Figure 7. Yaw With Batch Bias

9. SUMMARY

This paper uses the ERBS definitive attitude determination system to demonstrate two things: first, that gyro noise affects batch attitude accuracy, as seen in the observation residuals and batch junction discontinuities; and second, that Kalman filtering can reduce those residuals and eliminate the junction discontinuities as well. What was also demonstrated, albeit inadvertently, was that during times of incomplete attitude observability, the filtered attitude may diverge. Patchwork remedies were tried, such as using additional sensors or future observations, but the basic cause of the problem seems to have been the "tuning" of the filter. Had a smaller value been assumed for the gyro noise, the attitude would not have diverged as quickly as it did. In spite of the present incomplete success,

Kalman filtering still promises accuracy equal to or better than that of the batch estimator once the tuning process is mastered.

REFERENCES

1. Computer Sciences Corporation, CSC/TR-89/6021, *A Study of Definitive Attitude Determination*, M. Lee and J. Rowe, September 1989

2. --, CSC/SD-82/6013, *Earth Radiation Budget Satellite (ERBS) Attitude Ground Support System (AGSS) Functional Specifications and Requirements*, G. Nair, September 1982

3. R. L. Farrenkopf, "Generalized Results for Precision Attitude Reference Systems Using Gyros," AIAA Paper 74-903, AIAA Mechanics and Control of Flight Conference, Anaheim, CA, August 1974

4. J. R. Wertz, *Spacecraft Attitude Determination and Control*, D. Reidel Publishing Co., 1978, p. 269

5. O. Filla, T. Z. Willard, and D. Chu, *In-Flight Estimation of Gyro Noise*, Flight Mechanics/Estimation Theory Symposium, NASA–Goddard Space Flight Center, May 1990

6. E. Harvie, D. Chu, and M. Woodard, "The Accuracy of Dynamic Propagation," Flight Mechanics/Estimation Theory Symposium, NASA–Goddard Space Flight Center, May 1990

7. E. J. Lefferts, F. L. Markley, and M. D. Shuster, "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance*, September–October 1982, vol. 5, no. 5, pp. 417-429

FLIGHT MECHANICS/ESTIMATION THEORY SYMPOSIUM

MAY 22-24, 1990

SESSION 2

An Accuracy Assessment of Magellan Very Long Baseline Interferometry

D. B. Engelhardt*, G. R. Kronschnabl*, and J. S. Border‡
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Abstract

Very Long Baseline Interferometry (VLBI) measurements of the Magellan spacecraft's angular position and velocity were made during July through September, 1989, during the spacecraft's heliocentric flight to Venus. The purpose of this data acquisition and reduction was to verify this data type for operational use before Magellan is inserted into Venus orbit, in August, 1990. The accuracy of these measurements are shown to be within 20 nanoradians in angular position, and within 5 picoradians/sec in angular velocity. The media effects and their calibrations are quantified; the wet fluctuating troposphere is the dominant source of measurement error for angular velocity. The charged particle effect is completely calibrated with S and X-Band dual-frequency calibrations. Increasing the accuracy of the Earth platform model parameters, by using VLBI-derived tracking station locations consistent with the planetary ephemeris frame, and by including high frequency Earth tidal terms in the Earth rotation model, add a few nanoradians improvement to the angular position measurements. Angular velocity measurements were insensitive to these Earth platform modelling improvements.

Introduction

The Magellan spacecraft, launched May 4, 1989, is currently on a heliocentric trajectory to Venus, with a scheduled arrival on August 10, 1990. Upon arrival, Magellan will be inserted into a nearly polar orbit in order to systematically map the surface of Venus using a Synthetic Aperature Radar (SAR). The spacecraft orbit will be determined from radiometric measurements of the spacecraft velocity; highly accurate orbit reconstruction and prediction are required to effectively command the SAR, to provide picture-element registration in the ground processing of the radar image data, and to interpret measurements from an on-board altimeter. The nominal orbit about Venus will have a relatively short orbital period (3.15 hours) and low periapsis altitude (250 km), with an eccentricity of 0.38. This orbit geometry poses a substantial challenge to navigators in determining the spacecraft orbit within the stringent accuracy requirements because (1) conventional, two-way, coherent Doppler measurements are less sensitive to certain orbit parameters, namely, the orientation of the orbit about the Earth-spacecraft line-of-sight [Wood, 1986], and (2) uncertain irregularities in the Venus gravity field will perturb the spacecraft over a large fraction of each orbit.

---

* Member of Magellan Navigation Team, Navigation Systems Section.

‡ Member of Tracking Systems and Applications Section.

To meet orbit determination accuracy requirements, measurements of the spacecraft velocity in the plane-of-the-sky will be made [Mohan, 1985]. These data will provide information on the orbit parameters poorly determined with Doppler measurements, as Doppler measures the spacecraft velocity in only the spacecraft line-of-sight direction. Very Long Baseline Interferometry (VLBI) is used to provide this measurement by correlating the spacecraft radio downlink received at two widely separated ground antennas [Melbourne, 1977].

This VLBI technique has been experimentally proven with Voyager, a deep space probe [Border et al., 1982], and with Pioneer 12, a Venus orbiter [Esposito et al., 1983], but never proven operationally. To verify that the VLBI system can meet stringent operational and accuracy requirements for a planetary orbiter, a 9 week test was conducted during the early heliocentric cruise portion of Magellan's trajectory to Venus. Spacecraft and quasar VLBI measurements were acquired and processed during this time period, in a manner similar to the planned VLBI data acquisition and reduction in the orbit phase. This paper describes this test and the assessment of the VLBI measurement precision about a reference trajectory. The accuracies of the reference trajectory, the media calibrations, and the Earth platform models are presented. Their contribution to the measurement accuracy is included in the presentation of the VLBI measurement accuracy.

Magellan Radiometric Measurements

Three types of radiometric measurements of the spacecraft state are made: 2-way coherent Doppler, Differential One-Way Range (DOR), and Differential One-way Doppler (DOD). Doppler is collected at a single station, and the DOR and DOD are VLBI measurements, made by simultaneously recording 10 minute scans of the spacecraft downlink signal at two widely separated tracking stations. NASA's Deep Space Network (DSN) tracking stations acquire and relay the data to the Jet Propulsion Laboratory in Pasadena, California. The DSN tracking stations are spaced around the globe in California, Spain, and Australia. The VLBI intercontinental baselines are between the Spain-California and California-Australia tracking stations. Figure 1 shows the DSN sites. Simplified Earth-spacecraft geometry is shown in Figure 2. This geometry is simplified by assuming a distant spacecraft, where the range, ρ, is much greater than the baseline length, B. These three data types are briefly described below.

Doppler

Two-way coherent Doppler measures the spacecraft's velocity component in the line-of-sight from the antenna to the spacecraft, as shown by $\dot{\rho}_1$ and $\dot{\rho}_2$, at tracking stations 1 and 2, in Figure 2. The tracking station generates and transmits a signal with a stable frequency, and the spacecraft sends a coherent version of the received signal back to the same station where the Doppler frequency shift is determined. This Doppler frequency shift is proportional to the spacecraft velocity component in the line-of-sight direction by the factor (c/2f), where f is the signal frequency and c is the speed of light. The line-of-sight direction is specified by $\hat{e}_1$ in Figure 2.

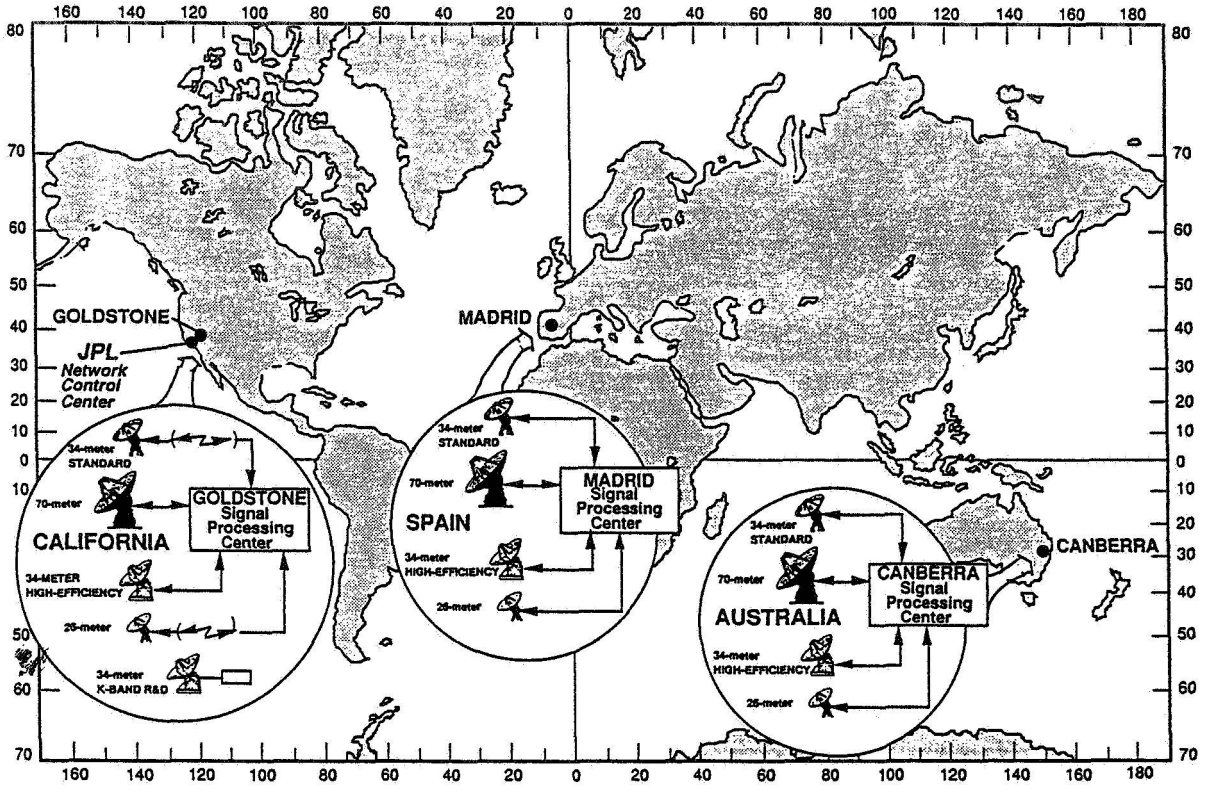


Figure 1. Deep Space Network Sites

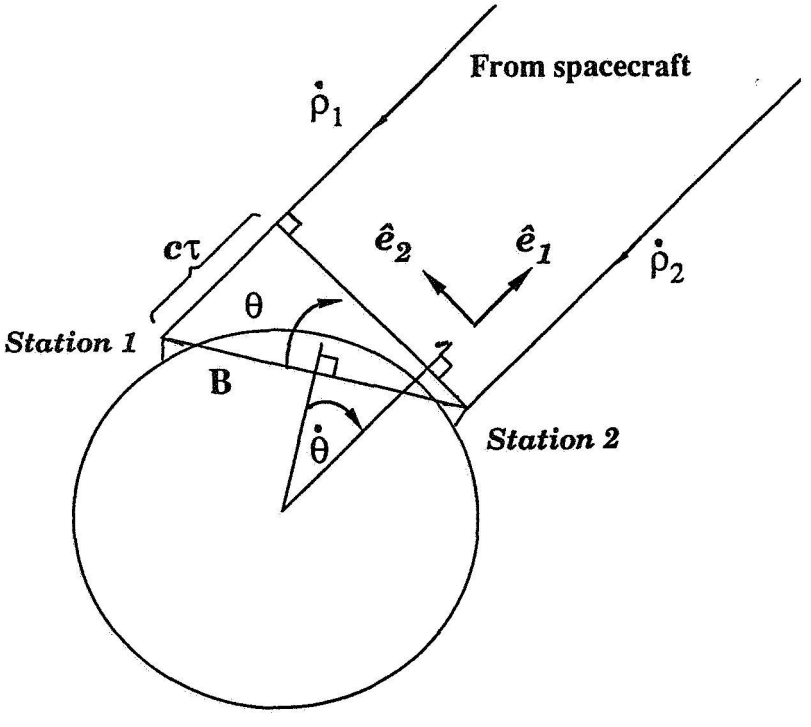


Figure 2. Simplified VLBI Geometry

DOR

Differential One-way Range (DOR) measures the geocentric angular position of the spacecraft with respect to the baseline; this quantity is shown by θ in Figure 2. This angle is determined by directly measuring τ, the delay of the spacecraft signal reception from station 1 to station 2. In Figure 2,

$$c\tau \cong B \sin\theta. \tag{1}$$

Given B, the basline length, θ is determined from the measurement of $c\tau$, the differential range between stations. τ is determined by differencing the phase of the signal received at each station [Border, 1972]. This delay is ambiguous by an integer number of cycles, however. Multiple downlink frequency tones are needed to provide the information to resolve this ambiguity. Harmonics of Magellan's telemetry subcarrier signal provide recordable side tones 2 to 35 MHz apart. The ambiguity in the 2 MHz bandwidth is c/2 Mhz/B, or 17 microradians, and is easily resolved from apriori knowledge. The 35 MHz bandwidth provides more measurement precision.

DOD

Differential One-way Doppler (DOD) measures the geocentric angular velocity of the spacecraft with respect to the baseline; this quantity is shown by $\dot{\theta}$ in Figure 2. This angular rate is determined by directly measuring $\dot{\tau}$, the time delay-rate of the spacecraft carrier signal from station 1 to station 2. Differentiating equation (1) with respect to time gives

$$c\dot{\tau} \cong B \cos\theta(\dot{\theta}). \tag{2}$$

$c\dot{\tau}$ is the differential velocity between stations, and the sought-after rate is:

$$\dot{\theta} \cong \frac{c\dot{\tau}}{B\cos\theta} \tag{3}$$

$\dot{\theta}$ is related to the spacecraft velocity, $\vec{V}_{s/c}$, by:

$$\dot{\theta} \cong \frac{(\vec{V}_{s/c} \cdot \hat{e}_2)}{R} + \text{Earth rotation of baseline} \tag{4}$$

where R is the Earth to spacecraft range, and $\hat{e}_2$ is the direction normal to the antenna-spacecraft line-of-sight, as shown in Figure 2. Assuming the earth rotation of the baseline is known, the measurement of $c\dot{\tau}$ is related to the spacecraft velocity, $\vec{V}_{s/c}$, by:

$$c\dot{\tau} \cong \frac{B\cos\theta}{R}(\vec{V}_{s/c} \cdot \hat{e}_2). \tag{5}$$

$\dot{\tau}$ is determined by rate of change of the carrier signal phase from each station.

DOR and DOD measurements can also be generated from the radio emissions of quasars for measurement calibration purposes. Quasar delay is obtained by signal cross-correlation [Thomas, 1972]. Since these quasar measurements are corrupted in almost the same way as spacecraft measurements are by tracking station location errors, media errors, and timekeeping errors, spacecraft and quasar measurements made in close temporal and spatial proximity are differenced to greatly reduce these major errors. These spacecraft-quasar differenced measurements are signified by ΔDOR and ΔDOD, or ΔVLBI in general.

The Magellan Trajectory

The radiometric data used for this accuracy analysis was collected from July 1, 1989 to August 11, 1989, early in Magellan's 15 month heliocentric cruise to Venus. The view of the spacecraft trajectory during this period, from the north ecliptic pole, is shown in Figure 3.

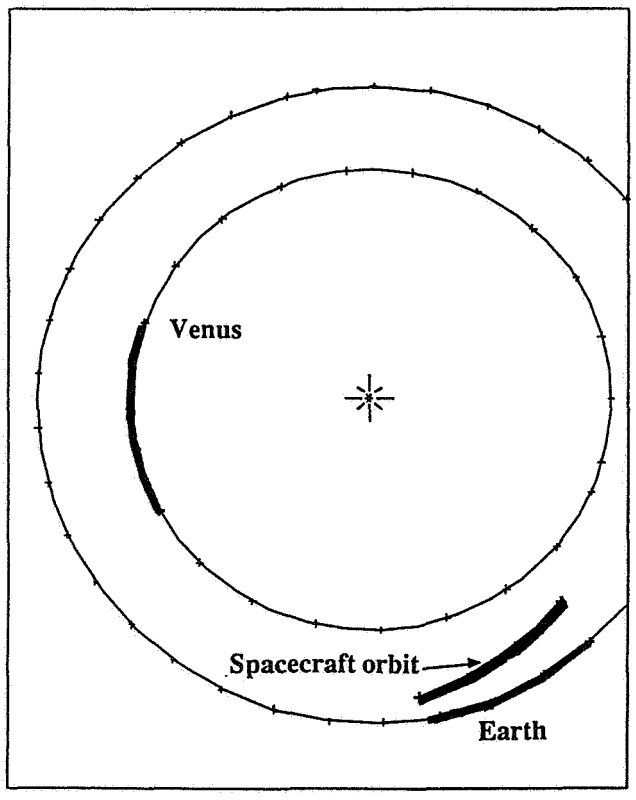


Figure 3. Trajectory View from North Ecliptic Pole

The VLBI measurement accuracy is quantified as an RMS value of measurement 'residuals'. A measurement residual is the difference of the acquired measurement and a

predicted measurement. The predicted measurement is computed using a reference spacecraft trajectory and models of the Earth platform, radio frequencies, and quasar locations. So the measurement RMS quantifies the data precision about a reference trajectory. This reference trajectory accuracy has been computed, so that the absolute measurement accuracy can be ascertained. A reference trajectory was determined from Doppler and ΔDOR data. From Equation (3), one can see that θ is an intermediate quantity in the computation of $\dot{\theta}$, given the measurement of $c\dot{\tau}$. To evaluate the ΔDOD data accuracy, then, the reference trajectory computation included ΔDOR data, to provide direct information of θ.

The trajectory accuracy is given in terms of uncertainty in geocentric spherical coordinates of the spacecraft on August 11, 1989 (at the end of the data arc). The right ascension, α, is measured with respect to the vernal equinox of 2000.0, and the declination, δ, is measured with respect to Earth's mean equator. These uncertainties take into account the measurement errors, as well as the known uncertainties in the station locations, spacecraft non-gravitational forces (i.e. solar radiation pressure), and quasar locations.* The trajectory uncertainties are given in Table 1 for orbit solutions using 3 different data sets: Doppler only, Doppler+ΔDOR, and Doppler+ΔDOD.

	Doppler Only	Doppler + ΔDOR	Doppler + ΔDOD
Range, (km)	77.1	73.9	73.1
α, (nanorads)	474.1	338.4	542.5
δ, (nanorads)	724.4	303.2	558.0
Range-rate, (mm/s)	0.196	0.148	0.277
$\dot{\alpha}$, (picorads/sec)	0.1448	0.1500	0.1410
$\dot{\delta}$, (picorads/sec)	0.5057	0.4833	0.2926

Table 1. Trajectory Uncertainty on August 11, 1989, computed with three data sets: Doppler only, Doppler+ΔDOR, and Doppler+ΔDOD.

The strength of each VLBI data type is apparent. The ΔDOR data brings in angular position information, reducing the angular position uncertainty by a factor of 2. The ΔDOR data arc was only 2 weeks long, compared to a 6 week arc of ΔDOD. For comparable data arcs, the ΔDOR would be expected to produce by far the most accurate measure of angular velocity during cruise. The ΔDOD data brings in angular velocity information, reducing the angular velocity uncertainty by a factor of 1.5.

* These uncertainties are: 20% in non-gravitational force values (solar radiation pressure and unbalanced spacecraft manuevers); 20 nanoradian radio/planetary frame-tie offset; station location errors of: 1.5 meters in distance from the pole, 1.0 meter in longitude, 10.0 meters in height above the equator; 10 cm relative station error; and 30 cm longitude error due to Earth rotation uncertainty.

Media Calibrations

This section describes the effects of the media on the VLBI measurements, and the calibrations used to minimize the measurement error due to the media. The success of the calibrations are shown with comparisons of uncalibrated and calibrated VLBI data residuals.

Radio signal propogation is delayed by charged particles in solar plasma and in the Earth's ionosphere, and by the Earth's troposphere. Only the corruption to the VLBI downlink signal needs to be calibrated because the uplink signal, being common to each of the two recorded downlink signals, cancels out in the correlation process. The charged particle delay is proportional to the Total Electron Content (TEC) along the signal path and inversely proportional to the signal frequency squared. The TEC changes with time, inducing an error in the delay-rate measurement.

Since the charged particle delay is frequency dependent, a calibration to the delay-rate measurement can be computed with a linear combination of two delay-rates recorded at separate frequencies, S and X band. The computation for X-Band delay-rate, in Hertz, is:

$$\text{Delay-rate correction} = \frac{\dot{\phi}_x - (f_x/f_s)\dot{\phi}_s}{1-(f_x/f_s)^2} \tag{6}$$

where $\dot{\phi}_x$ and $\dot{\phi}_s$ are the measured delay rates at X and S bands, and f_s and f_x are the S and X band frequencies [Wolff, 1985].

The average daytime flucuation in the X-band zenith ionosphere delay is on the order of 0.1 mm/sec over 10 minutes [Wu, 1982]. Solar plasma induces a delay rate error on the order of 0.01 mm/sec in X-band signal phase-rate over 10 minutes, 10° away from the Sun [Kahn, 1988].

The S-band downlink cannot be coherently phase-tracked when the signal path is within 5° of the Sun due to large solar plasma induced spectral broadening and variations in the received signal phase. This will occur in the middle of Magellan's prime mission, from October 13, 1990 to November 22, 1990. To calibrate for charged particles during this period, ionospheric delay rate will be measured independently by measuring the Faraday rotation of linearly polarized VHF signals from geosynchronous satellites over the tracking stations. The TEC along the line-of-sight is deduced from these measurements, and mapped to the Magellan spacecraft or quasar line-of-sight as a function of time. [Royden, 1980]

A comparison between the dual-frequency calibrations and the Faraday rotation calibrations are shown in Figure 4. The calibrations, in Hertz, are shown for three spacecraft and three quasar scans, made up of 20 second points. These data were acquired on August 10, 1989. The Faraday calibrations agree with the mean of the dual-frequency calibrations for each 10 minute scan, but the dual-frequency method provides calibrations for charged

particle flucuations within a 10 minute scan. The Faraday method provides calibrations for only the low-frequency variation of the ionosphere over a 10 minute scan.

The Earth's troposphere induces a radio signal delay from 2 meters at zenith, to 12 meters at 10° elevation. A seasonal model of the delays, generated from 2 years of radiosonde measurements made at the DSN tracking stations, can calibrate this delay to 6 cm in zenith, which scales to other elevation angles in proportion to the path length through the troposphere [Chao, 1971]. This model does not predict the fluctuating wet component of the troposphere well. A theoretical prediction of this error at 10° elevation is 0.07 mm/sec [Truehaft and Lanyi, 1987].

To verify the integrity of the VLBI data, the sub-second phase residuals from each station are plotted before compressing. A line is removed from these residuals, leaving only media information, so that any cycle slips can be corrected before observable generation. These plots are also useful to assess the size of media corruptions. In Figure 5, the S and X-Band phase residuals are plotted, for 3 separate 10 minute spacecraft scans, over 75 minutes on July 27, 1989. The phase received at the California station and the Madrid station are each plotted, as well as the station differenced phase. The scale is defined by the S and X wavelengths denoted on the right side of the plots. In scans 1 and 2, at both California and Spain, the same phase signatures appear, and the X-Band signature is larger than the S-Band. This is due to charged particle corruption to the uplink, which is multiplied by the spacecraft to X-Band frequency for downlink. This effect is eliminated in the station differenced phase for these scans. In scan 3, however, the phase is perturbed by the fluctuating wet troposphere, which has a larger effect at the higher X-Band frequency, and is not eliminated in the station differencing. The wet troposphere can fluctuate so fast that this effect cannot be removed by the quasar calibration, leaving this effect as the dominant VLBI error source for velocity measurements.

Figure 6 also shows phase residuals for 2 spacecraft scans, acquired over 2 hours on July 16, 1989. In both scans, the ionosphere delay over Australia is apparent because the X-Band phase signature is magnified in the S-Band phase, and remains in the station differenced phase. This effect can be completely eliminated by applying the S and X-Band dual-frequency calibration to each the spacecraft and quasar data points.

Media Calibration Results

DOR

Two sets of ΔDOR residuals, each computed with the reference trajectory described above, are shown in Figure 7. These residuals were computed with and without the seasonal troposphere calibration. The calibrated troposphere delay is about 5 meters. The post-calibration residual RMS is 0.183 meters, or 20 nanoradians in θ, the geocentric angular position.

Figure 8 shows residuals computed with and without Faraday rotation ionosphere calibrations. This calibration improves some scans (August 6 and 8), but degrades others

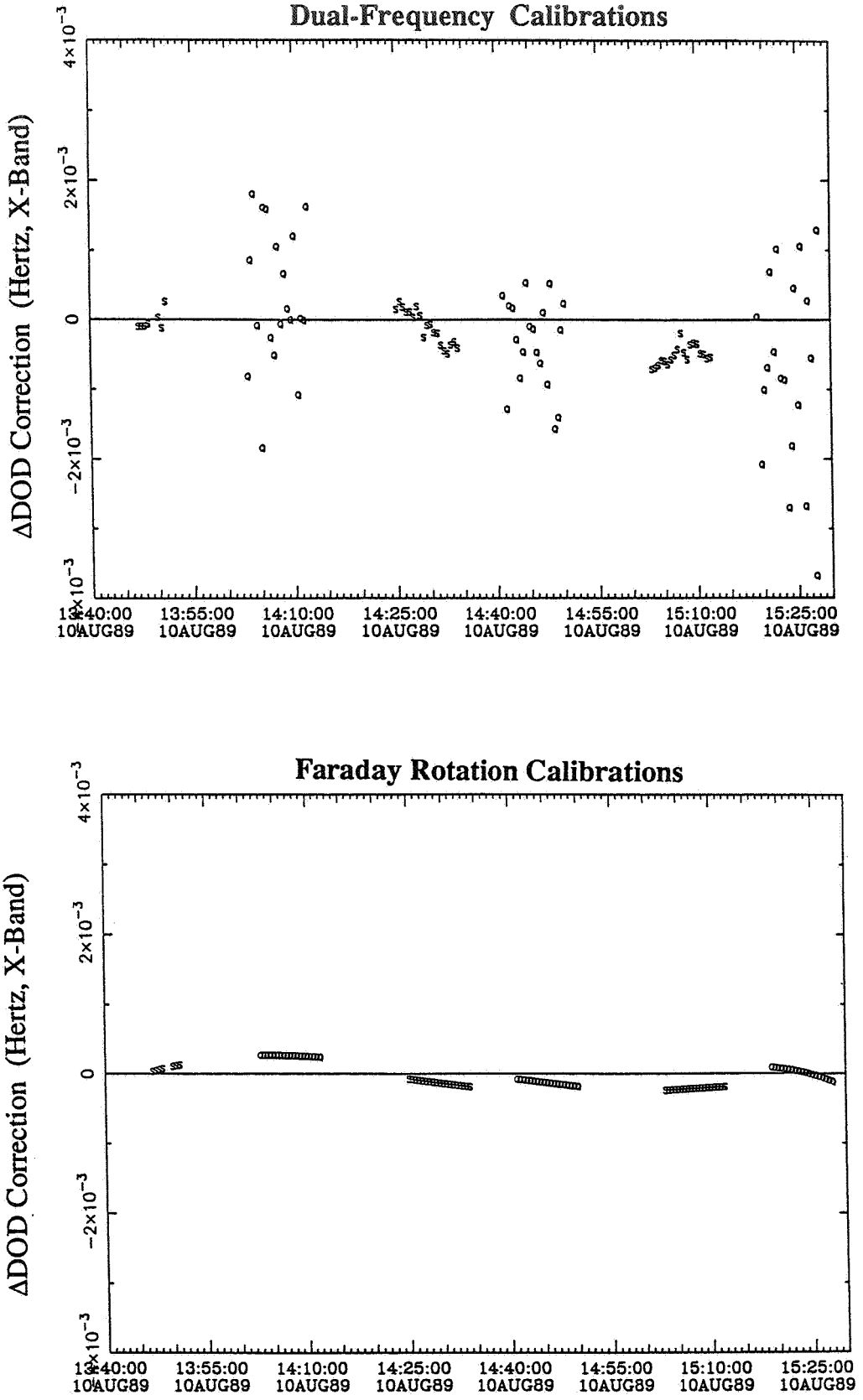


Figure 4. Charged Particle Calibrations for August 10, 1989

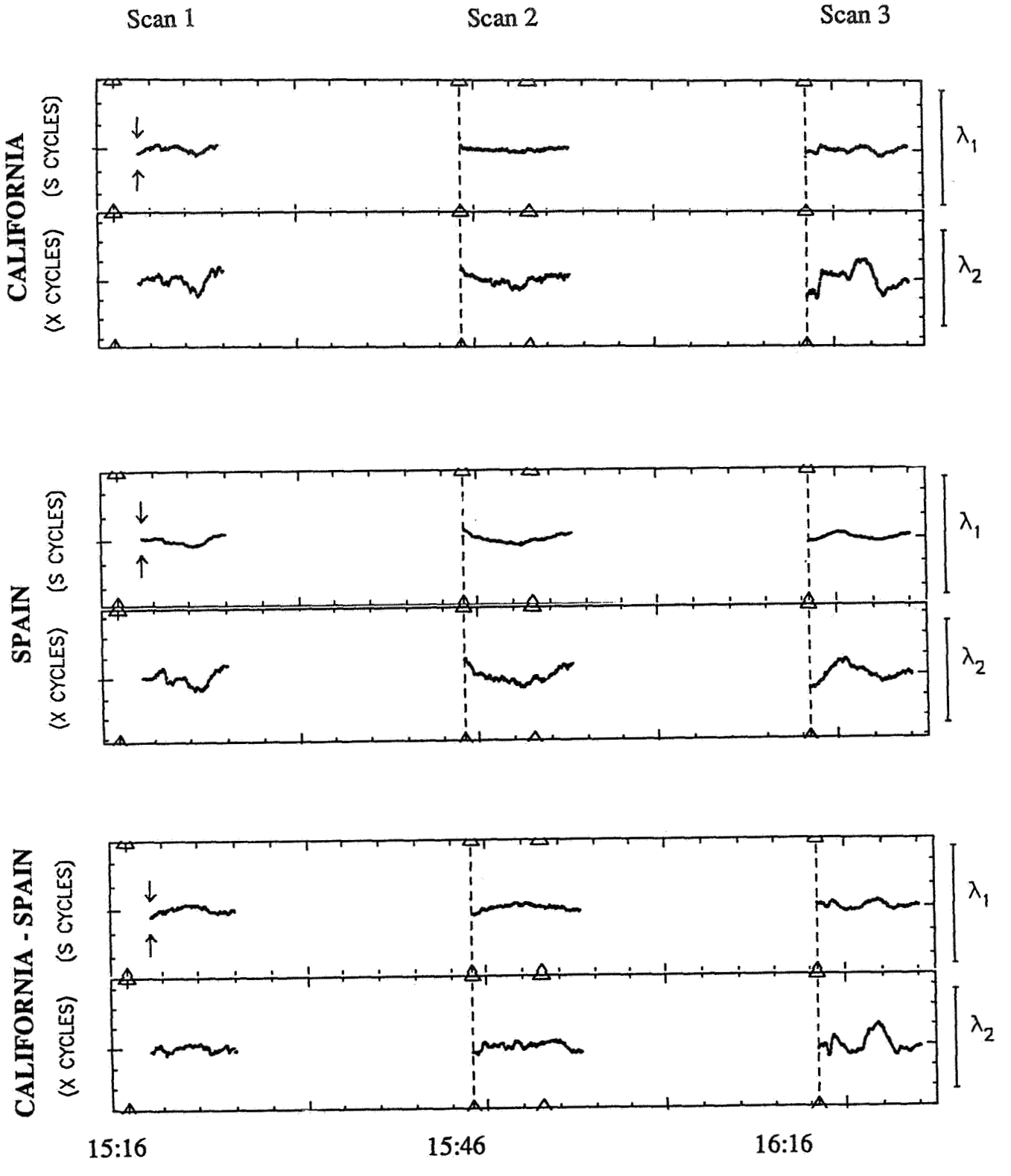


Figure 5. Uncompressed Phase Residuals for July 27, 1989

$\lambda_1 = 131$ mm, $\lambda_2 = 36$ mm

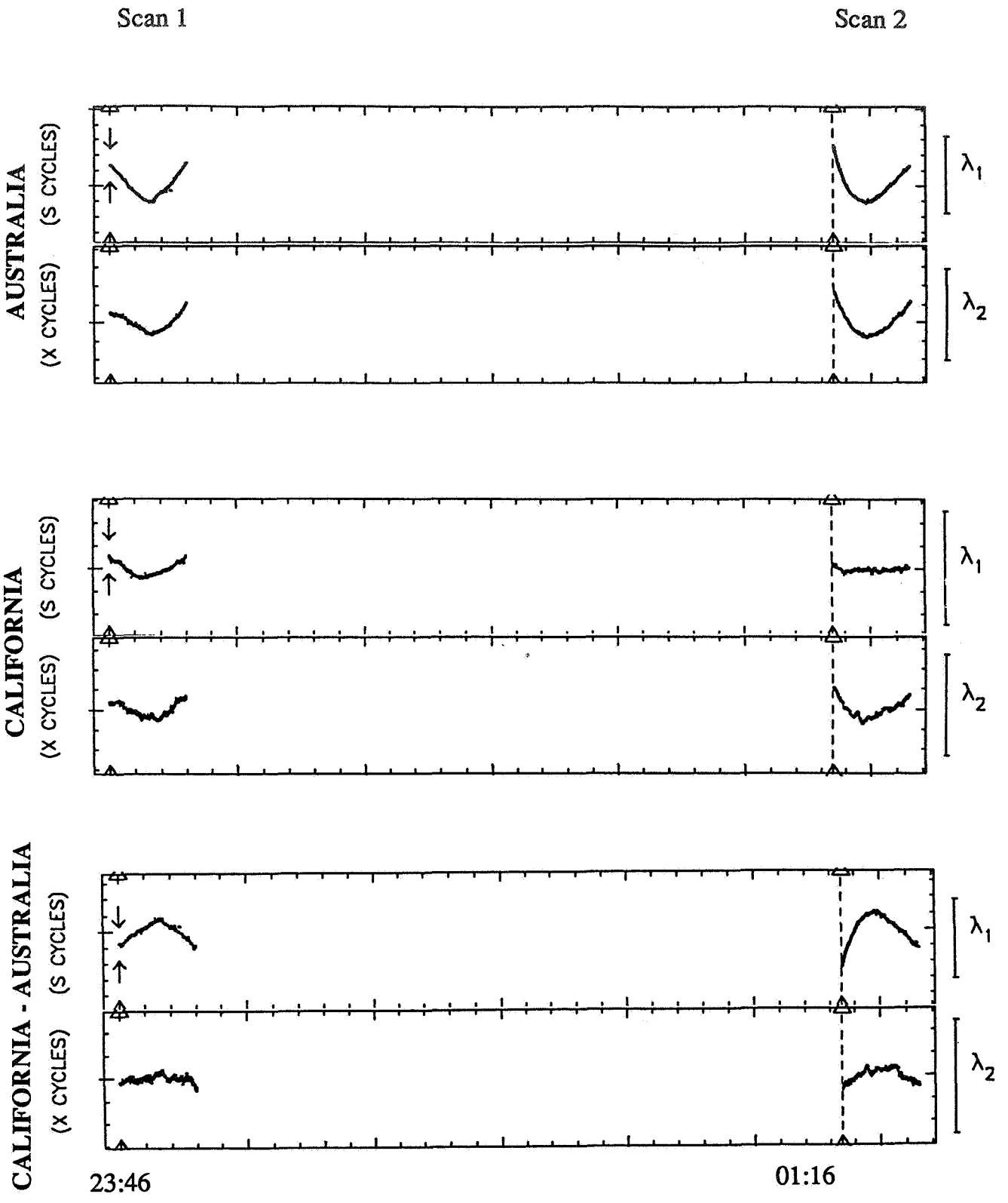


Figure 6. Uncompressed Phase Residuals for July 17, 1989

$\lambda_1 = 131$ mm, $\lambda_2 = 36$ mm

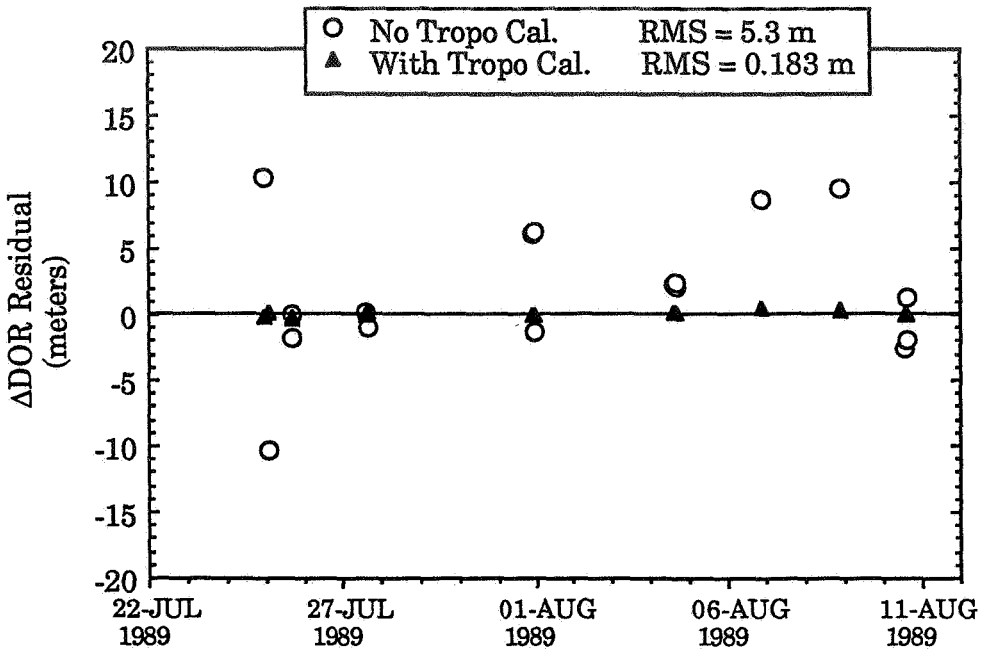


Figure 7. ΔDOR Residuals with and without Troposphere Calibration.

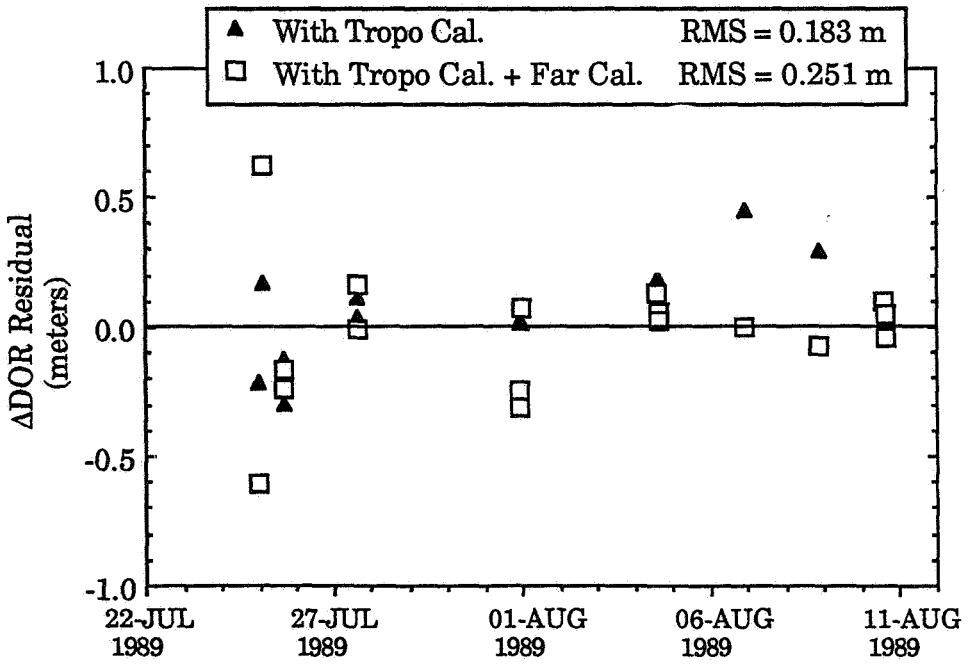


Figure 8. ΔDOR Residuals with and without Ionosphere Calibration.

(July 25), with a small net residual RMS degradation of 0.07 meters. This could be due to the error incurred when mapping the TEC from zenith to the low Magellan line-of-sight.

DOD

Two sets of ΔDOD residuals are shown in Figure 9. These residuals were computed with and without the seasonal troposphere calibration. The RMS troposphere delay-rate calibration is 0.66 mm/sec. The post-calibration residual RMS is 0.054 mm/sec, or 6 picoradians/sec in the geocentric angular velocity, $\dot{\theta}$.

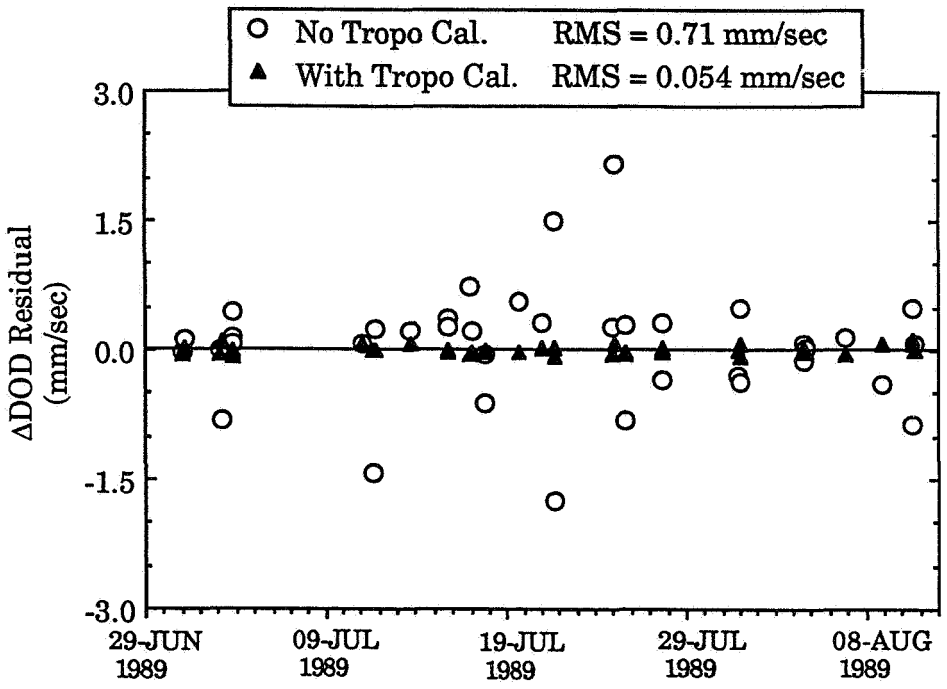


Figure 9. ΔDOD Residuals with and without Troposphere Calibration.

The calibration improvement using the quasar DOD measurement is shown in Figure 10, which plots the DOD residuals with and without the quasar calibration. Both data sets have been calibrated with the seasonal troposphere model. The quasar calibration improves the residual RMS by 0.05 mm/sec, or about 6 picoradians/sec in $\dot{\theta}$, the geocentric angular velocity. The quasar calibrates the ionosphere and troposphere components common between the two scans, station location errors, and station clock drift.

Figure 11 shows two sets of ΔDOD residuals, computed with and without charged particle calibrations. Dual-frequency (S and X-Band) calibrations were used, except for measurements where the S-Band measurement was not recorded; in these cases, the Faraday ionosphere calibration was used (20% of the data points). The charged particle calibration improves the data RMS by 0.01 mm/sec, or about 1 picoradian/sec. This is the expected delay-rate due to the solar plasma; most of the ionosphere delay-rate for this data set was calibrated with the quasar measurement. The resulting residual RMS is 0.045 mm/sec, or about 5 picoradians/sec.

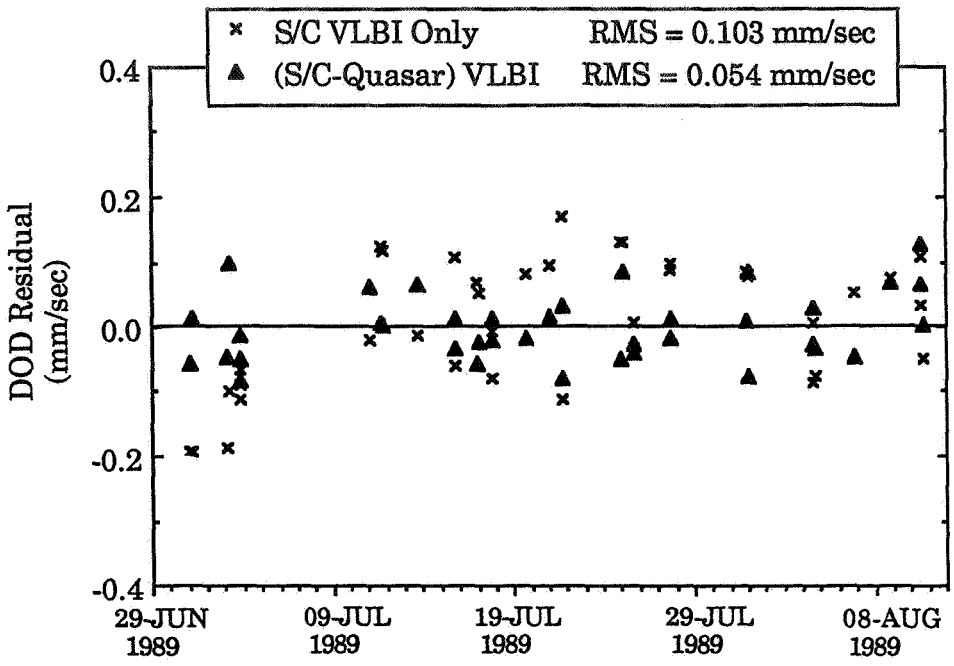


Figure 10. DOD Residuals with and without Calibration by Quasar.

Earth Platform Model

Errors in the tracking station location in inertial space induce errors in the computation of the ΔVLBI signal delay and delay-rate. An error in the angular position of the quasar also contributes to the error. The tracking station locations and quasar locations used for the ΔVLBI computations were derived from VLBI observations of quasars, and oriented to the planetary ephemeris frame. This orientation was constructed by matching the VLBI station locations to the Lunar Laser Ranging (LLR) station locations, which were used in the derivation of the planetary ephemeris reference frame [Folkner and Finger, 1989; Neill, 1984]. The uncertainty of this orientation is believed to be 20 nanoradians, though possible systematic errors are still being investigated. Figure 12 shows the ΔDOR residuals computed with these station locations and quasar locations. The ΔDOR residuals were also computed using tracking stations derived from planetary encounter tracking data [Moyer, 1988]. The quasar locations were oriented to this planetary frame by simply rotating the quasar right ascensions by the average offset between the longitudes of the Moyer station set and the standard VLBI station set [Ulvestad, 1989]. This average offset is 208 nanoradians. Figure 12 also shows the ΔDOR residuals computed with these station locations and quasar locations. The ΔDOR residual RMS was degraded by 5.4 cm, or 6 nanoradians in geocentric angular position, for this case.

The Earth's pole location and rotation rate (UT1) fluctuate due to deformations of the

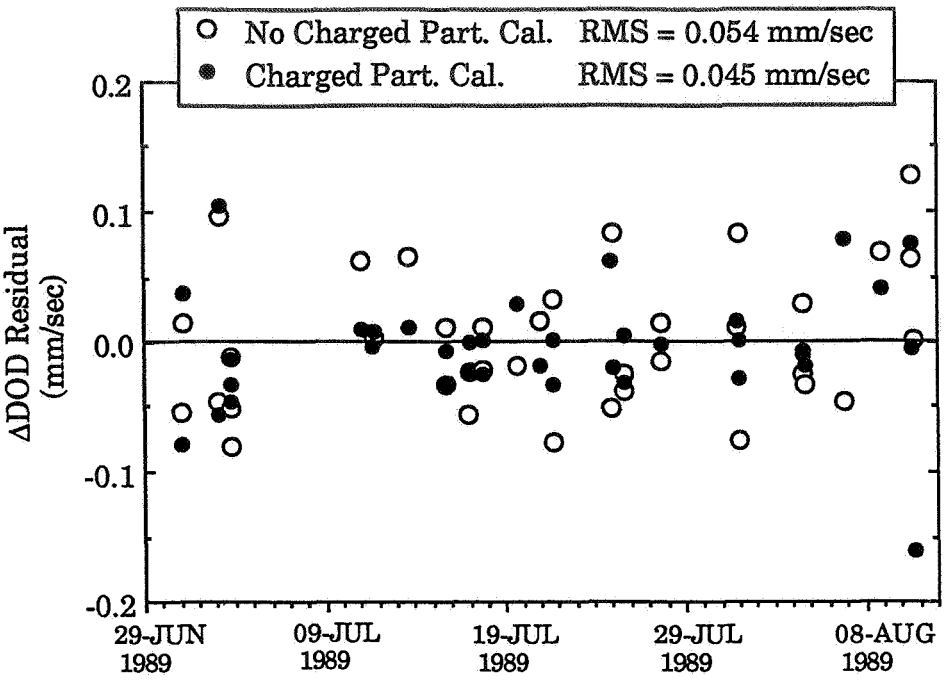


Figure 11. ΔDOD Residuals with and without Charged Particle Calibration.

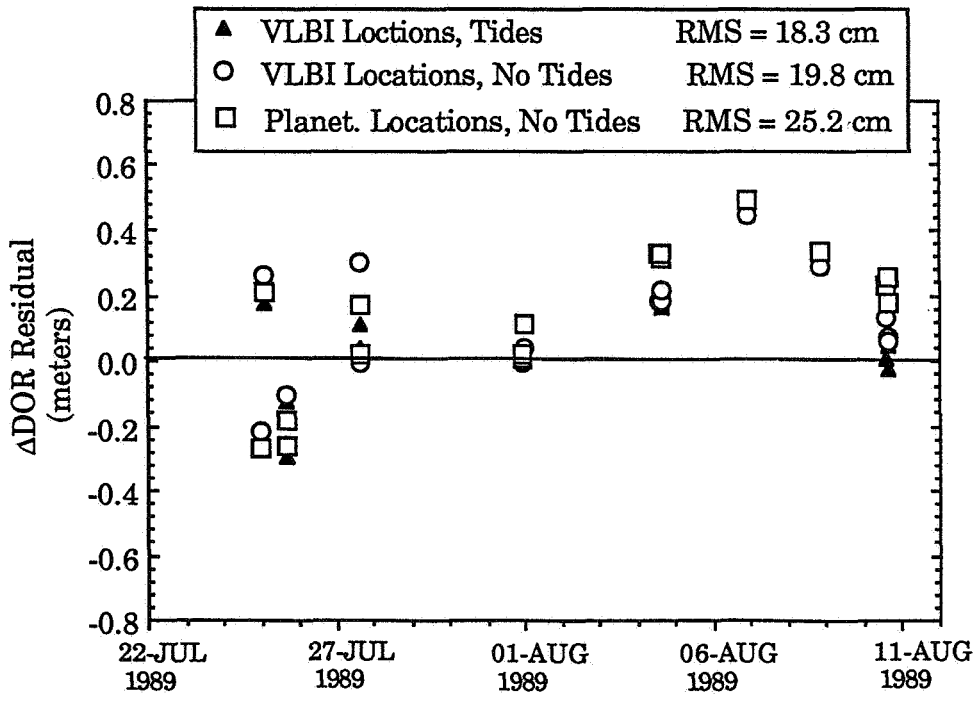


Figure 12. ΔDOR Residuals with alternate Earth Platform Models.

solid Earth and by exchanges of angular momentum between the solid and fluid parts of the Earth, as well as by exchanges of angular momentum with extraterrestrial objects. This motion is estimated using several sources of geodetic measurements. A Kalman filtering solution produces a smoothed time series with interpolation points spaced every 2 weeks. Tidal deformation of the Earth by the sun and moon produce periodic variations in the Earth's moment of inertia, and therefore in the Earth's rotation rate. These short-term tidal effects can be calculated from a theoretical model [Yoder et al., 1981], and included in this time series, with interpolation points every 2 days. Figure 12 shows ΔDOR residuals computed with and without these tidal effects. Inclusion of these tidal terms improved the residual RMS by 1.5 cm, or 1.7 nanoradians.

These improved levels of Earth platform modelling did not improve the ΔDOD measurement accuracy. As shown in Figure 12, these model improvements only provided slight improvement to the ΔDOR measurement accuracy.

Conclusions

Measurements of the Magellan spacecraft angular position and velocity are were made with ΔVLBI during the the spacecraft's heliocentric cruise from Earth to Venus. The angular measurement accuracy was 20 nanoradians in position and 5 picoradians/sec in velocity. The largest error contribution for velocity comes from the fluctuating wet troposphere; this media effect fluctuates too quickly to be calibrated with the quasar measurement. Charged particle effects are eliminated by dual-frequency calibrations. Faraday rotation measurements provide effective ionosphere calibrations when dual-frequency measurements are not available. Small improvements in the ΔDOR measurement accuracy, on the order of a few nanoradians, were obtained with an Earth platform model containing tracking station locations derived from quasar VLBI observations versus stations locations derived from planetary encounter tracking data. Modelling the high frequency effects in the Earth rotation model due to the solid Earth tides also improves the ΔDOR measurement accuracy by a few nanoradians. Each of these Earth platform models did not improve the ΔDOD measurement accuracy for this data set.

Acknowledgement

The work described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Astronautics Administration.

References

Border, J.S., F.F. Dovivan, S.G. Finley, C.E. Hildebrand, B. Moultrie, and L.J. Skjerve, "Determining Spacecraft Angular Position with Delta VLBI: The Voyager Demonstration", paper AIAA-82-1471, AIAA/AAS Astrodynamics Conference, San Diego, California, August, 1982.

Chao, C.C., "New Troposphere Range Corrections with Seasonal Adjustment", JPL Technical Report 32-1526, V. 6, pp.67-82, Jet Propulsion Laboratory, September/October 1971.

Esposito, P.B., F.F. Donivan, S.G. Finley, X.X. Newhall, C.B. Smith, S.C. Wu, "Narrowband Differential Interferometry Applied to Pioneer Venus Orbiter", AAS paper 83-310, AAS/AIAA Astrodynamics Specialist Conference, Lake Placid, New York, August, 1983.

Folkner, W.M., and M.H. Finger, "Station Location set and quasar catalog for use with ephemeris DE200", JPL internal document, September, 1989.

Kahn, R.D., and J.S. Border, "Precise Interferometric Tracking of Spacecraft at Low Sun-Earth-Probe Angles", paper AIAA-88-0572, AIAA 26th Aerospace Sciences Meeting, Reno, Nevada, January, 1988.

Melbourne, W.G., and D.W. Curkendall, "Radiometric Direction Finding: A New Approach to Deep Space Navigation", paper A78-31889, AAS/AIAA Astrodynamics Specialist Conference, Jackson Hole, Wyoming, September, 1977.

Moyer, T.D., "J2000 Station Location Sets for Planetary Ephemerides DE200 and DE202 and 1950 Station Coordinates for DE130", JPL internal document, August, 1988.

Mohan, S.N., "Challenges In Navigation of the Venus Radar Mapper (VRM) Spacecraft", paper IAF-85-259, IAF '85, Stockholm, Sweden, October, 1985.

Neill, A.E., "Absolute Geocentric DSN Station Locations and Radio-Planetary Frame Tie", JPL internal document, March, 1984.

Royden, H.N., D.W. Green, G.R. Walson, "Use of Faraday-Rotation Data from Beacon Satellites to Determine Ionospheric Corrections for Interplanetary Spacecraft Navigation", Proc. Satellite Beacon Symposium, Warszawa, Poland, May, 1980.

Thomas, J.B., "An Analysis of Long Baseline Radio Interferometry", JPL Technical Report 32-1526, Vol. VII, pp. 37-50, 1972.

Treuhaft, R.N. and G.E. Lanyi, "The effect of the dynamic wet troposphere on radio interferometric measurements", *Radio Science*, Volume 22, Number 2, pp. 251-265, March-April 1987.

Ulvestad, J.S. and O.J. Sovers, "Preliminary VLBI catalog for Magellan", JPL internal document, January, 1989.

Wood, L.J., "Orbit Determination Singularities in the Doppler Tracking of a Planetary Orbiter", *Journal of Guidance, Control, and Dynamics*, Volume 9, Number 4, page 485, July-August 1986.

Wolff, P.J., "Documentation of the equations in the CALGEN Program", JPL internal document, September, 1985.

Wu, S.C., "Error Estimation for ΔVLBI Angle and Angle Rate Measurements Over Baselines Between a Ground Station and a Geosynchronous Orbiter", TDA Progress Report 42-71, pp. 8-19, July-September 1982.

Yoder, C.F., J.G. Williams, M.E. Parke, "Tidal Variations of Earth Rotation", *Journal of Geophysical Research* , pp. 881-891, 1981.

JOINT JSC/GSFC TWO-TDRS NAVIGATION CERTIFICATION RESULTS FOR STS-29, STS-30, AND STS-32*

Thomas G. Schmidt, Edward T. Brown, and Valerie E. Murdock
ROCKWELL SPACE OPERATIONS COMPANY (RSOC)

James O. Cappellari, Jr., Evan A. Smith, Mark W. Schmitt, and James W. O'Malley
COMPUTER SCIENCES CORPORATION (CSC)

Flora B. Lowes
JOHNSON SPACE CENTER (JSC)

James B. Joyce
GODDARD SPACE FLIGHT CENTER (GSFC)

ABSTRACT

This paper describes the procedures used and the results obtained in the joint Johnson Space Center (JSC)/Goddard Space Flight Center (GSFC) navigation certification of the two-TDRS S-band tracking configuration for support of low- to medium-inclination (28.5 to 62 degrees) Shuttle missions (STS-29 and STS-30) and Shuttle rendezvous missions (STS-32). The objective of this certification effort was to certify the two-TDRS configuration for nominal STS on-orbit navigation support, thereby making it possible to significantly reduce the ground tracking support requirements for routine STS on-orbit navigation.

JSC had the primary responsibility for certification of the two-TDRS configuration for STS support, and GSFC supported the effort by performing Ground Network (GN) and Space Network (SN) tracking data evaluation, parallel orbit solutions, and solution comparisons.

In the certification process, two types of orbit determination solutions were generated by JSC and by GSFC for each tracking arc evaluated, one type using TDRS-East and TDRS-West tracking data combined with ground tracking data (the reference solutions) and one type using only TDRS-East and TDRS-West tracking data. The two types of solutions were then compared to determine the maximum position differences over the solution arcs and whether these differences satisfied the navigation certification criteria. The certification criteria were a function of the type of Shuttle activity in the tracking arc, i.e., quiet, moderate, or active. Quiet periods included no attitude maneuvers or ventings, moderate periods included one or two maneuvers or ventings, and active periods included more than two maneuvers or ventings.

This paper presents the results of the individual JSC and GSFC certification analyses for the STS-29, STS-30, and STS-32 missions and the joint JSC/GSFC conclusions regarding certification of the two-TDRS S-band configuration for STS support.

* This work was supported by the National Aeronautics and Space Administration (NASA)/Johnson Space Center (JSC), Houston, Texas, under Contract NAS 9-18000, and by NASA/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, under Contract NAS 5-31500.

1.0 INTRODUCTION

1.1 BACKGROUND

NASA is transitioning navigation support for Shuttle missions and for unmanned spacecraft from a primarily ground-based system (the Ground Network (GN)), utilizing Ground Spacecraft Tracking and Data Network (GSTDN) S-band tracking stations and Department of Defense (DOD) C-band tracking stations, to a primarily space-based system (the Space Network (SN)), utilizing the Tracking and Data Relay Satellite (TDRS) System (TDRSS). Currently, TDRSS consists of three geosynchronous satellites: TDRS-East, located at 41 degrees west longitude; TDRS-Spare, located at 171 degrees west longitude; and TDRS-West, located at 174 degrees west longitude.

The task of certifying the TDRSS for navigation support for the Space Transportation System (STS) was assigned to the TDRSS Orbit Determination and Navigation Working Group (TODNWG), a group composed of engineers from the Johnson Space Center (JSC) Navigation Section and from the Goddard Space Flight Center (GSFC) Flight Dynamics Facility (FDF). This group first met in 1982 to develop the STS/TDRSS Navigation Certification Plan (Reference 1). The first step in the certification of TDRSS for STS support was a joint JSC/GSFC single-TDRS certification effort, which took place between August 1983, the start of TDRSS tracking services, and the end of the STS-41G mission in October 1984. This effort was successful in certifying the augmented single-TDRS network (TDRS-East plus ground stations outside the TDRS-East coverage) as an adequate navigation system for noncritical Shuttle flight periods (see Reference 2).

The second step in the certification process was a joint JSC/GSFC two-TDRS certification effort with the objective of certifying the TDRS-East/TDRS-West configuration (without ground station augmentation) for nominal on-orbit navigation support of STS flights. A successful certification effort would mean that ground tracking support for routine STS on-orbit navigation could be significantly reduced or eliminated. Some ground-based tracking would still be necessary to support critical and high-activity periods.

This certification effort consisted of two phases: (1) certification of the two-TDRS configuration for navigation support of low- to medium-inclination (28.5 to 62 degrees) missions that do not include a rendezvous and (2) certification of the two-TDRS configuration for navigation support of low- to medium-inclination rendezvous missions. The STS certification missions were STS-29 and STS-30 (nonrendezvous missions) and STS-32 (a rendezvous mission). This paper documents the certification results obtained for these three missions at JSC and GSFC. A more detailed report of the certification results will be published at a later date (Reference 3).

JSC had the primary responsibility for certification of the two-TDRS network for STS support. GSFC supported this effort by performing tracking data evaluation for the GN and SN tracking data and by performing parallel orbit solutions and solution comparisons.

1.2 ORGANIZATION OF PAPER

Section 2 describes the JSC and GSFC certification criteria and procedures used in this study, and Section 3 presents the certification results. Conclusions are given in Section 4.

2.0 CERTIFICATION CRITERIA AND PROCEDURES

This section presents the two-TDRS navigation certification criteria and describes the procedures used in the orbit determination process and in the comparison of orbit solutions at both JSC and GSFC. It also defines the spacecraft (Shuttle and TDRS) characteristics and force modeling parameters used in the orbit solutions.

2.1 TWO-TDRS NAVIGATION CERTIFICATION CRITERIA

The criteria for evaluating the navigation certification results for the two-TDRS effort are documented in Reference 4. For each tracking data arc evaluated, orbit determination solutions were obtained using TDRS-East

and TDRS-West tracking data combined with ground tracking data (the reference solution) and using only the TDRS-East and TDRS-West tracking data. These solutions were then compared to determine the maximum position differences between the reference and two-TDRS solutions over the tracking data arc. These maximum differences could not exceed the certification criteria shown in Table 1. Quiet periods included no attitude maneuvers or ventings within the tracking arc; moderate periods included only one or two attitude maneuvers or ventings; and active periods included more than two attitude maneuvers or ventings and the 6-hour period just prior to deorbit ignition. TDRSS tracking of the Shuttle was rated satisfactory when 70 percent or more of the Doppler data were usable during scheduled support periods.

Table 1. Acceptance Criteria for STS Navigation Certification

EPHEMERIS COMPARISONS	MAXIMUM TOTAL POSITION DIFFERENCES (METERS)				
	STS ACTIVITY LEVEL				
	QUIET	MODERATE	ACTIVE	POSTINSERTION	RENDEZVOUS
TWO-TDRS-ONLY BTB VERSUS REFERENCE BET	320	420	570	420	700
TWO-TDRS-ONLY BET VERSUS REFERENCE BET	280	420	640	420	750

NOTES:

1. EPHEMERIS COMPARISONS SHALL SATISFY THE ABOVE CRITERIA IN 90 PERCENT OF THE CASES FOR EACH ACTIVITY LEVEL.
2. BTB = BATCH-TO-BATCH
   BET = BEST ESTIMATED TRAJECTORY

2.2 JSC CERTIFICATION PROCEDURES

The JSC two-TDRS certification procedures that were common to all three missions are described in Sections 2.2.1 and 2.2.2. The mission-specific procedures are described in Section 2.2.3. The reference and two-TDRS-only solutions generated by JSC used the spacecraft characteristics and force modeling parameters given in Table 2.

TDRS-East and TDRS-West ephemerides were updated every 6 hours using accurate GSFC-supplied vectors. The tracking data sampling rate was every 40 seconds for TDRSS S-band measurements, every 10 seconds for GN S-band measurements, and every 6 seconds for GN C-band measurements. Data weights (sigmas) used in the solutions were (1) 0.10 hertz for TDRS S-band Doppler measurements; (2) 18.3 meters, 0.0344 degree, and 0.60 hertz for GN S-band two-way range, angles, and Doppler measurements, respectively; and (3) 18.3 meters and 0.0229 degree for GN C-band two-way range and angle measurements, respectively.

2.2.1 Batch-to-Batch (BTB) Processing

JSC Ground Navigation normally processes data in a batch-to-batch (BTB) mode, where data batches are processed sequentially after the end of each tracking pass. A weighted least-squares differential correction of the spacecraft orbital estimate is performed according to the following equation:

$$\Delta x_i = [A^T W A + (k^n \Gamma)^{-1}]^{-1} [A^T W \Delta y + (k^n \Gamma)^{-1} \Delta x_{i-1}] \quad (1)$$

where

Δx = current and last corrections to the a priori position and velocity state vector estimate for each iteration
Δy = vector of residuals (observations minus the expected values)
A = matrix containing the partial derivatives of the data observations with respect to the Cartesian position and velocity components
W = diagonal matrix of observation data weights ($1/\sigma^2$)
Γ^{-1} = covariance from the last BTB solution, propagated to the time of the current batch
k^n = covariance multiplier

The multiplier k^n is applied to the a priori covariance to control the amount by which previous history constrains the current solution. The number of times the k value is applied, n, is controlled by the JSC navigator during processing. The multiplier can be applied to the entire covariance matrix (an (XYZ) $k\Gamma$). The in-plane or out-of-plane elements of the covariance can be selectively downweighted through a transformation of the covariance from Cartesian to UVW (radial, along-track, and cross-track) coordinates (a (UVW) $k\Gamma$). An in-plane (UV) $k\Gamma$ allows the current data to change the in-plane elements of the state vector while constraining out-of-plane changes. A (W) $k\Gamma$ affects only the out-of-plane covariance elements. A (UV) $k\Gamma$ followed by a (W) $k\Gamma$ is completely equivalent to an (XYZ) $k\Gamma$. TDRS BTB processing often employs in-plane covariance downweighting in an attempt to compensate for the weakness of Doppler-only TDRS data in orbital plane determination following trajectory perturbations.

2.2.2 Best-Estimated-Trajectory (BET) Processing

The mathematical basis of best-estimated-trajectory (BET) differential correction processing is essentially the same as for BTB processing, although in practice there are several differences between the two modes. Rather than processing a single pass of tracking data from one station, BET processing considers data arcs that contain measurement information from several tracking passes. The solved-for state vector can include up to three vents, whose start and stop times are specified by the JSC navigator. These solved-for vents are often used to account for unmodeled thrusting due to attitude and translational maneuvers. BET processing does not normally use an a priori covariance, because this would unrealistically constrain the solved-for position, velocity, and vent force solutions. Angle measurements are usually excluded from BET processing.

The two-TDRS-only BET processing was performed over the same data arcs used in the reference BET. Quiet periods were certified during STS-29 and therefore were not included in the STS-30 and STS-32 processing. Each data arc began and ended with TDRSS data so that the reference and two-TDRS-only BETs would be directly comparable. Adjacent arcs included one or two common batches to minimize discontinuities between successive trajectories. The start and stop times for each BET arc were defined as the points of minimum trajectory difference within the overlap portions of the surrounding arcs. Solution residuals were minimized by solving for the Shuttle state vector and for vents that were placed to match actual trajectory perturbations as closely as possible. Vents were not solved for in the overlap regions. Solution quality was judged on the basis of residuals beyond the data arc, as computed from the propagated solution vector, solution statistics, and the reasonableness of solved-for parameters.

2.2.3 Mission-Specific Procedures

The JSC two-TDRS certification procedures that were mission specific are discussed below for STS-29, STS-30, and STS-32.

2.2.3.1 STS-29 PROCEDURES

JSC Ground Navigation maintained two realtime BTB sequences (chains) throughout the entire mission. The reference BTB chain was initialized on the revolution 2 C-band ground pass from Kwajalein. The two-TDRS-only BTB chain was initialized with a BET solution over the first two post-OMS-2 TDRSS passes (OMS is the Orbital Maneuvering System). A Shuttle body-axis correction vent was modeled during the on-orbit timeframe to account for unmodeled translation effects of attitude control thrusting. Both the reference and two-TDRS BET solutions included this vent, meaning that any solved-for vent force was in addition to the already

modeled forces. A reference BET and a two-TDRS-only BET were generated for 23 data arcs spanning the entire mission.

2.2.3.2 STS-30 PROCEDURES

JSC Ground Navigation maintained two realtime BTB chains throughout the entire mission. The reference BTB chain was initialized on the revolution 2 C-band ground pass from Kaena Point, Hawaii. The two-TDRS BTB chain was initialized on a BET solution over the first two on-orbit TDRS passes. Constant Shuttle body-axis correction vents were modeled. Both the reference and two-TDRS BET solutions included these modeled vents, meaning that any solved-for vent force was in addition to the already modeled forces. The STS-30 reference BET consisted of 19 data arcs spanning the period between OMS-2 and the deorbit burn.

2.2.3.3 STS-32 PROCEDURES

JSC Ground Navigation maintained the reference BTB chain throughout the entire mission. The two-TDRS BTB chain, which spanned only the rendezvous period, was initialized on a TDRS-East solution from the well-established reference BTB chain. Five reference and three two-TDRS-only BET arcs were processed during the rendezvous certification period. Constant Shuttle body-axis correction vents were modeled. Both the reference and two-TDRS BET solutions included these modeled vents, meaning that any solved-for vent force was in addition to the already modeled forces.

2.3 GSFC/FDF CERTIFICATION PROCEDURES

The GSFC/FDF certification processing and procedures are described below. The GSFC solutions were generated using the spacecraft parameters and force modeling parameters given in Table 3. The TDRS-East and TDRS-West ephemerides used in conjunction with SN tracking data were those generated as part of the normal FDF daily operations.

The tracking data types used in the solutions were (1) TDRSS S-band Doppler measurements, (2) GN S-band range and range-rate measurements, and (3) GN C-band range measurements. The tracking data sampling rate was every fourth observation for the TDRS S-band measurements and every observation for the GN S-band and C-band measurements. The data weights (sigmas) used in the solutions were (1) 0.25 hertz for the TDRSS S-band Doppler measurements; (2) 20 meters and 10 centimeters per second for the GN S-band range and range-rate measurements, respectively; and (3) 20 meters for the GN C-band range measurements.

2.3.1 GSFC FDF Batch Processing

The GSFC FDF uses a differential correction process to estimate the spacecraft orbit and associated parameters. This process uses a Bayesian weighted least-squares estimation algorithm with an a priori covariance matrix. The Cowell equations of motion are integrated with a predictor/corrector algorithm. For low-eccentricity orbits (such as TDRS and Shuttle), GSFC/FDF uses a fixed integration step size, in contrast to JSC, which uses a variable step size. The orbit processing at GSFC/FDF is essentially equivalent to the JSC BET processing.

2.3.2 GSFC FDF Procedures

For each certification tracking data arc, the GSFC/FDF generated a reference solution including both GN and SN tracking data and a two-TDRS solution including only SN tracking data. The orbit solutions for each tracking data arc were initially generated with no attempt to model the thrusting activities within the arc, as the FDF does not have an STS thrust modeling capability comparable to the JSC modeling. For those cases where the maximum position differences exceeded the certification criteria, the solutions were regenerated applying along-track thrust components provided by JSC. Finally, the reference and two-TDRS solutions were compared to determine the maximum position difference between the two solutions for each of the tracking data arcs. Whenever successive tracking data arcs overlapped, overlap comparisons were performed for both the reference solutions and the two-TDRS solutions.

Table 2. Spacecraft Characteristics and Force Modeling Parameters (JSC)

PARAMETERS	STS-29, STS-30, STS-32	TDRS-EAST AND TDRS-WEST
CROSS-SECTIONAL AREA	ATTITUDE-DEPENDENT	N/A
TYPE OF INTEGRATION	ENCKE	ENCKE
INTEGRATION STEPSIZE	52 SECONDS	333 SECONDS
INTEGRATION COORDINATE SYSTEM	MEAN OF 1950.0	MEAN OF 1950.0
GEOPOTENTIAL MODEL	GODDARD EARTH MODEL-10 (GEM-10) 7x7	GEM-10 7X7
ATMOSPHERIC DENSITY MODEL	JACCHIA-LINEBERRY MODEL WITH 90-DAY MEAN SOLAR FLUX (1970-71)	N/A
SOLAR REFLECTIVITY COEFFICIENT (C_R)	NOT USED	NOT USED
DRAG COEFFICIENT (C_D)	2.0	N/A
SOLAR/LUNAR/PLANETARY FILE	JPL DE-19	JPL DE-19
SOLVED-FOR PARAMETERS	STATE (POSITION AND VELOCITY), VENTS, AND MANEUVERS	N/A

Table 3. Spacecraft Characteristics and Force Modeling Parameters (GSFC/FDF)

PARAMETERS	STS-29, STS-30, STS-32	TDRS-EAST AND TDRS-WEST
CROSS-SECTIONAL AREA	120 METERS^2	40 METERS^2
TYPE OF INTEGRATION	FIXED-STEP COWELL	FIXED-STEP COWELL
INTEGRATION STEPSIZE	45 SECONDS	600 SECONDS
INTEGRATION COORDINATE SYSTEM	MEAN OF 1950.0	MEAN OF 1950.0
GEOPOTENTIAL MODEL	GODDARD EARTH MODEL-9 (GEM-9) 7x7	GEM-9 8x8
ATMOSPHERIC DENSITY MODEL	HARRIS-PRIESTER	N/A
SOLAR REFLECTIVITY COEFFICIENT (C_R)	1.5	SOLVED FOR
DRAG COEFFICIENT (C_D)	2.0	N/A
SOLAR/LUNAR/PLANETARY FILE	JPL DE-118	JPL DE-118
SOLVED-FOR PARAMETERS*	STATE (POSITION AND VELOCITY) AND DRAG VARIATION PARAMETER	STATE, C_R

*THRUST MODELING WAS APPLIED FOR SELECTED SOLUTION ARCS.

3.0 CERTIFICATION RESULTS

The JSC and GSFC/FDF certification results for the STS-29, STS-30, and STS-32 missions are described below.

3.1 STS-29

STS-29 was launched into a circular 28.5-degree inclination, 296-kilometer altitude orbit on March 13, 1989. The primary objective of the STS-29 mission was to deploy the third operational Tracking and Data Relay Satellite (TDRS-4). Twenty-three tracking data arcs were used for navigation certification during this mission. The thrusting activity level, the start and stop times, and the number of GN and SN tracking passes for each tracking data arc are given in Table 4.

Table 4. Tracking Arc Definition, Thrusting Activity Level, and Number of Tracking Passes for STS-29

ARC NO.	VENTING ACTIVITY LEVEL	TRACKING INTERVAL				NO. OF TRACKING PASSES			
		START		END		GROUND NETWORK		TDRSS	
		DATE	GMT (HHMMSS)	DATE	GMT (HHMMSS)	S-BAND/ C-BAND	TOTAL	TDRS-E/ TDRS-W	TOTAL
1	ACTIVE	3/13/89	15:53:30	3/13/89	21:10:40	0 / 13	13	3 / 4	7
2	ACTIVE		21:25:20	3/14/89	03:51:10	0 / 4	4	4 / 4	8
3	QUIET	3/14/89	02:25:20		08:48:40	0 / 3	3	3 / 6	9
4	MODERATE		08:01:30		13:21:10	0 / 3	3	3 / 5	8
5	ACTIVE		12:11:30		18:56:40	1 / 11	12	5 / 3	8
6	MODERATE		18:34:00	3/15/89	00:47:30	0 / 3	3	4 / 6	10
7	QUIET	3/15/89	00:17:30		06:11:30	0 / 6	6	4 / 5	9
8	ACTIVE		05:51:30		12:17:10	0 / 3	3	4 / 4	8
9	MODERATE		11:33:10		16:44:20	0 / 4	4	3 / 4	7
10	MODERATE		17:09:00		23:35:10	0 / 1	1	4 / 5	9
11	QUIET		22:03:00	3/16/89	04:23:40	1 / 5	6	4 / 5	9
12	QUIET	3/16/89	02:51:30		09:11:20	1 / 3	4	5 / 5	10
13	ACTIVE		08:27:00		14:03:20	0 / 4	4	3 / 4	7
14	ACTIVE		12:34:20		17:11:20	0 / 3	3	3 / 3	6
15	ACTIVE		15:46:30		21:09:30	0 / 2	2	4 / 3	7
16	ACTIVE		20:40:10	3/17/89	03:38:20	0 / 4	4	6 / 4	10
17	QUIET	3/17/89	02:12:40		08:27:40	0 / 4	4	4 / 4	8
18	ACTIVE		07:59:00		13:15:20	0 / 4	4	4 / 3	7
19	ACTIVE		12:30:00		18:09:20	2 / 11	13	4 / 4	8
20	ACTIVE		17:40:00		23:46:20	0 / 1	1	4 / 5	9
21	QUIET		20:12:20	3/18/89	06:06:50	1 / 5	6	4 / 7	11
22	MODERATE	3/18/89	04:56:30		09:45:10	0 / 3	3	3 / 4	7
23	ACTIVE		08:55:50		13:35:10	0 / 12	12	3 / 3	6

The GSFC tracking data evaluation for STS-29 is documented in Reference 5. A total of 8 S-band and 108 C-band GN on-orbit passes were evaluated. Anomalies were encountered in four of the S-band passes and three of the C-band passes. Approximately 80 percent of the TDRS-East view periods and 84 percent of the TDRS-West view periods had at least 70 percent usable data. With Doppler compensation and GN times deleted, the TDRS-East and TDRS-West data were approximately 95 percent usable.

The JSC and GSFC orbit determination results for STS-29 are presented in Sections 3.1.1 and 3.1.2, respectively.

3.1.1 JSC Orbit Determination Results for STS-29

This section describes the JSC STS-29 two-TDRS BTB and BET certification results.

The maximum position differences between the two-TDRS-only BTB and reference BET solutions (and their radial, cross-track, and along-track components) are shown in Table 5. Three of the two-TDRS-only BTB

Table 5. Maximum Position Differences Between Two-TDRS BTB and Reference BET Solutions for STS-29 (JSC)

ARC NO.	VENTING ACTIVITY LEVEL	MAXIMUM POSITION DIFFERENCE BETWEEN THE TWO-TDRS AND REFERENCE SOLUTIONS (METERS)					NUMBER OF TWO-TDRS SOLUTIONS PASSED/ TOTAL NUMBER	PERCENT PASSED
		RADIAL	CROSS-TRACK	ALONG-TRACK	TOTAL (RSS)	CERTIFICATION CRITERION		
1	ACTIVE	19	-372	-58	377	* 420	7 / 7	100
2	ACTIVE	-72	-324	416	532	570	8 / 8	100
3	QUIET	28	71	-165	182	320	7 / 7	100
4	MODERATE	132	120	-295	345	420	6 / 6	100
5	ACTIVE	-37	7	-276	278	570	7 / 7	100
6	MODERATE	13	148	269	307	420	9 / 9	100
7	QUIET	5	150	-64	164	320	8 / 8	100
8	ACTIVE	-130	188	-169	284	570	6 / 6	100
9	MODERATE	241	-38	103	265	420	7 / 7	100
10	MODERATE	33	377	42	381	420	8 / 8	100
11	QUIET	-1	-160	20	162	320	6 / 6	100
12	QUIET	25	30	-166	170	320	7 / 7	100
13	ACTIVE	-67	-163	304	352	570	5 / 5	100
14	ACTIVE	-20	-139	-80	162	570	3 / 3	100
15	ACTIVE	-115	1102	215	1129	570	3 / 5	60
16	ACTIVE	-17	-141	-11	142	570	9 / 9	100
17	QUIET	-21	-150	83	173	320	6 / 6	100
18	ACTIVE	14	-454	-256	522	570	7 / 7	100
19	ACTIVE	-113	1	-42	121	570	4 / 4	100
20	ACTIVE	369	-522	-375	742	570	5 / 6	83
21	QUIET	-71	-28	-155	173	320	8 / 8	100
22	MODERATE	-67	52	-194	212	420	6 / 6	100
23	ACTIVE	62	-171	40	187	570	6 / 6	100

* ARC 1 WAS A POSTINSERTION ARC

Table 6. Maximum Position Differences Between Two-TDRS BET and Reference BET Solutions for STS-29 (JSC)

ARC NO.	VENTING ACTIVITY LEVEL	MAXIMUM POSITION DIFFERENCE BETWEEN THE TWO-TDRS AND REFERENCE SOLUTIONS (METERS)					PASS/ FAIL
		RADIAL	CROSS-TRACK	ALONG-TRACK	TOTAL (RSS)	CERTIFICATION CRITERION	
1	ACTIVE	-54	217	-102	246	* 420	PASS
2	ACTIVE	-11	-75	-116	139	640	PASS
3	QUIET	-12	-59	27	66	280	PASS
4	MODERATE	6	94	-95	134	420	PASS
5	ACTIVE	25	-90	81	123	640	PASS
6	MODERATE	-23	62	108	127	420	PASS
7	QUIET	-5	-37	-11	39	280	PASS
8	ACTIVE	-7	-44	36	57	640	PASS
9	MODERATE	4	17	-20	26	420	PASS
10	MODERATE	3	28	36	46	420	PASS
11	QUIET	1	68	53	86	280	PASS
12	QUIET	-17	-92	9	94	280	PASS
13	ACTIVE	26	119	-49	132	640	PASS
14	ACTIVE	23	-79	-22	85	640	PASS
15	ACTIVE	-15	162	168	234	640	PASS
16	ACTIVE	-2	67	-6	67	640	PASS
17	QUIET	-7	-14	23	28	280	PASS
18	ACTIVE	4	-30	-9	32	640	PASS
19	ACTIVE	23	-148	-104	182	640	PASS
20	ACTIVE	-36	103	53	121	640	PASS
21	QUIET	8	42	-13	45	280	PASS
22	MODERATE	-5	68	-55	87	420	PASS
23	ACTIVE	33	-371	12	373	640	PASS

* ARC 1 WAS A POSTINSERTION ARC

solutions exceeded the certification criteria out of a total number of 151 solutions, corresponding to a 98-percent pass ratio. Arc 15 produced two failures and arc 20 produced one. The maximum total position difference of 1129 meters occurred during arc 15, primarily due to cross-track position differences.

Another measure of two-TDRS-only BTB solution accuracy is provided by comparisons with the reference BTB chain. Two-TDRS-only BTB differences in inclination and ascending node estimates did not subside until the fifth TDRSS batch after OMS-2. Total position differences were less than 610 meters for the entire postinsertion and deploy timeframe. During the deorbit preparation period, total position differences were below 244 meters.

The STS-29 reference and two-TDRS-only BET consisted of the 23 tracking data arcs described in Table 4. To more accurately model Shuttle trajectory perturbations, JSC Ground Navigation solved for 34 vents in both the reference and two-TDRS-only BET solutions. The first 32 vents had identical start and stop times. Comparable vent forces and associated energy changes were obtained between the reference and two-TDRS-only solutions. The majority of the semimajor axis changes were positive in sign, ranging from 12 meters to 1293 meters. This is normal for Shuttle flights due to translational effects from attitude control and attitude maneuvers.

Minimum trajectory overlap position differences between successive BET solutions were usually less than 360 meters. The two-TDRS-only BET produced comparable position differences relative to the reference BET. Comparisons were performed between the two-TDRS-only and reference BET solutions. Maximum position differences during the 23 data arcs are shown in Table 6. Every two-TDRS-only BET passed the certification criteria. Arc 23 produced the largest total position difference of 373 meters.

3.1.2 GSFC/FDF Orbit Determination Results for STS-29

The GSFC/FDF orbit determination results for STS-29 are presented in Table 7, which gives the maximum position differences between the two-TDRS and reference solutions (with along-track thrust modeling included in the solutions for arcs 1, 4, 5, 8, 13 and 19). The last column of Table 7 indicates whether the two-TDRS solution passed or failed the certification criterion for each arc. All the arcs passed the certification criteria except arc 8. The JSC vent solution for this arc showed that a large radial thrust component was applied, which explains why the application of an along-track thrust component in the FDF solution did not succeed in reducing the maximum position difference for this arc. Maximum overlap position differences between successive two-TDRS solutions ranged from 153 to 2453 meters. In summary, the GSFC/FDF analysis showed that 22 of the 23 arcs satisfied the two-TDRS navigation certification criteria.

3.2 STS-30

STS-30 was launched into a circular 28.85-degree inclination, 230-kilometer altitude orbit on May 4, 1989. The primary objective of the STS-30 mission was to deploy the Magellan interplanetary spacecraft. Nineteen tracking data arcs were used for navigation certification during this mission. The thrusting activity level, the start and stop times, and the number of GN and SN tracking passes for each tracking data arc are given in Table 8.

The GSFC tracking data evaluation results for STS-30 are documented in Reference 6. A total of nine S-band on-orbit passes were evaluated. Anomalies were encountered in four of the S-band passes and four of the C-band passes. Approximately 80 percent of the TDRS-East view periods and 70 percent of the TDRS-West view periods had at least 70 percent usable data. With Doppler compensation and GN times deleted, the TDRS-East data were approximately 90 percent usable and the TDRS-West data were approximately 94 percent usable.

The JSC and GSFC orbit determination results for STS-30 are presented in Sections 3.2.1 and 3.2.2, respectively.

3.2.1 JSC Orbit Determination Results for STS-30

STS-30 was the second of two low-inclination certification flights during which JSC and GSFC assessed the accuracy of two-TDRS orbit determination. This section describes the JSC two-TDRS BTB and BET certification results for each of the 19 data arcs.

Table 7. Maximum Position Differences Between Two-TDRS and Reference Solutions for STS-29 (With Along-Track Thrust Modeling) (GSFC/FDF)

ARC NO.	VENTING ACTIVITY LEVEL	MAXIMUM POSITION DIFFERENCE BETWEEN THE TWO-TDRS AND REFERENCE SOLUTIONS (METERS)					PASS/ FAIL
		RADIAL	CROSS-TRACK	ALONG-TRACK	TOTAL (RSS)	CERTIFICATION CRITERION	
1	ACTIVE	21	186	192	268	* 420	PASS
2	ACTIVE	52	47	430	436	640	PASS
3	QUIET	-11	45	-40	61	280	PASS
4	MODERATE	3	-17	-379	380	420	PASS
5	ACTIVE	31	207	318	381	640	PASS
6	MODERATE	0	104	380	394	420	PASS
7	QUIET	3	-60	-8	60	280	PASS
8	ACTIVE	-42	-499	488	699	640	FAIL
9	MODERATE	14	-20	-147	150	420	PASS
10	MODERATE	-1	0	15	15	420	PASS
11	QUIET	16	0	112	113	280	PASS
12	QUIET	15	-211	157	263	280	PASS
13	ACTIVE	-77	-36	342	352	640	PASS
14	ACTIVE	-38	-173	26	179	640	PASS
15	ACTIVE	22	77	117	142	640	PASS
16	ACTIVE	4	115	244	269	640	PASS
17	QUIET	9	-62	-72	96	280	PASS
18	ACTIVE	9	165	188	250	640	PASS
19	ACTIVE	-18	344	-408	534	640	PASS
20	ACTIVE	-7	-26	46	53	640	PASS
21	QUIET	5	-31	27	41	280	PASS
22	MODERATE	-2	-90	-87	125	420	PASS
23	ACTIVE	85	-11	-519	525	640	PASS

* ARC 1 WAS A POSTINSERTION ARC

NOTE: ALONG-TRACK THRUST MODELING WAS APPLIED FOR ARCS 1, 4, 5, 8, 13, AND 19.

Table 8. Tracking Arc Definition, Thrusting Activity Level, and Number of Tracking Passes for STS-30

ARC NO.	VENTING ACTIVITY LEVEL	TRACKING INTERVAL				NO. OF TRACKING PASSES			
		START		END		GROUND NETWORK		TDRSS	
		DATE	GMT (HHMMSS)	DATE	GMT (HHMMSS)	S-BAND/ C-BAND	TOTAL	TDRS-E/ TDRS-W	TOTAL
1	ACTIVE	5/4/89	19:45:20	5/5/89	01:01:50	0 / 21	21	3 / 4	7
2	ACTIVE		22:52:00		05:09:00	0 / 7	7	5 / 4	9
3	QUIET	5/5/89	04:49:40		10:59:50	0 / 4	4	4 / 4	8
4	ACTIVE		10:09:10		16:00:00	0 / 4	4	3 / 4	7
5	MODERATE		15:56:10		20:19:40	0 / 4	4	4 / 3	7
6	ACTIVE		19:08:00	5/6/89	01:32:50	0 / 7	7	4 / 6	10
7	ACTIVE	5/6/89	00:00:50		05:21:40	1 / 1	2	4 / 3	7
8	QUIET		04:49:40		11:13:10	0 / 4	4	4 / 4	8
9	QUIET		09:42:10		14:15:50	0 / 1	1	4 / 3	7
10	ACTIVE		12:56:50		18:16:30	0 / 3	3	4 / 3	7
11	ACTIVE		17:44:30		23:49:00	0 / 6	6	5 / 5	10
12	ACTIVE		22:24:50	5/7/89	04:47:50	1 / 4	5	5 / 5	10
13	QUIET	5/7/89	04:12:10		11:26:20	0 / 5	5	4 / 5	9
14	ACTIVE		10:35:00		16:53:00	0 / 4	4	5 / 5	10
15	ACTIVE		15:57:20		22:26:00	0 / 7	7	6 / 5	11
16	ACTIVE		20:56:50	5/8/89	03:35:20	0 / 3	3	5 / 5	10
17	QUIET	5/8/89	03:06:00		10:01:30	1 / 5	5	5 / 3	8
18	ACTIVE		09:10:50		14:55:00	0 / 3	3	3 / 5	8
19	ACTIVE		12:28:50		18:40:10	0 / 12	12	7 / 6	13

The maximum position differences between the two-TDRS-only BTB and reference BET solutions are shown in Table 9. These results satisfied the certification criteria 92 percent of the time. In many cases, however, the reference and two-TDRS-only ascending node estimates differed by about 0.005 degree. This caused large out-of-plane position differences, with the result that several arcs failed the TDRS BTB acceptance criteria. In addition, the two-TDRS BTB processing required several revolutions to recover from ascending node errors induced by the separation maneuver on revolution 5 and by several other orbit perturbations during the flight. The ascending node errors are not a general characteristic of two-TDRS postmaneuver processing, however. For example, the post-OMS-2 plane was fixed within three TDRS passes. Analysis of the failed two-TDRS-only realtime BTB processing revealed that the errors were probably due to over use of the (UV) $k\Gamma$, which so constrained the BTB solutions that actual plane changes were not allowed.

Two-TDRS-only BTB processing met the postinsertion acceptance criterion with the exception of the first two solutions, which had significant (–0.012-degree) ascending node errors. The two-TDRS-only BTB chain was initialized on a one-orbit BET solution over the first two on-orbit TDRSS passes, which should not have changed the plane appreciably. The first TDRS-West pass outside of the initialization data arc corrected most of the plane error, and the two-TDRS-only BTB processing met the acceptance criteria throughout the rest of postinsertion. The large initial ascending node error highlights the weakness of TDRSS Doppler-only data in determining the orientation of the orbit plane.

Two-TDRS-only BTB processing during the predeorbit phase compared favorably with the reference BET. Position differences were for the most part less than 300 meters and were at all times under the 570-meter certification criterion for active periods. None of the ascending node errors seen in earlier processing were evident during the deorbit preparation period.

The STS-30 reference BET consisted of the 19 data arcs described above, within which JSC Ground Navigation solved for vent forces to more accurately model trajectory perturbations. All two-TDRS-only BET arcs met the acceptance criteria, as shown in Table 10. The maximum position difference between the reference and two-TDRS-only BET solutions, seen in arc 12, was 536 meters. Minimum overlap position difference comparisons between successive BET solutions were normally less than 300 meters. Arcs 4, 5, 12, and 14 initially failed the acceptance criteria but passed after postmission reprocessing. It is interesting to note that every failure was due to inclination and ascending node errors and the associated cross-track position error.

The majority of orbital energy changes during the flight were positive. Solved-for vents accounted for semimajor axis changes ranging from 20 to 293 meters. The arc 2 solution also solved for the large OMS separation maneuver following deployment of the Magellan spacecraft.

3.2.2 GSFC/FDF Orbit Determination Results for STS-30

The GSFC/FDF orbit determination results for STS-30 are presented in Table 11, which gives the maximum position differences between the two-TDRS and reference solutions (with along-track thrust modeling included in the solutions for arcs 2 and 15). The deployment of the Magellan spacecraft during arc 2 was also modeled. The last column of Table 11 shows that all the arcs passed the certification criteria. Except for arc 2, the maximum overlap position differences between successive two-TDRS solutions ranged from 100 to 1560 meters. In summary, the GSFC/FDF analysis showed that all 19 of the 19 valid arcs satisfied the navigation certification criteria.

3.3 STS-32

STS-32 was launched into a circular 28.5-degree inclination, 352-kilometer altitude orbit on January 9, 1990. The major objective of this mission was to retrieve the Long-Duration Exposure Facility (LDEF) spacecraft and return it to Earth. As the certification efforts for STS-29 and STS-30 had already certified the two-TDRS configuration for nominal STS support, the certification effort for STS-32 was focused on the period around the rendezvous activities. Consequently, there were only five tracking data arcs used for navigation certification during this mission. The thrusting activity level, the start and stop times, and the number of GN and SN tracking passes for each tracking data arc are given in Table 12.

The GSFC tracking data evaluation is documented in Reference 6. A total of 167 C-band GN on-orbit passes were evaluated. Anomalies were encountered in two of the C-band passes. Approximately 72 percent of the

Table 9. Maximum Position Differences Between Two-TDRS BTB and Reference BET Solutions for STS-30 (JSC)

ARC NO.	VENTING ACTIVITY LEVEL	MAXIMUM POSITION DIFFERENCE BETWEEN THE TWO-TDRS AND REFERENCE SOLUTIONS (METERS)					NUMBER OF TWO-TDRS SOLUTIONS PASSED/ TOTAL NUMBER	PERCENT PASSED
		RADIAL	CROSS-TRACK	ALONG-TRACK	TOTAL (RSS)	CERTIFICATION CRITERION		
1	ACTIVE	116	-488	-38	503	* 420	5 / 7	71
2	ACTIVE	-87	605	-253	662	570	7 / 8	88
3	QUIET	41	136	-148	205	320	8 / 8	100
4	ACTIVE	-61	54	217	232	570	7 / 7	100
5	MODERATE	-19	14	-403	404	420	4 / 4	100
6	ACTIVE	10	378	386	540	570	8 / 8	100
7	ACTIVE	-8	433	381	577	570	6 / 7	86
8	QUIET	33	203	126	242	320	8 / 8	100
9	QUIET	-56	-158	130	213	320	3 / 3	100
10	ACTIVE	-37	-651	-11	653	570	6 / 7	86
11	ACTIVE	-98	181	-234	311	570	8 / 8	100
12	ACTIVE	135	-187	505	555	570	6 / 6	100
13	QUIET	-4	136	203	245	320	9 / 9	100
14	ACTIVE	81	-226	399	466	570	7 / 7	100
15	ACTIVE	-44	79	-342	354	570	8 / 8	100
16	ACTIVE	-31	206	84	225	570	8 / 8	100
17	QUIET	-8	-34	-212	215	320	8 / 8	100
18	ACTIVE	-4	-429	460	629	570	3 / 4	75
19	ACTIVE	48	28	265	270	570	5 / 5	100

* ARC 1 WAS A POSTINSERTION ARC

Table 10. Maximum Position Differences Between Two-TDRS BET and Reference BET Solutions for STS-30 (JSC)

ARC NO.	VENTING ACTIVITY LEVEL	MAXIMUM POSITION DIFFERENCE BETWEEN THE TWO-TDRS AND REFERENCE SOLUTIONS (METERS)					PASS/ FAIL
		RADIAL	CROSS-TRACK	ALONG-TRACK	TOTAL (RSS)	CERTIFICATION CRITERION	
1	ACTIVE	-8	96	-83	127	* 420	PASS
2	ACTIVE	-23	172	185	253	640	PASS
3	QUIET						
4	ACTIVE	-30	-141	22	145	640	PASS
5	MODERATE	20	-168	177	245	420	PASS
6	ACTIVE	-33	133	6	137	640	PASS
7	ACTIVE	-38	109	209	239	640	PASS
8	QUIET						
9	QUIET	-10	58	49	77	280	PASS
10	ACTIVE	-8	-95	-208	228	640	PASS
11	ACTIVE	-56	-234	209	319	640	PASS
12	ACTIVE	-88	-202	-489	536	640	PASS
13	QUIET						
14	ACTIVE	-135	62	-418	444	640	PASS
15	ACTIVE	7	-194	-231	302	640	PASS
16	ACTIVE	8	-4	113	114	640	PASS
17	QUIET						
18	ACTIVE	16	-144	165	220	640	PASS
19	ACTIVE	4	-121	112	165	640	PASS

* ARC 1 WAS A POSTINSERTION ARC

NOTE: TWO-TDRS BET COMPARISONS WERE NOT PERFORMED FOR QUIET ARCS 3, 8, 13, AND 17

Table 11. Maximum Position Differences Between Two-TDRS and Reference Solutions for STS-30 (With Along-Track Thrust Modeling) (GSFC/FDF)

ARC. NO.	VENTING ACTIVITY LEVEL	MAXIMUM POSITION DIFFERENCE BETWEEN THE TWO-TDRS AND REFERENCE SOLUTIONS (METERS)					PASS/ FAIL
		RADIAL	CROSS-TRACK	ALONG-TRACK	TOTAL	CERTIFICATION CRITERION	
1	ACTIVE	13	147	177	231	* 420	PASS
2	ACTIVE	9	306	370	480	640	PASS
3	QUIET	15	-85	99	132	280	PASS
4	ACTIVE	-21	-299	362	470	640	PASS
5	MODERATE	1	4	87	87	420	PASS
6	ACTIVE	12	114	-36	120	640	PASS
7	ACTIVE	24	279	-421	506	640	PASS
8	QUIET	8	-29	-38	48	280	PASS
9	QUIET	2	80	61	101	280	PASS
10	ACTIVE	29	410	476	629	640	PASS
11	ACTIVE	-40	-423	267	501	640	PASS
12	ACTIVE	15	254	222	338	640	PASS
13	QUIET	-1	-2	34	34	280	PASS
14	ACTIVE	6	-430	420	601	640	PASS
15	ACTIVE	20	172	-398	434	640	PASS
16	ACTIVE	1	126	111	168	640	PASS
17	QUIET	3	-22	58	63	280	PASS
18	ACTIVE	-1	-164	176	240	640	PASS
19	ACTIVE	-102	65	461	476	640	PASS

* ARC 1 WAS A POSTINSERTION ARC

NOTE: ALONG-TRACK THRUST MODELING WAS APPLIED FOR ARCS 2 AND 15.

Table 12. Tracking Arc Definition, Thrusting Activity Level, and Number of Tracking Passes for STS-32

ARC NO.	VENTING ACTIVITY LEVEL	TRACKING INTERVAL				NO. OF TRACKING PASSES			
		START		END		GROUND NETWORK		TDRSS	
		DATE	GMT (HHMMSS)	DATE	GMT (HHMMSS)	S-BAND/ C-BAND	TOTAL	TDRS-E/ TDRS-W	TOTAL
1	ACTIVE	1/11/90	12:28:20	1/11/90	18:42:00	0 / 10	10	3 / 4	7
2	QUIET		17:22:50		23:36:30	1 / 4	5	4 / 5	9
3	QUIET		22:10:00	1/12/90	04:25:50	1 / 3	4	4 / 5	9
4	ACTIVE	1/12/90	03:54:00		10:22:00	0 / 4	4	4 / 4	8
5	ACTIVE		09:32:50		15:04:30	0 / 10	10	3 / 5	8

TDRS-East view periods and 81 percent of the TDRS-West view periods had at least 70 percent usable data. With Doppler compensation and GN times deleted, the TDRS-East data were approximately 81 percent usable and the TDRS-West data were approximately 93 percent usable.

The JSC and GSFC orbit determination results for STS-32 are presented in Sections 3.3.1 and 3.3.2, respectively.

3.3.1 JSC Orbit Determination Results

The maximum position difference comparisons between the two-TDRS-only BTB and reference BET solutions are shown in Table 13. One TDRSS BTB solution exceeded the certification criterion out of 25 total solutions, corresponding to a 96-percent pass ratio. Arc 5, which contained three attitude maneuvers, two rendezvous burns, and four midcourse correction maneuvers, produced the sole two-TDRS-only BTB failure; this failure occurred in close proximity to a rendezvous maneuver.

Total position differences between the two-TDRS-only and reference BTB solutions were usually less than 152 meters. Inclination and ascending node differences were acceptable, indicating that the two-TDRS-only BTB processing determined the correct orbital plane.

The STS-32 rendezvous certification period included five reference and three two-TDRS-only BET arcs. Comparable vent forces and associated energy changes were obtained for the two-TDRS-only and reference BET solutions. Semimajor axis changes for the solved-for vents ranged from 62 to 5068 meters. Minimum overlap position differences between successive BET solutions were less than 150 meters.

Maximum position differences between the two-TDRS-only and reference BET solutions for arcs 1, 4, and 5 are shown in Table 14. Each arc satisfied the certification criteria. Arc 1 produced the largest total position difference of 439 meters.

3.3.2 GSFC/FDF Orbit Determination Results

The GSFC/FDF orbit determination results for STS-32 are presented in Table 15, which gives the maximum position differences between the two-TDRS and reference solutions (with along-track thrust modeling included for arcs 1, 4, and 5). The two quiet arcs (arcs 2 and 3) passed the certification criteria, but all three active arcs (arcs 1, 4, and 5) failed the certification criteria. Arc 1 showed the largest maximum position difference, which can be attributed to three large ventings in both the radial and cross-track directions. Arcs 4 and 5 also included large ventings in the radial and cross-track directions. This explains why the application of along-track thrust components in the GSFC solutions did not succeed in significantly reducing the maximum position differences for these arcs. Except for arc 1, the maximum overlap position differences between successive two-TDRS solutions ranged from 99 to 5066 meters. In summary, the GSFC/FDF analysis showed that two of the five arcs satisfied the two-TDRS navigation certification criteria. However, GSFC/FDF was unable to corroborate the JSC results for the three active arcs because of thrust modeling limitations.

4.0 SUMMARY AND CONCLUSIONS

This section presents the JSC and GSFC/FDF conclusions from this certification study and discusses additional considerations.

4.1 CERTIFICATION CONCLUSIONS

During STS-29, STS-30, and STS-32, JSC Ground Navigation certified that the two-TDRS network is an effective tool for Shuttle navigation. Two-TDRS-only BTB processing satisfied the certification requirements for over 90 percent of the solutions obtained during each flight. BET processing met the acceptance criteria for every data arc considered in the three flight certification efforts. The BET state vector and vent solutions were quite similar to those of the reference solutions, and direct comparisons showed that these results were uniform throughout the data arcs.

Table 13. Maximum Position Differences Between Two-TDRS BTB and Reference BET Solutions for STS-32 (JSC)

ARC NO.	VENTING ACTIVITY LEVEL	MAXIMUM POSITION DIFFERENCE BETWEEN THE TWO-TDRS AND REFERENCE SOLUTIONS (METERS)					NUMBER OF TWO-TDRS SOLUTIONS PASSED/ TOTAL NUMBER	PERCENT PASSED
		RADIAL	CROSS-TRACK	ALONG-TRACK	TOTAL (RSS)	CERTIFICATION CRITERION		
1	ACTIVE	-35	-164	-187	251	700	7 / 7	100
2	QUIET	-	-	-	-	N/A	-	-
3	QUIET	-	-	-	-	N/A	-	-
4	ACTIVE	55	88	-34	109	700	9 / 9	100
5	ACTIVE	43	698	359	786	700	8 / 9	89

NOTE: TWO-TDRS-ONLY SOLUTIONS WERE NOT PERFORMED DURING THE QUIET PERIODS.

Table 14. Maximum Position Differences Between Two-TDRS BET and Reference BET Solutions for STS-32 (JSC)

ARC NO.	VENTING ACTIVITY LEVEL	MAXIMUM POSITION DIFFERENCE BETWEEN THE TWO-TDRS AND REFERENCE SOLUTIONS (METERS)					PASS/ FAIL
		RADIAL	CROSS-TRACK	ALONG-TRACK	TOTAL (RSS)	CERTIFICATION CRITERION	
1	ACTIVE	-61	363	239	439	750	PASS
2	QUIET	-	-	-	-	N/A	-
3	QUIET	-	-	-	-	N/A	-
4	ACTIVE	-10	-11	24	28	750	PASS
5	ACTIVE	-146	382	-67	415	750	PASS

NOTE: TWO-TDRS-ONLY SOLUTIONS WERE NOT PERFORMED DURING THE QUIET PERIODS.

Table 15. Maximum Position Differences Between Two-TDRS and Reference Solutions for STS-32 (With Along-Track Thrust Modeling) (GSFC/FDF)

ARC NO.	VENTING ACTIVITY LEVEL	MAXIMUM POSITION DIFFERENCE BETWEEN THE TWO-TDRS AND REFERENCE SOLUTIONS (METERS)					PASS/ FAIL
		RADIAL	CROSS-TRACK	ALONG-TRACK	TOTAL	CERTIFICATION CRITERION	
1	ACTIVE	265	10,232	17,346	20,141	750	FAIL
2	QUIET	1	16	48	50	280	PASS
3	QUIET	20	-51	-32	64	280	PASS
4	ACTIVE	-251	-139	2,333	2,350	750	FAIL
5	ACTIVE	1,267	-1822	5,596	6,020	750	FAIL

The GSFC/FDF orbit determination results for STS-29, STS-30, and STS-32 corroborated the JSC certification results. The GSFC/FDF analysis showed that 43 of the 47 certification arcs passed the certification criteria. GSFC/FDF was unable to corroborate the JSC results for the remaining four certification arcs because of thrust modeling limitations.

The GSFC/FDF tracking data evaluation showed that approximately 80 percent of the TDRS-East and TDRS-West view periods had at least 70 percent usable data. With Doppler compensation and GN times deleted, the TDRS-East data were approximately 91 percent usable and the TDRS-West data were approximately 94 percent usable.

The two-TDRS certification effort has established the ability of TDRSS to detect and correct for unmodeled orbital energy changes. In the JSC BET processing, the presence of TDRSS data results in more precise vent solutions, because of the close proximity of the TDRSS data to the solved-for events. In the JSC BTB mode, the effects on orbital energy from drag mismodeling, translational effects from Shuttle venting and attitude control, attitude maneuvers, and burn mismodeling are picked up faster and more accurately than was the case with the ground-only network. In addition, the increased communications coverage from the two-TDRS network allows more timely and accurate translational maneuver confirmation, which leads to faster recovery of the JSC Ground Navigation state vector solution. Finally, TDRSS data, used in conjunction with C-band ground data, give good early state vector solutions. This capability has been used several times in the recent past to update and significantly improve the onboard state vector in the revolution following OMS-2.

4.2 CONCLUDING REMARKS

The successful two-TDRS certification effort has already resulted in a significant reduction in on-orbit C-band radar support. In some circumstances, however, the processing of TDRSS Doppler data alone has proved to be insufficient. A normal TDRSS tracking pass duration ranges from 35 to 60 minutes. The JSC Ground Navigation software automatically splits TDRSS batches in the event of a modeled translational maneuver or a change in the telemetry bit rate or transmitter frequency. The resulting shorter TDRSS batches often give adequate results during quiet and moderate activity periods, but a short TDRSS batch often does not give a good BTB solution during very active periods, such as a multiburn rendezvous sequence. In addition, two-TDRS-only BTB processing sometimes needs extra time to recover from orbital plane errors induced by unmodeled or poorly modeled translational maneuvers. In a two-TDRS-only environment, the requirement for accuracy may, in some cases, have to be traded off against the need for a timely state vector.

By comparison, a C-band ground pass will give at least a good local solution in under 10 minutes of tracking. As a result, C-band ground tracking is still required for periods having strict state vector accuracy requirements, for periods requiring state vectors soon after an event, such as a deployment, and for postmaneuver processing.

Another consideration for future C-band scheduling is the availability of usable TDRSS tracking data. For example, TDRS stationkeeping maneuvers and Shuttle attitude-related antenna blockages can result in periods of unusable TDRSS tracking data. During the STS-32 mission, two TDRS-West maneuvers resulted in unusable TDRS-West tracking data for two revolutions after the first maneuver and one revolution after the second maneuver. Consequently, the TDRS-East satellite tracking was augmented by ground C-band trackers during those periods.

5.0 ACKNOWLEDGMENTS

The authors wish to acknowledge the contributions of Jon Weaver (JSC), Osvaldo Cuevas (GSFC), William Heilman (RSOC), Dale Schultz (CSC), Taesul Lee (CSC), Elizabeth Smith (CSC), and Shuby Ambardekar (Bendix Field Engineering Corporation).

6.0 REFERENCES

1. TDRSS Orbit Determination and Navigation Working Group, *STS/TDRSS Navigation Certification Plan*, J. Teles (GSFC) and R. Osburn (JSC), February 1983

2. TDRSS Orbit Determination and Navigation Working Group, *The Joint JSC/GSFC STS/TDRSS Navigation Certification Interim Report: Single-TDRS Results*, J. B. Joyce (GSFC) and J. K. Weaver (JSC), April 1988

3. TDRS Orbit Determination and Navigation Working Group, *The Joint JSC/GSFC STS/TDRSS Navigation Certification Final Report: Two-TDRS Results*, J. B. Joyce (GSFC) and F. B. Lowes (JSC), to be published

4. TDRSS Orbit Determination and Navigation Working Group, *Addendum 1 to the Joint JSC/GSFC STS/TDRSS Navigation Certification Plan, Revision 1*, J. B. Joyce (GSFC) and F. B. Lowes (JSC), May 1989

5. Bendix Field Engineering Corporation, SSG 89-84, *STS-29 Metric Tracking Data Report*, S. P. Ambardekar, March 1989

6. Bendix Field Engineering Corporation, SSG 89-147, *STS-30 Metric Tracking Data Report*, S. P. Ambardekar, May 1989

7. Bendix Field Engineering Corporation, SSG 90-26, *STS-32 Metric Tracking Data Report*, S. P. Ambardekar, January 1990

IONOSPHERIC REFRACTION EFFECTS ON ORBIT DETERMINATION USING THE ORBIT DETERMINATION ERROR ANALYSIS SYSTEM*

C. P. Yee, D. A. Kelbel, T. Lee, and J. B. Dunham
COMPUTER SCIENCES CORPORATION

G. D. Mistretta
GODDARD SPACE FLIGHT CENTER

ABSTRACT

The influence of ionospheric refraction on orbit determination has been studied through the use of the Orbit Determination Error Analysis System (ODEAS). This paper presents the results of a study of the orbital state estimate errors due to the ionospheric refraction corrections, particularly for measurements involving spacecraft-to-spacecraft tracking links.

In current operational practice at the Goddard Space Flight Center (GSFC) Flight Dynamics Facility (FDF), the ionospheric refraction effects on the tracking measurements are modeled in the Goddard Trajectory Determination System (GTDS) using the Bent ionospheric model. While GTDS has the capability of incorporating the ionospheric refraction effects for measurements involving ground-to-spacecraft tracking links, such as those generated by the Ground Spaceflight Tracking and Data Network (GSTDN), it does not have the capability to incorporate the refraction effects for spacecraft-to-spacecraft tracking links for measurements generated by the Tracking and Data Relay Satellite System (TDRSS). The lack of this particular capability in GTDS raised some concern about the achievable accuracy of the estimated orbit for certain classes of spacecraft missions that require high-precision orbits. Using an enhanced research version of GTDS, some efforts have already been made to assess the importance of the spacecraft-to-spacecraft ionospheric refraction corrections in an orbit determination process. While these studies were performed using simulated data or real tracking data in definitive orbit determination modes, the study results presented in this paper were obtained by means of covariance analysis simulating the weighted least-squares method used in orbit determination.

The current operational version of ODEAS has the capability to compute ionospheric refraction corrections for range and range-rate measurements for both GSTDN and TDRSS data using the Bent ionospheric model. Using this capability in ODEAS, this study demonstrates how the magnitude and the characteristics of the spacecraft-to-spacecraft ionospheric refraction depend on such factors as spacecraft altitudes, solar activity, tracking geometry, and the local solar times of measurements. This study also provides error analysis results showing how the ionospheric refraction affects the spacecraft state estimate errors for routine orbit determination of spacecraft missions such as Gamma Ray Observatory (GRO), depending on the tracking geometries and the tracking pass lengths. ODEAS is a software system particularly convenient for analyses of this type. The results obtained from this study will provide a quick reference for ionospheric refraction effects on orbit determination and will be useful for assessing orbit accuracy requirements for many future spacecraft missions.

---

* This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, under Contract NAS 5-31500.

1.0 INTRODUCTION

The Earth's ionosphere extends from approximately 80 kilometers to beyond 1000 kilometers from the surface of the Earth and consists of ionized particles. As a tracking signal travels through this medium, the refraction or bending of the signal path occurs due to the presence of free electrons. The magnitude of refraction is dependent upon the electron densities along the path of the tracking signal and the frequency of the signal. To achieve high accuracy in determining a spacecraft orbit, it is therefore necessary to properly correct the tracking data for the refraction effects.

In the current operational practice at the Goddard Space Flight Center (GSFC) Flight Dynamics Facility (FDF), the ionospheric refraction effect is computed in the Goddard Trajectory Determination System (GTDS) for ground-to-spacecraft tracking links using a global electron density model known as the Bent model (References 1 and 2). However, no refraction correction is performed for spacecraft-to-spacecraft links for Tracking and Data Relay Satellite System (TDRSS) data; therefore, TDRSS tracking data are only partly corrected for ionospheric refraction. To circumvent this deficiency, a geometrical editing criterion can be employed in GTDS to reject the tracking data subject to significant ionospheric disturbance. This practice reduces the amount of tracking data actually used for the orbit determination process. It is therefore desirable to know the characteristics and the magnitude of the ionospheric refraction corrections and how they influence the accuracy of orbit determination. Such information will be useful to an orbit analyst in assessing the accuracy requirements of the spacecraft missions and will provide insights into ways of achieving more accurate spacecraft orbits by minimizing the orbit errors associated with the ionospheric refraction corrections. Using an enhanced research version of GTDS, some efforts have already been made to assess the importance of the spacecraft-to-spacecraft ionospheric refraction corrections in an orbit determination process (Reference 3). While these studies were performed using simulated data or real tracking data in definitive orbit determination modes, the study results presented in this paper were obtained by means of covariance analysis simulating the weighted least-squares method used in orbit determination.

This paper presents results obtained by using the Orbit Determination Error Analysis System (ODEAS). ODEAS is a general-purpose linear orbit determination error analysis system which can be used as a covariance analysis tool to analyze the orbit determination errors resulting from various systematic error sources and random noises. Throughout the study, the Earth's ionosphere was modeled as a nonhomogeneous medium based upon the Bent model. The study results can be divided into two parts. The first part presents the magnitude and the general characteristics of ionospheric refraction correction as a function of spacecraft-to-spacecraft tracking geometry, solar activity, orbit inclination, user spacecraft orbital height, and local times of TDRS and the user spacecraft at the time of measurement. The second part investigates the spacecraft orbit determination errors that can result from neglecting the ionospheric refraction effects in spacecraft-to-spacecraft links. Due to the linearity assumptions involved in the orbit estimation process, the results obtained can also be extended to address the effects of inexact modeling of ionospheric refraction in spacecraft-to-spacecraft tracking links by simple linear scaling. The results for a spacecraft mission such as the Gamma Ray Observatory (GRO) are discussed.

A brief background of the Bent model is provided in Section 1.1 of this paper. Section 1.2 describes the mathematical procedures used in ODEAS for modeling the ionospheric refraction correction for spacecraft-to-spacecraft links using the Bent model. Section 2 discusses the study results, and Section 3 presents conclusions and recommendations.

1.1 BENT MODEL

The Bent model is an empirical world-wide electron density model named after its original developers R. B. Bent et al. (Reference 2). This model was derived in the early 1970s from analysis of the many thousands of satellite and ground-based ionospheric soundings and satellite electron density measurements available at the time. It was extensively evaluated before incorporation into GTDS (Reference 4). The model provides the electron density profiles (electron density as a function of altitude) derived by incorporating the effects of variations in latitude, longitude, solar activity, geomagnetic activity, seasonal variations, and local time. Using this model, an electron density profile similar to that shown in Figure 1 can be constructed for any geographic location around the globe. This profile is represented in the Bent model using seven curve sections, including two parabolic sections and five exponential tail sections.

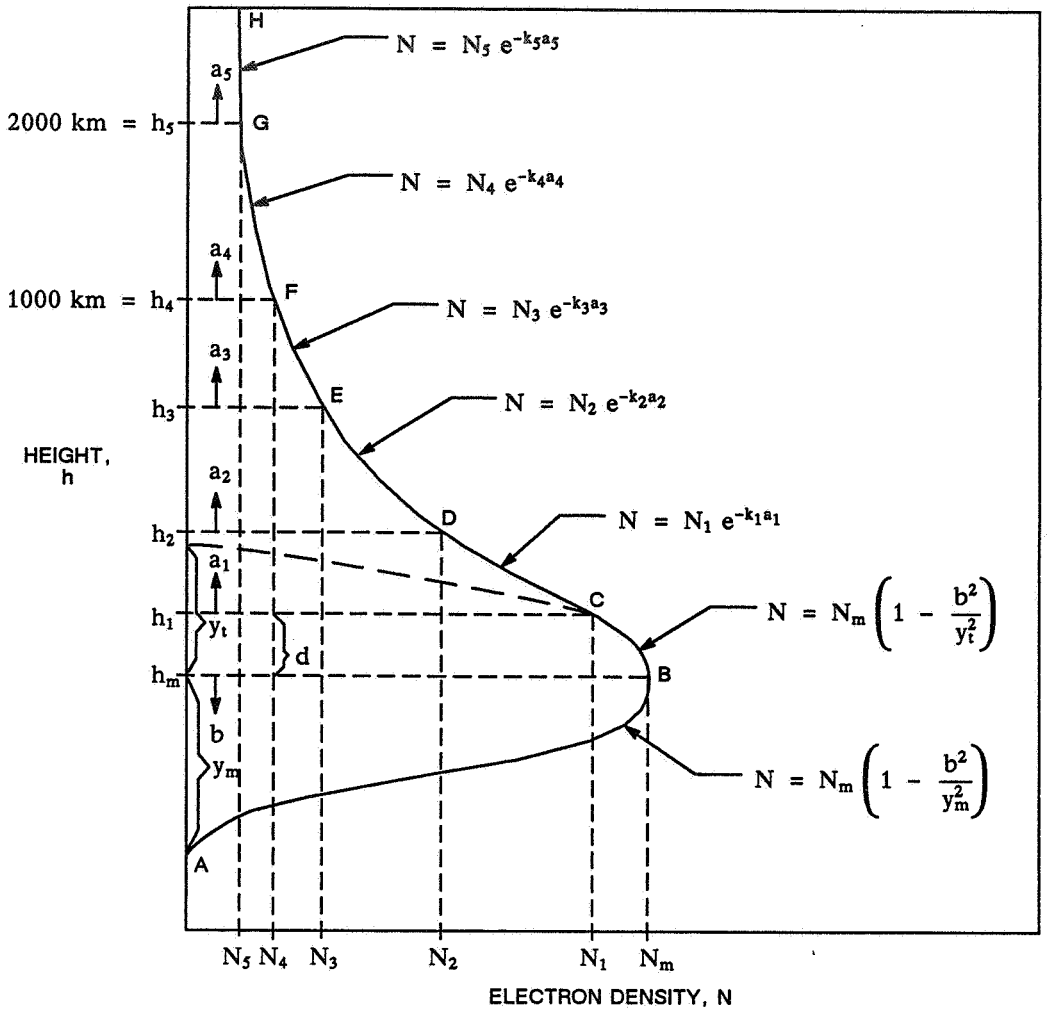


Figure 1. Empirical Worldwide Electron Density Profile

The computation of the electron density profile consists of evaluating parameters specifying the upper and lower parts of the parabolic section including the maximum density, N_m, the corresponding altitude, h_m, and the exponential constants specifying the five exponential tail sections. This model has been successfully used in GTDS for correcting the ground-to-spacecraft tracking links for ionospheric refraction. The detailed mathematical description of the model as implemented in GTDS is described in Reference 1. Although the model has been successful in correcting the ground-to-spacecraft tracking measurements, the validity of the model for correcting spacecraft-to-spacecraft measurements, especially for high-altitude regions, has yet to be thoroughly evaluated for application in GTDS.

The following section discusses the ionospheric refraction corrections using the Bent model for the spacecraft-to-spacecraft links as implemented in ODEAS. Reference 5 provides the mathematical details of the model implementation.

1.2 IONOSPHERIC REFRACTION CORRECTION FOR SPACECRAFT-TO-SPACECRAFT LINKS

The ionospheric refraction effects can be characterized in terms of the variable local index of refraction, n, of the medium through which the signal is propagated. This index of refraction can be expressed as

$$n = 1 - N_I \tag{1}$$

where N_I is the refractivity and is given in terms of the electron density, N (electrons per cubic meter), and the signal transmission frequency, f (hertz), as follows:

$$N_I = 40.3 \frac{N}{f^2} \tag{2}$$

For a given measurement, the signal frequency, f, is a constant, and the electron density, N, varies with the altitude, longitude, latitude, and local time of the signal path. The range correction due to ionospheric refraction, $\Delta\varrho_I$, can be computed by integrating N_I as expressed in Equation (2) along the signal path, as follows:

$$\Delta\varrho_I = \frac{40.3}{f^2} \int N \, ds \tag{3}$$

where ds denotes an infinitesimal ray path length.

For a spacecraft-to-spacecraft tracking link, the integral expressed in Equation (3) must be evaluated by a numerical method, because the electron density, N, cannot easily be expressed as a function of the ray path, s. In ODEAS, this integral is evaluated using a 12-point Gaussian quadrature scheme.

Figure 2 illustrates the two tracking configurations for a spacecraft-to-spacecraft tracking link that are relevant for the effects of ionospheric refraction. For configuration 1, (β > 90 degrees), the integral of Equation (3) is evaluated from the user spacecraft at U to the maximum ionospheric height at Q, which is set at 3000 kilometers. For configuration 2, (β < 90 degrees), the integral is evaluated in two parts divided at point P, the point of lowest approach for the signal path. One integration is done from point P to Q and the other from point P to U.

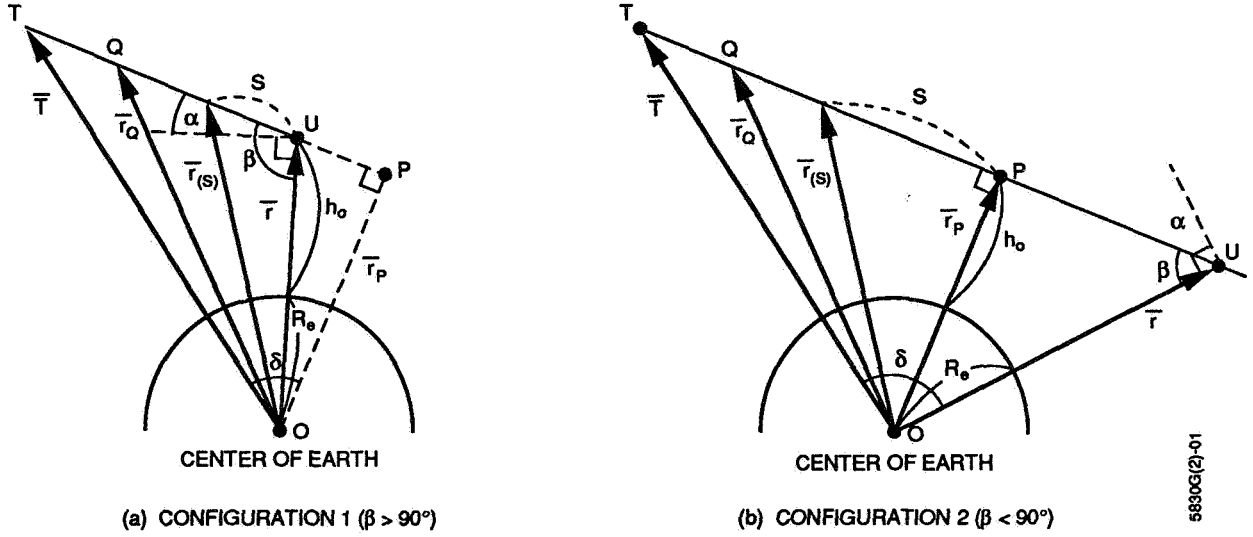


Figure 2. Spacecraft-to-Spacecraft Segment (Two Possible Configurations Shown)

In the ODEAS implementation, to achieve computational efficiency, a new electron density profile may not always be recomputed at each integration point along the ray path. A new electron density profile is computed for point s_i whenever the geographic longitude and latitude of this point are different from those of the previous integration point on the ray path, s_{i-1}, by more than a specified tolerance, i.e.,

$$|\phi_i - \phi_{i-1}| + |\lambda_i - \lambda_{i-1}| \geq \text{Tol} \tag{4}$$

where

(ϕ_i, λ_i) and $(\phi_{i-1}, \lambda_{i-1})$ = geographic latitude and longitude pairs for the points s_i and s_{i-1}, respectively

Tol = user-specified tolerance level

Therefore, by adjusting the tolerance level, the number of electron density profiles computed in the integration can be controlled. Throughout the study, the tolerance level for recomputing the electron density profile was set to 1 degree. This setting ensures recomputation of electron density profiles at each integration point along the ray path for most of the tracking geometries except for those where the integration points are very close to each other, such as in the case of large β angles, in which case the range ionospheric corrections obtained by setting Tol = 1 degree and Tol = 0.001 degree differed by less than 2 percent.

Range-rate corrections can be computed from range corrections as follows:

$$\Delta \dot{\varrho}_I = - \frac{\Delta\varrho_I(t_i) - \Delta\varrho_I(t_{i-1})}{\Delta t} \tag{5}$$

where

$\Delta\varrho_I(t_i)$ and $\Delta\varrho_I(t_{i-1})$ = two successive range corrections computed at time t_i and t_{i-1}, respectively

Δt = time difference between two measurements (usually the Doppler count interval)

2.0 RESULTS

The results of this study are presented in two parts. Section 2.1 discusses the general characteristics of ionospheric refraction correction as a function of various parameters. Section 2.2 describes the orbit determination error analysis results that can be expected by ignoring the ionospheric refraction effects in spacecraft-to-spacecraft tracking links using the routine orbit determination scenario of a GRO-type spacecraft mission as an example.

2.1 GENERAL CHARACTERISTICS OF IONOSPHERIC REFRACTION CORRECTIONS

As stated earlier, the magnitude of ionospheric refraction correction is a function of tracking signal frequency and the electron densities along the signal path as it traverses the ionosphere. This is expressed mathematically in Equation (3). For a given measurement, the frequency, f, is a constant; and the ionospheric refraction correction is complicated due to the nature of electron density representation. This section presents the magnitude and the general characteristics of ionospheric refraction correction as a function of various parameters that affect the electron densities. The parameters studied were limited to spacecraft-to-spacecraft tracking geometry, solar activity, orbital height and inclination, and local solar times of TDRS and the user spacecraft at the time of measurement. Although the ionospheric refraction can also be affected by the seasonal variation, this effect was not studied parametrically.

To study the effects of these parameters, 18 different specially designed circular orbits were investigated using ODEAS. These basic spacecraft orbits were constructed by the combination of six different orbital heights (350, 550, 750, 950, 1150, and 1350 kilometers) and three different inclinations (28.5, 63.14, and 99.03 degrees). For each basic run scenario, a 24-hour definitive period was simulated using TDRSS two-way range and range-rate tracking. The tracking signals were modeled as S-band signals, with a frequency of 2100 megahertz, and a data sampling rate of 30 seconds. As shown in Equation (3), the ionospheric refraction correction is proportional to the inverse of the signal frequency squared; and, therefore, the results obtained can be scaled appropriately for different signal frequencies. For example, the ionospheric refraction correction value obtained with a K-band signal would yield a value approximately 50 times smaller than the S-band signal.

A measurement was scheduled whenever the user spacecraft was visible to the TDRS. This was done to ensure the collection of a fairly large sample of tracking data with a variety of tracking geometries. Ionospheric refraction corrections were computed for all the spacecraft-to-spacecraft range and range-rate data. The results obtained are described below.

2.1.1 Tracking Geometry

The spacecraft-to-spacecraft tracking geometry can be characterized by the relative positions of the relay spacecraft (TDRS), the user spacecraft, and the Earth. In the TDRSS environment, the relay spacecraft is in geosynchronous orbit at a height of approximately 35,800 kilometers above the Earth's surface. Figure 2 illustrates the two tracking geometries. In Figure 2a, both the relay and the user spacecraft are on the same side of the Earth; whereas in Figure 2b, the relay and the user spacecraft are on the opposite sides of the Earth.

The elevation angle of the relay with respect to the user, α, can be defined as the line-of-sight elevation angle of the relay spacecraft as measured from the local horizontal plane at the user. This elevation angle can be computed by

$$\alpha = \beta - 90 \text{ degrees} \tag{6}$$

With this definition, the elevation angle, α, assumes a positive value in Figure 2a and a negative value in Figure 2b.

The height of ray path, h_0, is the height of a point of closest approach on the ray path. For Figure 2a, h_0 is the same as the user spacecraft orbital height. The central angle, δ, is defined as the angle between the position vectors of TDRS and the user spacecraft, subtended at the center of the Earth.

It can be seen in Figure 2 that as the elevation angle, α, decreases, there is a greater possibility that the ray path will traverse a more "diverse" ionosphere, which consists of many different electron density profiles due to bigger possibility of variations in latitude, longitude, and local times. Conversely, as the elevation angle increases, the ray path encounters the ionosphere with less diversity. Figures 3a and 3b illustrate the average ionospheric refraction for range and range-rate data as a function of elevation angle for various orbital heights at a 28.5-degree inclination. The ionospheric refraction correction value plotted at a given orbital height at a given elevation angle was obtained as an average value in the case of range correction, or as a root-mean-square (RMS) value in the case of range-rate correction, from the collection of tracking data taken at that specified elevation angle during a 24-hour period. RMS values were used for range-rate data because the ionospheric refraction corrections for range-rate data could assume positive and negative values, and the RMS can represent the mean magnitude of the correction regardless of sign.

Figures 4a, 4b and Figures 5a, 5b illustrate the same type of plots for orbital inclinations of 63.14 and 99.03 degrees. To see the spread of ionospheric refraction correction values obtained at each elevation angle, Figures 6a and 6b illustrate the maximum, minimum, and average or RMS values of the ionospheric corrections for a 28.5-degree inclination, 350-kilometer altitude orbit.

All these figures demonstrate the fact that the ionospheric refraction correction for both range and range-rate measurements decreases as the elevation angle increases. The elevation angle dependence is more pronounced for lower orbital heights (350 and 550 kilometers) and less pronounced for higher altitudes (1150 and 1350 kilometers).

Figures 3 through 5 also illustrate the functional dependence of ionospheric refraction corrections on the orbital height. It can be seen from these figures that the magnitude of corrections are generally higher for lower orbital heights. At a 350-kilometer altitude, the average ionospheric refraction corrections for range measurements are about 1.5 to 2 times as high as those values observed at a 550-kilometer altitude. As the orbital height increases, the correction values become smaller and its dependence on the orbital height become less significant. For the spacecraft orbital heights of 950, 1150, and 1350 kilometers, the corrections are found to be in the same neighborhood.

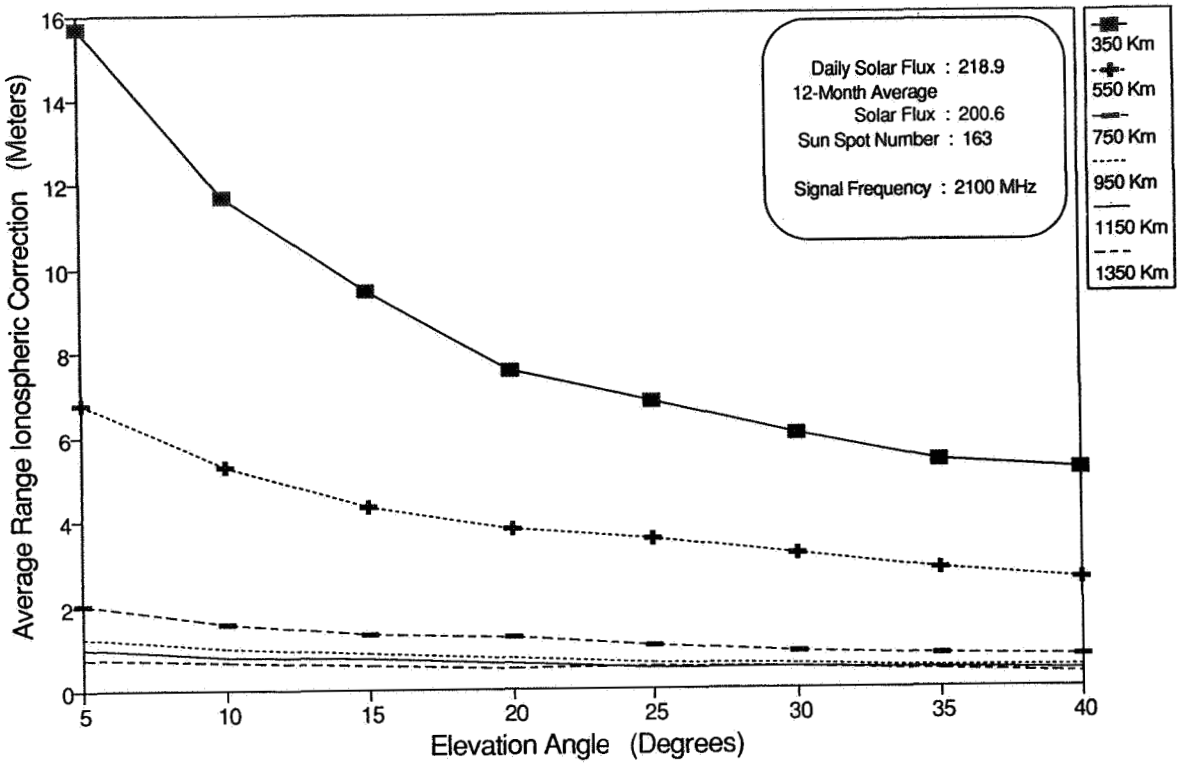


Figure 3a. Average Range Ionospheric Correction Versus Elevation Angle for a 28.5-Degree Inclination for Various Altitudes

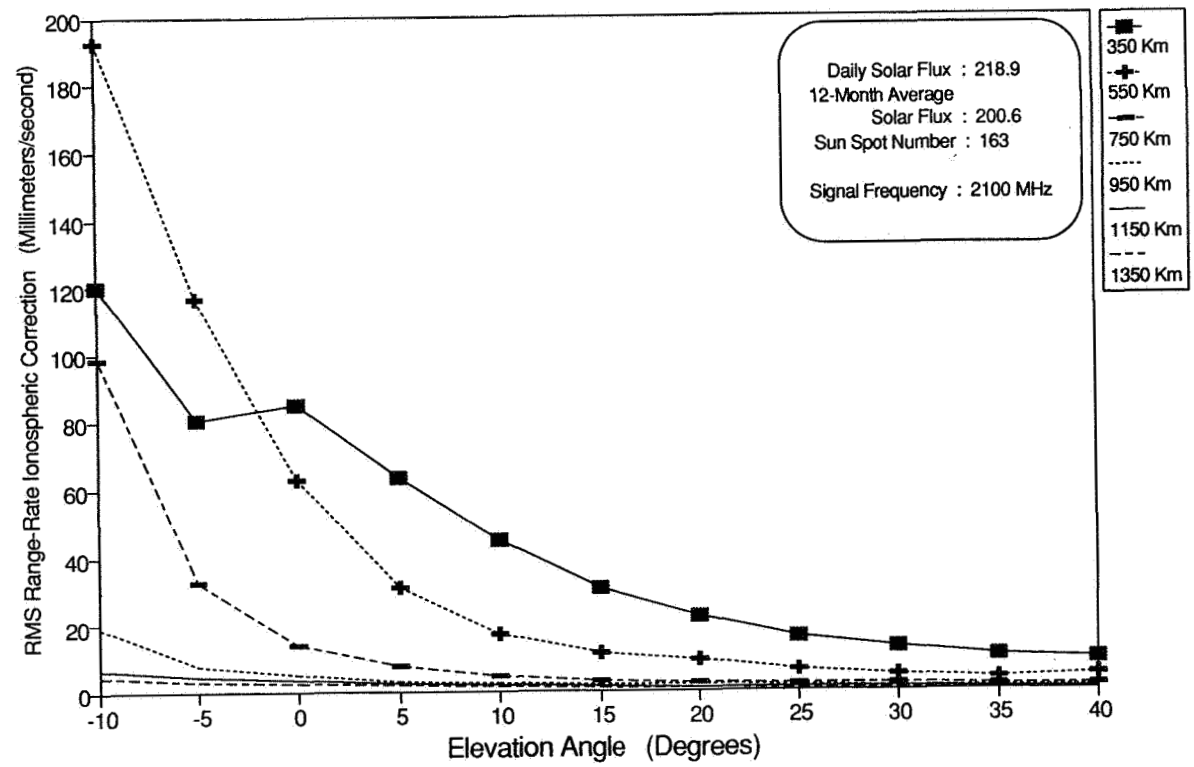


Figure 3b. RMS Range-Rate Ionospheric Correction Versus Elevation Angle for a 28.5-Degree Inclination for Various Altitudes

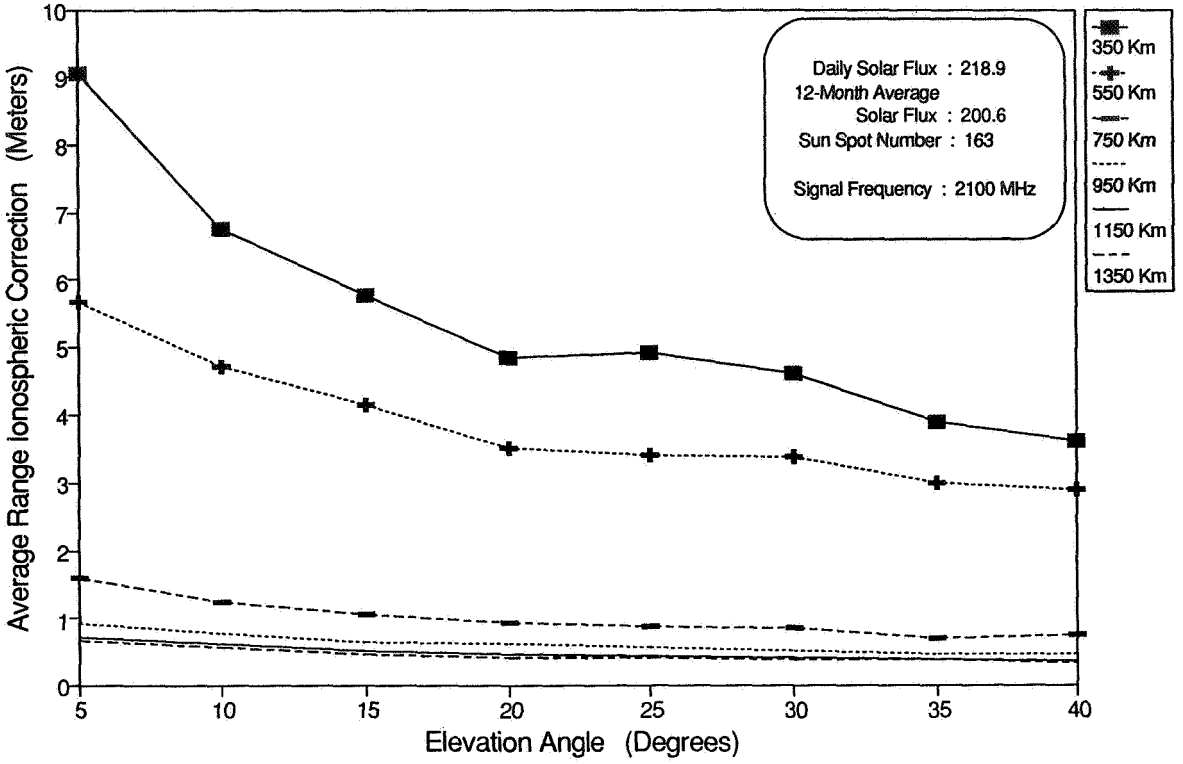


Figure 4a. Average Range Ionospheric Correction Versus Elevation Angle for a 63.14-Degree Inclination for Various Altitudes

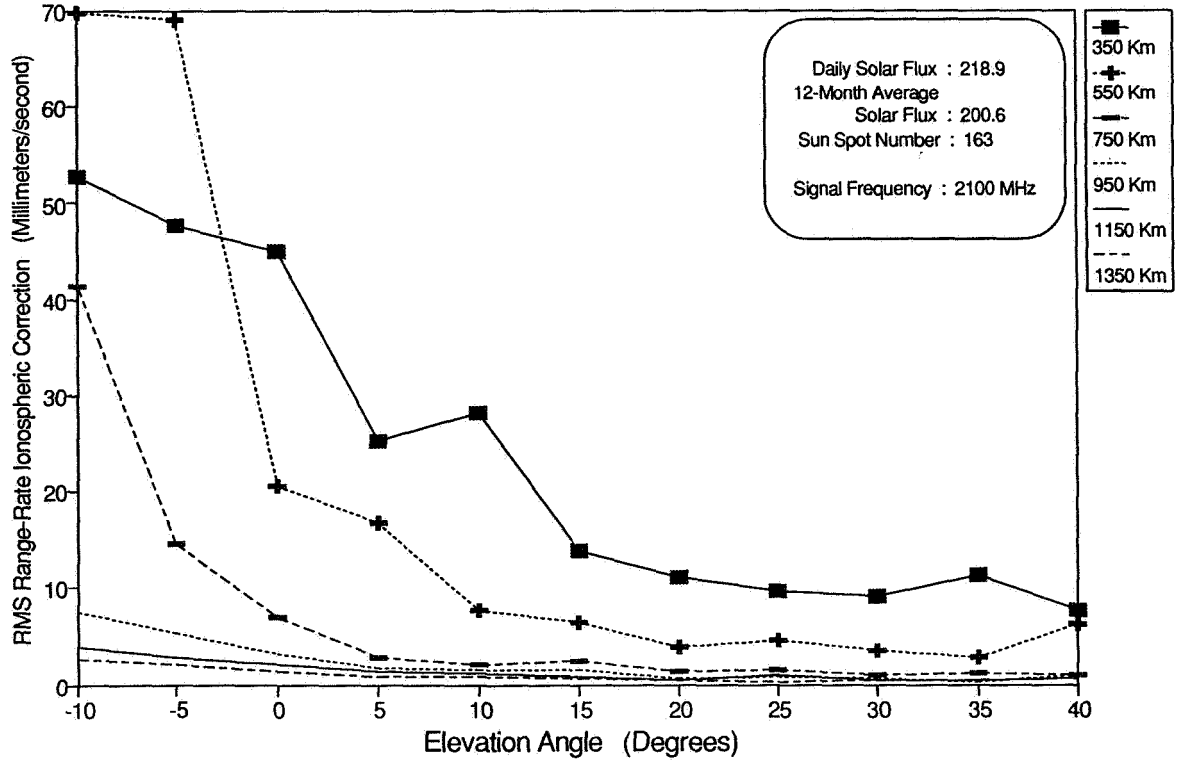


Figure 4b. RMS Range-Rate Ionospheric Correction Versus Elevation Angle for a 63.14-Degree Inclination for Various Altitudes

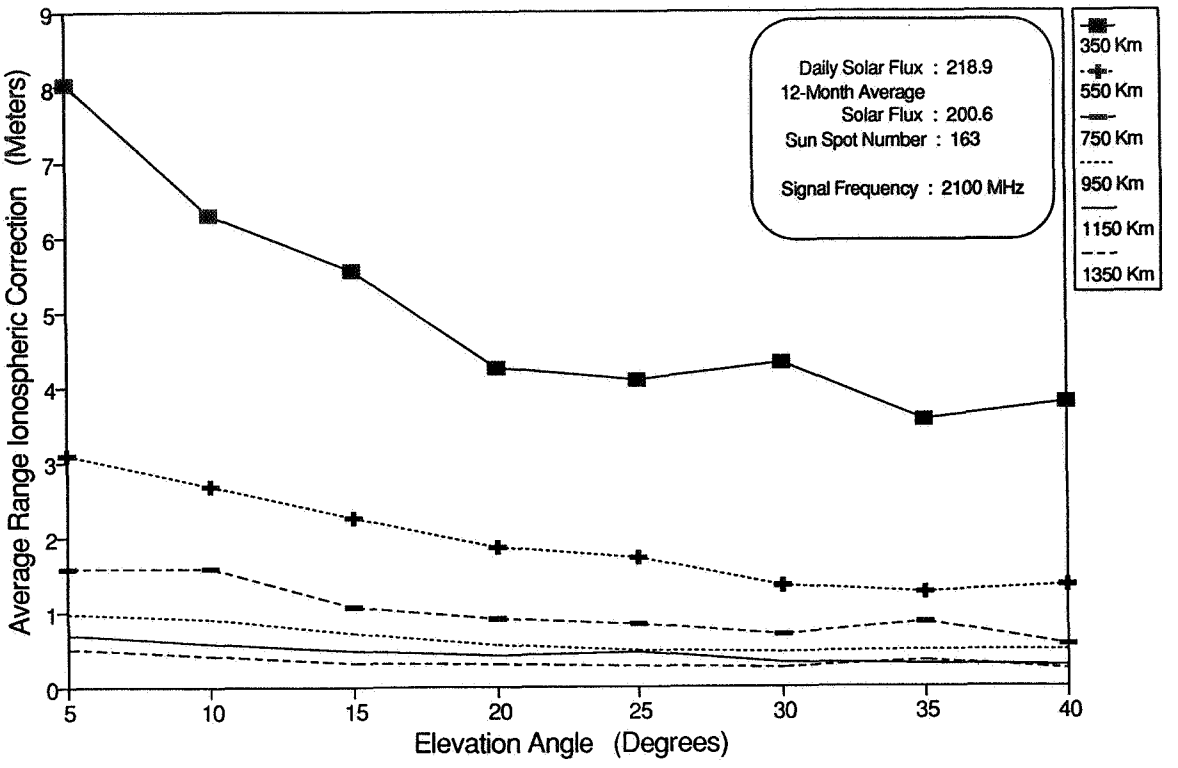


Figure 5a. Average Range Ionospheric Correction Versus Elevation Angle for a 99.03-Degree Inclination for Various Altitudes

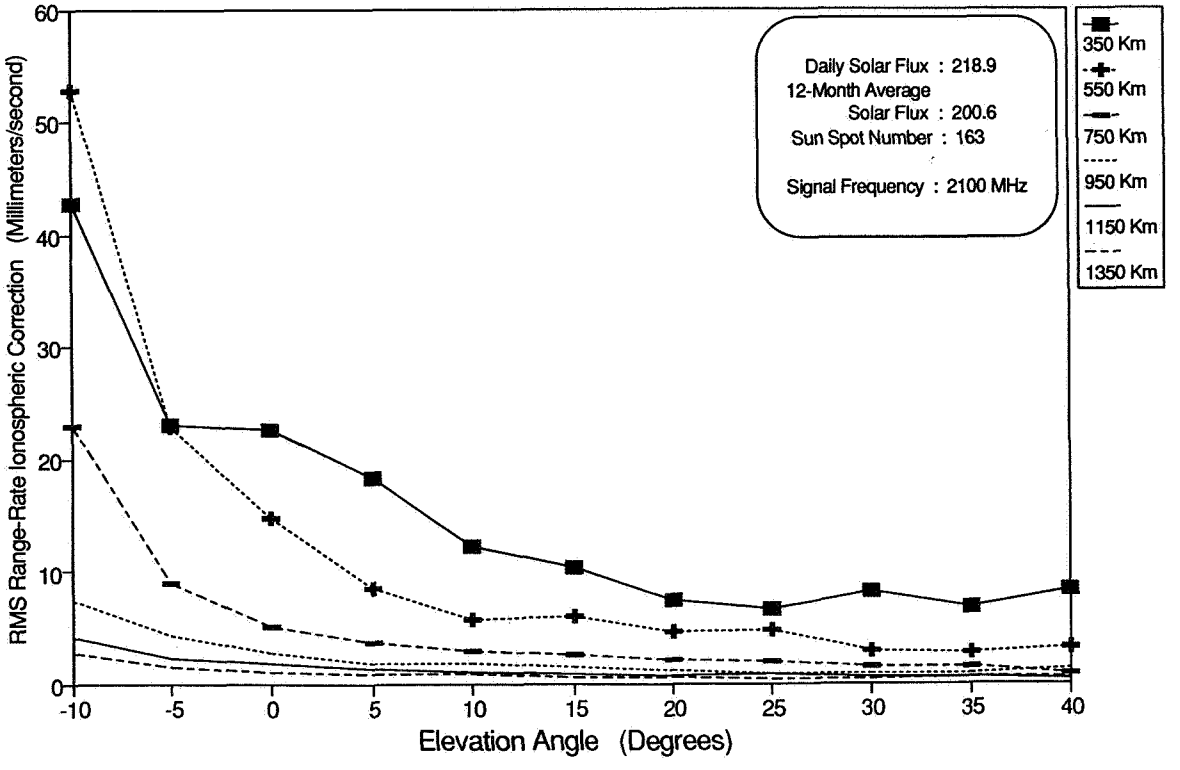


Figure 5b. RMS Range-Rate Ionospheric Correction Versus Elevation Angle for a 99.03-Degree Inclination for Various Altitudes

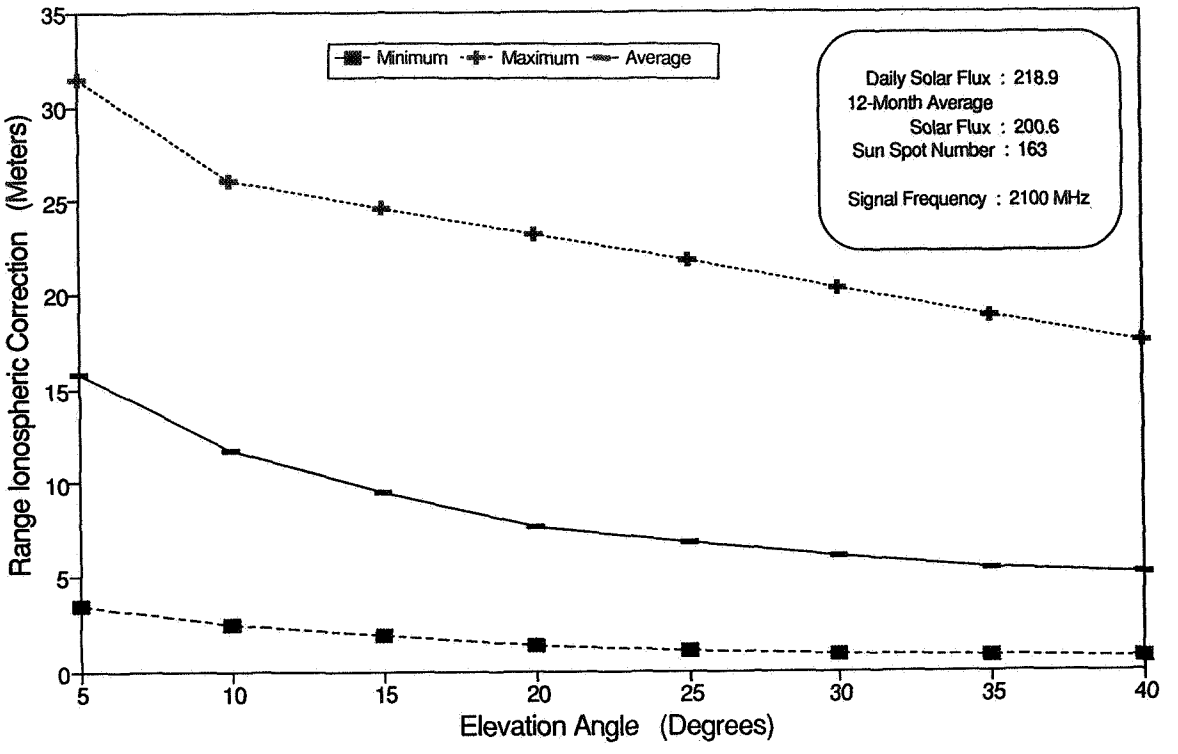


Figure 6a. Minimum, Maximum, and Average Range Ionospheric Correction Versus Elevation Angle for a 28.5-Degree Inclination and a 350-Kilometer Altitude

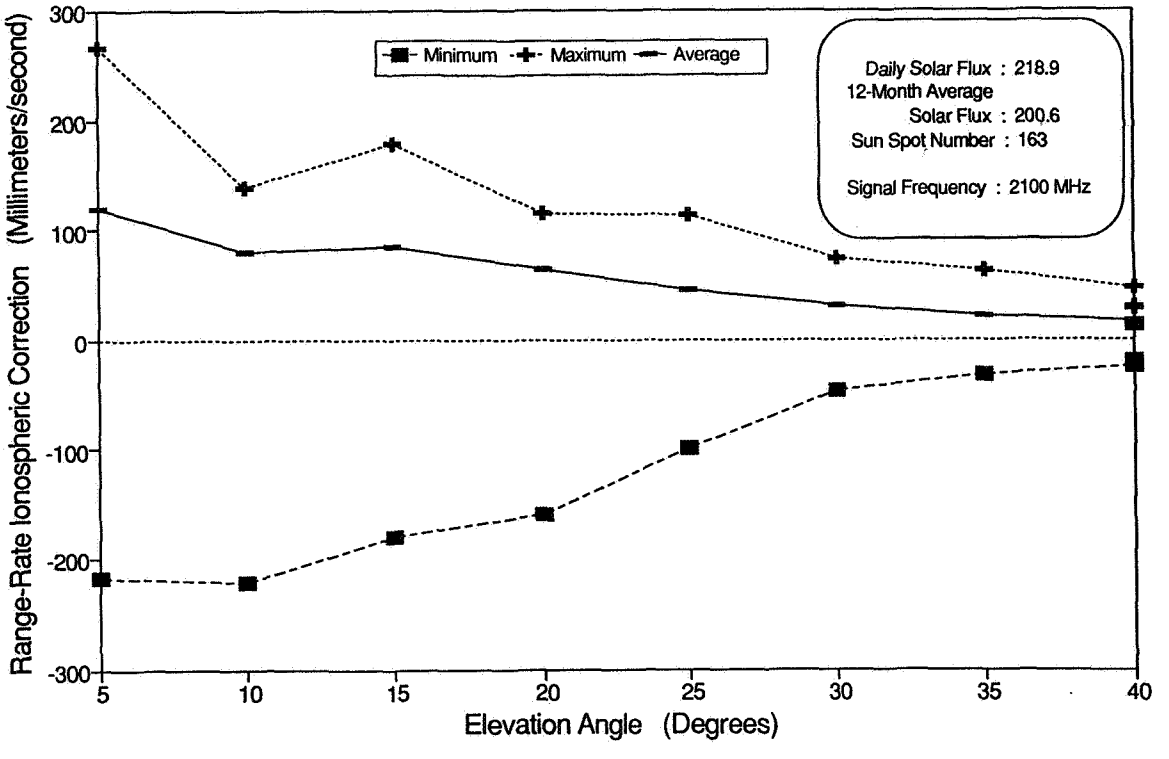


Figure 6b. Minimum, Maximum, and RMS Range-Rate Ionospheric Correction Versus Elevation Angle for a 28.5-Degree Inclination and a 350-Kilometer Altitude

As noted earlier, the upper bound of the ionosphere was set at 3000 kilometers. This choice is somewhat arbitrary. Although the contributions from altitudes above 3000 kilometers are not included in the results presented in this paper, their magnitudes can be estimated by evaluating the integral of Equation (3) all the way to the TDRS altitude. Analyzing these magnitudes obtained for a few sample tracking passes, it was found that they could vary significantly according to the altitudes of the user spacecraft. For a spacecraft at an altitude of 350 kilometers, for example, the contribution from above 3000 kilometers to the measurement correction was, in general, less than 5 percent of the contribution from below 3000 kilometers. However, for a spacecraft at 1350 kilometers, the contribution from above 3000 kilometers was found to be as high as 40 percent of the contribution from below 3000 kilometers.

2.1.2 Local Solar Times of Measurement

The ionospheric electron density for any location over the Earth's surface is also a function of the local solar time. This effect has been incorporated in the Bent electron density model. Because of this functional dependence, it is conceivable that the ionospheric refraction corrections for a particular range or range-rate measurement can be dependent on the local solar times of the relay and user spacecraft. For the spacecraft orbits studied, the ionospheric refraction correction was found to be affected more significantly by the local times of user spacecraft, with maximum corrections occurring around local times of 12 to 16 hours. This effect was consistently observed for all the spacecraft orbits studied.

Figure 7 shows the average range ionospheric refraction corrections as a function of the local solar time of the user spacecraft at the time of measurement. The spacecraft orbits shown are at three different orbital inclinations—28.5, 63.14, and 99.03 degrees—at a fixed altitude of 350 kilometers. This figure indicates a big increase in ionospheric refraction correction values around the user spacecraft local times of 12 to 16 hours, peaking at approximately 15 hours. The same trend was also observed for RMS range-rate corrections.

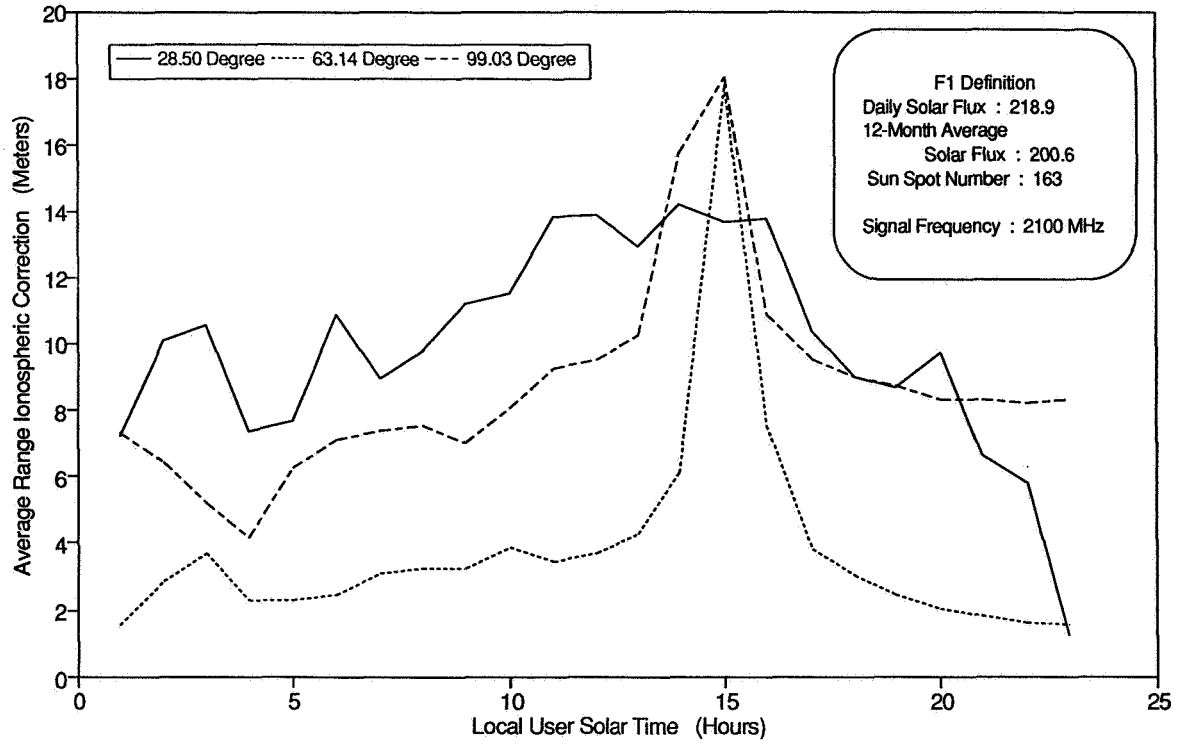


Figure 7. Average Range Ionospheric Correction Versus Local User Solar Time for a 350-Kilometer Altitude and Various Inclinations

If this phenomenon manifested by the Bent model closely matches the real ionospheric behavior, this may suggest that to effectively exclude the tracking data subject to high ionospheric disturbance, it might also be necessary to take into account the local time of the user spacecraft as a data editing criterion, together with the

tracking geometry considerations. A data editing criterion based on tracking geometry characterized by the height of ray path and the central angle is presently employed in the operational GTDS. In this practice, tracking data obtained at a height of ray path below a certain prescribed value and a central angle greater than a certain prescribed value are excluded. The prescribed values used are specific to mission requirements.

2.1.3 Solar Activity

To study the effects of solar activity on the ionospheric refraction correction, three sets of solar data were used, as summarized in Table 1. F1, F2, and F3 represent cases of moderately high, medium, and low solar activities, respectively.

Table 1. Solar Data Used for Ionospheric Refraction Simulations

SIMULATION NUMBER	DAILY SOLAR FLUX*	12-MONTH AVERAGE SOLAR FLUX*	SUN SPOT NUMBER
F1	218.9	200.6	163
F2	159.2	111.4	56
F3	74.4	76.1	20

* UNITS = 10^{-22} WATTS/METER^2/HERTZ

Figures 8 and 9 illustrate the average ionospheric refraction correction of range data as a function of elevation angle and solar activity for 350-kilometer and 1350-kilometer altitudes, respectively, at a 63.14-degree inclination. It can be seen that the ionospheric refraction correction values increase as the solar activity increases. For range corrections as shown in Figure 8, the values for the F1 curve are about two times higher than the corresponding F2 curve, and F2 is about two times larger than the corresponding F3 curve. Similar trends were also observed for RMS range-rate corrections.

To study the sensitivity of ionospheric refraction corrections to the daily and 12-month average or mean solar flux values, range corrections for two tracking passes, one for a 350-kilometer orbit and one for a 1350-kilometer orbit, were computed using different daily and mean solar flux values. The daily and mean solar flux values were varied in steps of 50 units. Whenever the mean solar flux was changed, the corresponding sunspot number was also changed according to an empirical formula that relates the sunspot number to the mean solar flux (Reference 3).

The results are summarized in Table 2. The values in cases 1, 3, and 5 were obtained using a common daily solar flux level of 218.9 with different mean solar flux levels. The values in cases 2, 3, and 4 were obtained with the same mean solar flux level of 200.6 and different daily solar flux levels. These results indicate that, in general, the refraction correction depends more strongly on the mean solar flux than on the daily flux. This general trend, however, does not appear to hold for the part of a pass with lower elevation angles.

2.1.4 Orbital Inclination

From Figures 3 through 5, it can be determined that the average or RMS values of ionospheric refraction obtained at various elevation angles for low-inclination (28.5-degree) orbits tend to be higher than those of high-inclination (99.03-degree) orbits. This tendency is found to be more pronounced in the low-altitude region (350 kilometers). At the 1350-kilometer altitude, the ionospheric refraction corrections for all three orbital inclinations are in the same neighborhood.

The observed dependence of the measurement refraction corrections on orbital inclination could be, at least partially, due to the strong dependence of the maximum electron density (N_m) and the corresponding height (h_m) on the latitude. Table 3 presents the values of these two parameters as a function of latitude. These results were obtained using the Bent model at a longitude of 318 degrees and a local solar time of 15 hours.

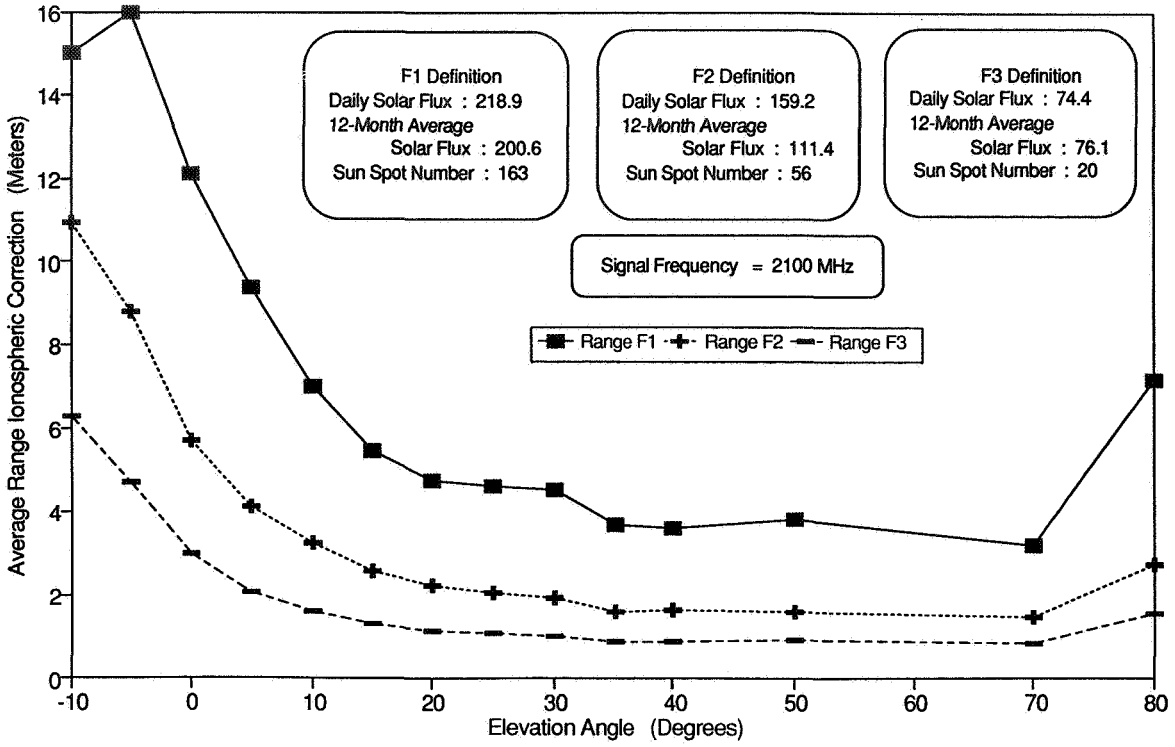


Figure 8. Average Range Ionospheric Correction Versus Elevation Angle for a 63.14-Degree Inclination, a 350-Kilometer Altitude, and Various Flux Levels

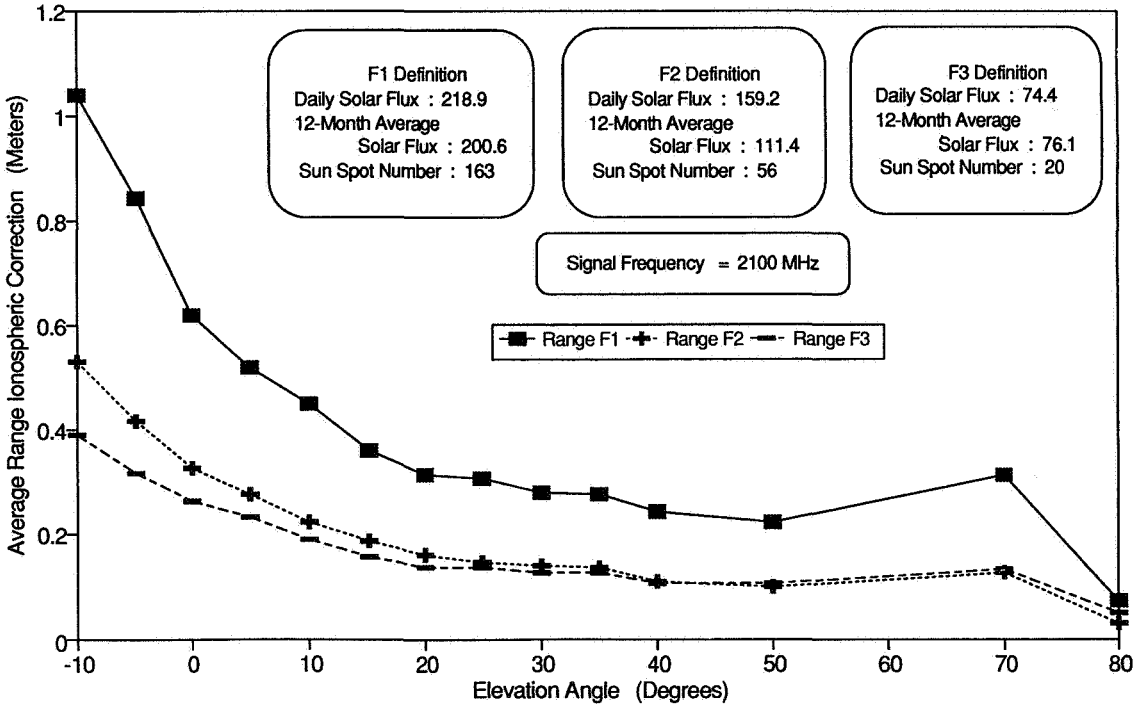


Figure 9. Average Range Ionospheric Correction Versus Elevation Angle for a 63.14-Degree Inclination, a 1350-Kilometer Altitude, and Various Flux Levels

Table 2. Daily and Mean Solar Flux Dependence of Ionospheric Refraction Corrections for Range Measurements

EPOCH: FEBRUARY 1, 1990; $0^h0^m0^s$
ORBIT: 350-KILOMETER ALTITUDE; 28.5-DEGREE INCLINATION

TIME (MINUTES FROM EPOCH)	ELEVATION ANGLE (DEGREES)	RANGE CORRECTIONS (METERS)				
		CASE 1	CASE 2	CASE 3	CASE 4	CASE 5
5	7.6	15.15	20.23	24.05	28.23	24.42
10	24.6	2.35	3.64	4.20	4.87	5.76
15	41.4	1.08	1.61	1.79	1.97	2.52
20	55.6	0.80	1.14	1.26	1.39	1.73
25	61.3	0.74	1.02	1.13	1.25	1.55
30	53.7	1.22	1.70	1.96	2.20	2.77
35	38.8	2.00	2.92	3.58	4.34	5.12
40	22.0	1.58	2.14	2.47	2.74	3.11
45	5.0	5.10	7.04	8.13	9.35	10.01
DAILY SOLAR FLUX*		218.9	168.9	218.9	268.9	218.9
12-MONTH MEAN SOLAR FLUX*		150.6	200.6	200.6	200.6	250.6

EPOCH: FEBRUARY 1, 1990; $0^h0^m0^s$
ORBIT: 1350-KILOMETER ALTITUDE; 28.5-DEGREE INCLINATION

TIME (MINUTES FROM EPOCH)	ELEVATION ANGLE (DEGREES)	RANGE CORRECTIONS (METERS)				
		CASE 1	CASE 2	CASE 3	CASE 4	CASE 5
5	2.6	0.61	0.89	1.09	1.33	1.51
10	16.4	0.18	0.26	0.29	0.34	0.39
15	30.1	0.10	0.13	0.13	0.14	0.17
20	43.3	0.07	0.09	0.09	0.09	0.11
25	54.6	0.07	0.08	0.08	0.08	0.09
30	60.5	0.07	0.08	0.08	0.09	0.09
35	57.4	0.08	0.09	0.09	0.10	0.11
40	47.4	0.10	0.12	0.14	0.17	0.21
45	34.6	0.12	0.14	0.15	0.18	0.19
50	20.9	0.15	0.18	0.19	0.19	0.21
DAILY SOLAR FLUX*		218.9	168.9	218.9	268.9	218.9
12-MONTH MEAN SOLAR FLUX*		150.6	200.6	200.6	200.6	250.6

*SOLAR FLUX UNITS = 10^{-22} WATTS/METER^2/HERTZ

As can be seen from this table, both h_m and N_m, especially the latter, increase in the lower latitude region. The parameter N_m acts as a scale factor for the entire electron density profile, thus affecting the electron densities computed in all different height regions. The variation of the maximum electron density height (h_m) can also contribute to an increase in the refraction corrections for orbits with lower inclinations. The values of N_m and h_m as a function of latitude were examined for several other longitudes, and all showed the trend similar to that described in Table 3.

Table 3. Latitude Variation of the Maximum Electron Density (N_m) and the Corresponding Height (h_m)

LATITUDE (DEGREES)	h_m (KILOMETERS)	N_m (10^{10} ELECTRONS/ METER^3)	LATITUDE (DEGREES)	h_m (KILOMETERS)	N_m (10^{10} ELECTRONS/ METER^3)
88.47	382.13	60.708906	-88.40	379.49	33.860825
82.38	363.65	49.976265	-78.17	361.63	45.700007
73.69	342.92	65.968391	-70.16	352.95	54.567174
61.99	325.73	118.187695	-59.07	346.42	79.233256
51.39	319.81	170.900019	-47.15	345.39	124.645481
44.31	320.52	192.604224	-38.04	349.87	164.401967
34.30	328.01	201.218013	-30.12	361.66	196.020009
21.03	352.92	213.455625	-18.44	409.45	215.195575
19.38	398.08	272.916559	-7.19	464.41	170.750405

NOTES: EPOCH = FEBRUARY 1, 1990
DAILY SOLAR FLUX = 218.9 (10^{-22} WATTS/METER^2/HERTZ)
12-MONTH MEAN SOLAR FLUX = 200.6 (10^{-22} WATTS/METER^2/HERTZ)
LONGITUDE = 318 DEGREES
LOCAL SOLAR TIME = 15^h

2.2 EFFECTS OF NEGLECTING SPACECRAFT-TO-SPACECRAFT IONOSPHERIC REFRACTION ON ORBIT DETERMINATION ACCURACY

As stated in Section 1, the current operational practice using GTDS ignores the ionospheric refraction effects for the spacecraft-to-spacecraft links of TDRSS measurements. It is therefore desirable to know for different types of spacecraft missions the magnitude and characteristics of the spacecraft orbit errors that can result by ignoring these effects. As the ionospheric refraction effect is functionally dependent upon various parameters that influence the electron density, a comprehensive study of this kind would require extensive investigations involving combinations of various mission scenarios and various parameters. This section presents the orbit determination error analysis results for the routine orbit determination scenario of a GRO-type spacecraft.

Because the primary interest of the study was the ionospheric refraction effects, the study was limited to investigating how ionospheric refraction of spacecraft-to-spacecraft links influences the orbit determination accuracy. A detailed orbit analysis investigation for the GRO spacecraft mission using ODEAS can be found in Reference 6.

2.2.1 Error Analysis Results

The GRO spacecraft is scheduled to be launched in late 1990. In this study, the nominal operational orbit of the GRO spacecraft was modeled as a near-circular orbit, with an orbital height of 350 kilometers and an inclination of 28.5 degrees. The orbital period for this orbit was 91.54 minutes. Routine orbit determination of the GRO spacecraft was simulated by processing 30-hour definitive data arcs in which the orbital state of the GRO spacecraft and the atmospheric drag scaling factor (ϱ_1) were simultaneously estimated together with the orbital elements of TDRS-East and TDRS-West. To see the effects of ionospheric refraction on the spacecraft orbit errors for different tracking geometries and tracking pass lengths, four tracking scenarios (T1 through T4) were simulated; the details of these scenarios are given in Table 4. All four tracking scenarios involved one TDRSS tracking pass per GRO revolution, alternating between TDRS-East and TDRS-West. For scenarios T1 and T2, tracking pass lengths were limited to 5 minutes each; whereas in scenarios T3 and T4, tracking pass lengths were limited to 20 minutes each. In scenarios T1 and T3, the tracking measurements are subject to relatively high ionospheric refraction effects; whereas in scenarios T2 and T4, the ionospheric refraction effects are relatively small.

Table 4. Simulated Tracking Scenarios for GRO Routine Orbit Determination

TRACKING* SCENARIO	RELAY SPACECRAFT	MINIMUM ELEVATION ANGLE (DEGREES)	MAXIMUM ELEVATION ANGLE (DEGREES)	AVERAGE ELEVATION ANGLE (DEGREES)	NUMBER OF MEASUREMENTS	IONOSPHERIC CORRECTIONS		
						AVERAGE RANGE (METERS)	RMS RANGE-RATE (MILLIMETERS/ SECOND)	AVERAGE RANGE-RATE (MILLIMETERS/ SECOND)
T1	TDRS-EAST	-12.9	7.53	-2.7	310 (RANGE) 310 (RANGE-RATE)	18.2	98.66	6.43
	TDRS-WEST	-13.3	10.5	-2.1	310 (RANGE) 310 (RANGE-RATE)	17.1	84.99	1.16
T2	TDRS-EAST	58.8	88.53	71.1	310 (RANGE) 310 (RANGE-RATE)	2.84	2.59	-0.94
	TDRS-WEST	43.5	84.6	63.8	310 (RANGE) 310 (RANGE-RATE)	4.14	7.15	0.99
T3	TDRS-EAST	-15.1	69.13	24.2	1210 (RANGE) 1210 (RANGE-RATE)	7.97	51.55	-0.07
	TDRS-WEST	-13.3	86.6	24.2	1210 (RANGE) 1210 (RANGE-RATE)	8.57	48.20	-0.79
T4	TDRS-EAST	33.2	88.53	60.4	1210 (RANGE) 1210 (RANGE-RATE)	3.05	3.95	-0.88
	TDRS-WEST	11.4	88.6	55.7	1210 (RANGE) 1210 (RANGE-RATE)	4.70	10.08	1.33

NOTES:
1. TDRSS TWO-WAY RANGE AND TWO-WAY RANGE RATE WITH TDRSS S-BAND SINGLE-ACCESS (SSA) SERVICE
2. 3σ RANGE NOISE = 1.5 METERS; WEIGHT = 90 METERS
   3σ RANGE-RATE NOISE = 0.004 METER/SECOND; WEIGHT = 0.1 METER/SECOND
3. DATA RATE = ONE RANGE AND ONE RANGE-RATE MEASUREMENT EVERY 10 SECONDS

* T1 AND T2 = 5-MINUTE TRACKING PASS PER GRO REVOLUTION, ALTERNATING BETWEEN TDRS-EAST AND TDRS-WEST
T3 AND T4 = 20-MINUTE TRACKING PASS PER GRO REVOLUTION, ALTERNATING BETWEEN TDRS-EAST AND TDRS-WEST

For each of the error analysis simulations, the ionospheric refraction corrections for TDRS-East-to-GRO and TDRS-West-to-GRO tracking links were included as considered error sources, with their respective 3σ errors set to be 100 percent of the spacecraft-to-spacecraft ionospheric refraction correction. Error analysis results obtained for the four tracking scenarios are summarized in Table 5. The contribution from the range and range-rate measurements are given separately for spacecraft-to-spacecraft tracking involving the TDRS-East and TDRS-West relay spacecraft. Total contributions from the range and range-rate measurements listed in the table are obtained by treating these quantities as fully correlated parameters. The maximum position errors listed are the maximum errors encountered during the 30-hour data arc for each position component error, which may not necessarily occur at the same time.

Error analysis results for scenarios T1 and T2 show that the GRO spacecraft position errors contributed by spacecraft-to-spacecraft ionospheric refraction can vary significantly, depending on the geometry of the tracking measurements included in the orbit estimation process. As shown in Table 4, the number of tracking passes and measurements included in T1 and T2 are the same, except that the measurements included in T1 have substantially lower elevation angles than those of T2 and are therefore subject to higher ionospheric disturbance than T2. Because of this, the GRO spacecraft position errors contributed by the spacecraft-to-spacecraft ionospheric refraction are found to be substantially higher in T1 than in T2, as shown in Table 5. In scenario T1, the maximum root-sum-square (RSS) position errors contributed by the TDRS-East and TDRS-West tracking measurements are 57.82 and 91.90 meters respectively; whereas in scenario T2, the same position error contributions are only 3.78 and 9.76 meters, respectively. The results suggest that by properly selecting the tracking measurements with favorable tracking geometries, the maximum RSS position errors can be reduced by as much as 50 to 80 meters. The same trend can be observed by comparison of scenarios T3 and T4.

Comparison of scenarios T1 with T3 and T2 with T4 also show the effects of tracking pass lengths on the spacecraft position errors. In the tracking scenarios studied, the spacecraft-to-spacecraft ionospheric refraction effects on the GRO spacecraft position errors are found to be generally smaller for longer tracking pass lengths. It should be noted that in all the simulated cases, the tracking scenarios associated with the longer tracking pass lengths also have better tracking geometries. Therefore, the effects observed in the above scenarios may be due to the combined effect of longer tracking pass lengths and better geometries.

To illustrate the relationship between the ionospheric refraction correction versus the resulting spacecraft position errors, Table 6 lists the ionospheric refraction corrections and the corresponding maximum RSS position errors contributed separately from the range and range-rate measurements. A general trend can be observed from this table indicating that an increase in the average value (for range measurement) or the RMS value (for range-rate measurement) of ionospheric refraction correction results in an increase in the maximum RSS position errors. However, this general trend does not hold very well between scenarios T4 and T2, where the magnitudes of corrections are small.

3.0 CONCLUSIONS AND RECOMMENDATIONS

The magnitudes of ionospheric refraction correction for spacecraft-to-spacecraft tracking links can vary depending upon the various parameters that affect the electron densities. The parameters studied include the local solar times of the measurements; solar activity; spacecraft inclination; and the tracking geometry, which is characterized by (1) the elevation angle of the TDRS with respect to the TDRS user spacecraft and (2) the orbital height. The following general conclusions can be made regarding the spacecraft-to-spacecraft ionospheric refraction:

- The ionospheric refraction corrections for both range and range-rate measurements increase with the decrease in the elevation angle of the TDRS with respect to the TDRS user spacecraft. This suggests that to avoid high ionospheric disturbance in tracking data, data taken at very low elevation angles should be excluded. For high spacecraft altitudes (950 kilometers and up), the magnitudes of ionospheric corrections are much smaller than those for the lower altitudes, but the overall trend of elevation angle dependence still exists.

- The ionospheric refraction correction is found to be a strong function of the local solar time of the TDRS user spacecraft at the time of measurement. It was found that measurements taken between 12 to

Table 5. GRO Spacecraft Position Errors Contributed by Spacecraft-to-Spacecraft Ionospheric Refraction Effects for Different Simulated Tracking Scenarios

TRACKING* SCENARIO	IONOSPHERIC REFRACTION ERROR SOURCE		POSITION ERROR AT EPOCH (METERS)				POSITION ERROR AT END OF DATA (METERS)				MAXIMUM POSITION ERROR COMPONENT (METERS)			
	SPACECRAFT	MEASURE-MENT TYPE	HEIGHT	CROSS-TRACK	ALONG-TRACK	RSS	HEIGHT	CROSS-TRACK	ALONG-TRACK	RSS	HEIGHT	CROSS-TRACK	ALONG-TRACK	RSS
T1	TDRS-EAST	RANGE	1.10	1.01	18.2	18.2	-0.863	-0.863	11.4	11.5	-1.10	1.27	18.19	18.25
		RANGE-RATE	-8.95	-26.6	-48.2	55.8	11.9	-41.6	-28.0	51.5	-12.94	-47.63	-53.28	60.09
		TOTAL	-7.86	-27.6	-30.0	41.6	11.5	-42.5	-16.6	47.0	12.93	-48.87	-49.69	57.82
	TDRS-WEST	RANGE	-1.74	-2.06	-14.4	14.6	1.493	0.461	-17.0	17.1	1.96	-2.08	17.00	17.07
		RANGE-RATE	12.5	-67.4	3.26	68.6	10.2	-9.11	75.6	76.8	20.03	-47.63	80.92	96.17
		TOTAL	10.7	-69.4	-11.2	71.1	11.7	-8.65	58.6	60.2	-19.63	-71.54	68.51	91.90
T2	TDRS-EAST	RANGE	0.189	-0.218	-0.097	0.304	-0.011	-1.01	0.441	1.10	0.198	1.06	0.443	1.147
		RANGE-RATE	0.495	2.01	0.011	2.07	0.22	-1.47	1.48	2.10	0.652	2.39	1.98	2.95
		TOTAL	0.685	1.79	-0.086	1.92	0.209	-2.48	1.92	3.14	0.826	-2.94	2.38	3.78
	TDRS-WEST	RANGE	0.410	-0.145	-0.165	0.465	-0.053	-1.50	1.16	1.90	0.419	-1.54	1.16	1.899
		RANGE-RATE	1.39	2.64	-5.80	6.52	-0.056	-2.93	4.87	5.69	-1.565	-3.77	-7.75	8.386
		TOTAL	1.80	2.49	-5.97	6.71	-0.109	-4.43	6.03	7.49	-1.971	-4.91	-8.57	9.76
T3	TDRS-EAST	RANGE	0.626	-0.626	3.61	3.72	0.119	-0.741	4.04	4.11	0.766	-1.036	4.993	5.008
		RANGE-RATE	-3.18	-0.057	-18.0	18.3	5.55	-15.9	-7.35	18.4	-5.681	16.058	-18.236	21.925
		TOTAL	-2.55	-0.683	-14.4	14.6	5.43	-16.64	-3.31	17.81	-5.72	-16.90	-14.66	20.13
	TDRS-WEST	RANGE	-0.448	-2.48	-4.46	5.12	0.930	-0.320	-3.01	3.17	0.930	-2.555	-4.456	5.118
		RANGE-RATE	0.660	-9.29	-4.97	10.6	4.136	-16.5	9.65	19.6	4.835	20.069	14.287	21.637
		TOTAL	0.212	-11.77	-9.43	15.08	5.06	-16.82	6.64	18.78	5.52	-3.86	13.27	23.18
T4	TDRS-EAST	RANGE	0.181	-0.191	-0.072	0.272	-0 .0302	-0.965	0.404	1.05	0.183	-1.014	0.404	1.084
		RANGE-RATE	-0.237	2.12	-0.536	2.20	0.278	-1.76	0.622	1.89	0.4497	2.648	1.063	2.795
		TOTAL	0.418	1.93	-0.607	2.07	0.248	-2.73	1.026	2.92	0.589	-3.212	1.327	3.48
	TDRS-WEST	RANGE	0.466	-0.0007	0.030	0.467	-0.133	-1.34	1.28	1.85	0.466	1.349	1.278	1.855
		RANGE-RATE	0.682	2.21	-2.90	3.71	0.103	-5.31	-0.418	5.32	-0.823	5.596	-3.704	6.623
		TOTAL	1.15	2.21	-2.87	3.80	-0.030	-6.65	0.862	6.71	-1.252	6.859	04.537	8.185

* T1 AND T2 = 5-MINUTE TRACKING PASS PER GRO REVOLUTION, ALTERNATING BETWEEN TDRS-EAST AND TDRS-WEST
T3 AND T4 = 20-MINUTE TRACKING PASS PER GRO REVOLUTION, ALTERNATING BETWEEN TDRS-EAST AND TDRS-WEST

Table 6. Relationship Between the Spacecraft-to-Spacecraft Ionospheric Refraction Corrections Versus the Corresponding GRO Spacecraft Position Errors

TRACKING* SCENARIO	RELAY SPACECRAFT	AVERAGE IONOSPHERIC CORRECTION RANGE (METERS)	MAXIMUM POSITION ERROR DUE TO RANGE (METERS)	RMS IONOSPHERIC CORRECTION RANGE RATE (MILLIMETERS/ SECOND)	MAXIMUM POSITION ERROR DUE TO RANGE RATE (METERS)
T1	TDRS-EAST	18.2	18.25	98.66	60.09
T3	TDRS-EAST	7.97	5.01	51.55	21.93
T4	TDRS-EAST	3.05	1.08	3.95	2.80
T2	TDRS-EAST	2.84	1.15	2.59	2.95
T1	TDRS-WEST	17.1	17.07	84.99	96.17
T3	TDRS-WEST	8.57	5.12	48.20	21.64
T4	TDRS-WEST	4.70	1.86	10.08	6.62
T2	TDRS-WEST	4.14	1.90	7.15	8.39

* T1 AND T2 = 5-MINUTE TRACKING PASS PER GRO REVOLUTION, ALTERNATING BETWEEN TDRS-EAST AND TDRS-WEST
T3 AND T4 = 20-MINUTE TRACKING PASS PER GRO REVOLUTION, ALTERNATING BETWEEN TDRS-EAST AND TDRS-WEST

16 hours of local solar time tend to have significantly higher ionospheric disturbance than those taken at different local solar times. If this phenomenon manifested by the Bent model closely matches the real ionospheric behavior, this may suggest a need to take into account the local time of the user spacecraft as a data editing criterion, together with the tracking geometry considerations, to effectively exclude the tracking data subject to high ionospheric disturbance. A data editing criterion based on tracking geometry, characterized by the height of the ray path and the central angle, is presently employed in the operational GTDS. To verify this local solar time effect, future studies should include assessing the orbit determination accuracy using real tracking data to determine whether more accurate spacecraft orbits can be realized by using a data editing criterion that excludes those tracking data taken at a user spacecraft local solar time around 15 hours and at very low elevation angles.

- The ionospheric refraction correction is found to be a strong function of solar activity. For a spacecraft orbit at a 350-kilometer altitude and 28.5-degree inclination, the ionospheric refraction correction values obtained at a moderately high solar activity [average solar flux = 200.6, daily solar flux = 218.9 (units = 10^{-22} watts/ meters^2/hertz)] can be twice as large as the corrections obtained at a relatively lower solar activity [average solar flux = 111.4, daily solar flux = 159.2 (units = 10^{-22} watts/ meters^2/hertz)]. It was also found that the ionospheric refraction effect is more sensitive to the 12-month average solar flux level than the daily solar flux level.

- The maximum electron density value, N_m, as defined on the Bent electron density profile, has a strong latitude dependence. It was found that N_m values tend to be higher in low latitude (near-equatorial) regions than in the high-latitude (polar) regions. It is therefore conceivable that in a particular spacecraft orbit, the ionospheric refraction effects may tend to be high if many of the tracking measurements were taken when the user spacecraft was near the equatorial region. This may explain why the ionospheric refraction corrections observed in the study tend to be generally higher for the low-inclination (28.5-degree) orbits than those of the high-inclination (99.03-degree) orbits.

From the error analysis results of the GRO spacecraft, the following conclusions can be drawn:

- The spacecraft position errors contributed by spacecraft-to-spacecraft ionospheric refraction can vary significantly, depending on the geometry of the tracking measurements included in the orbit estimation

process. In a routine orbit determination involving 30 hours of data with 5 minutes of TDRSS tracking per GRO revolution, selecting the tracking measurements with favorable tracking geometries can reduce the spacecraft maximum RSS position errors from 57.82 meters to 3.78 meters (in the case of TDRS-East-to-GRO ionospheric refraction contribution) and from 91.90 meters to 9.76 meters (in the case of TDRS-West-to-GRO ionospheric refraction contribution).

- An increase in the mean value of the ionospheric refraction correction is found to be associated with an increase in the maximum RSS position errors. However, this general trend does not hold very well where the magnitudes of the ionospheric refraction corrections are small or the geometric quality of the processed data is appreciably altered. This general relationship also implies that parameters affecting the mean ionospheric refraction values, such as solar activity, will also affect the corresponding spacecraft orbit errors in a similar fashion.

ACKNOWLEDGMENT

The authors wish to thank A. Schanzle (Computer Sciences Corporation) for his thoughtful comments and M. Radomski (Computer Sciences Corporation) for providing reference materials and some useful discussions during preparation of this paper.

REFERENCES

1. Goddard Space Flight Center, Flight Dynamics Division, FDD/552-89/001, *Goddard Trajectory Determination System (GTDS) Mathematical Theory, Revision 1*, A. C. Long, J. O. Cappellari, Jr., C. E. Velez, and A. J. Fuchs (editors), July 1989

2. Bent, R. B., Llewellyn, S. K., Nesterczuk, G., Schmid, P. E., "The Development of a Highly Successful Worldwide Empirical Ionospheric Model and Its Use in Certain Aspects of Space Communication and Worldwide Total Electron Content Investigations," *Proceedings of the Ionospheric Effects Symposium*, Arlington, Virginia, January 20-22, 1975

3. Computer Sciences Corporation, CSC/TM-89/6037, *TDRSS User Orbit Determination Error Due to Ionospheric Refraction*, M. Radomski and A. Binebrink, September 1989

4. Goddard Space Flight Center, X-591-73-281, *NASA-GSFC Ionospheric Corrections to Satellite Tracking Data*, P. E. Schmid, R. B. Bent, S. K. Llewellyn, G. Nesterczuk, and S. R. Rangaswamy, December 1973

5. Goddard Space Flight Center, Flight Dynamics Division, FDD/554-90/029, *Orbit Determination Error Analysis System (ODEAS) Mathematical Specifications, Revision 1*, C. P. Yee and T. Lee, January 1990

6. Goddard Space Flight Center, Flight Dynamics Division, FDD/554-90/013, *Gamma Ray Observatory (GRO), 1990 Flight Dynamics Analysis Report 6, Orbit Determination Analysis*, J. Dunham, February 1990

A COMPARISON OF OPERATIONALLY DETERMINED ATMOSPHERIC DENSITIES FROM SATELLITE ORBIT SOLUTIONS AND THE EXOSPHERIC TEMPERATURE FROM THE JACCHIA-ROBERTS MODEL*

D. T. Ward, E. A. Smith, and M. C. Phenneger
Computer Sciences Corporation (CSC)

ABSTRACT

Operational orbit determination by the Flight Dynamics Division at the Goddard Space Flight Center has yielded a data base of orbit solutions covering the onset of solar cycle 22. Solutions for nine satellites include an estimated drag adjustment parameter (ϱ_1) determined by the Goddard Trajectory Determination System (GTDS). The ϱ_1 is used to evaluate correlations between density variations and changes in the following: 10.7-centimeter wavelength solar flux ($F_{10.7}$), the geomagnetic index A_p, and two exospheric temperatures (T_c and T_∞) adapted from the Jacchia-Roberts atmospheric density model in GTDS. T_c depends on the daily and 81-day centered mean $F_{10.7}$; T_∞ depends on T_c and the geomagnetic index K_p values. The highest correlations are between density and T_∞. Correlations with T_c and $F_{10.7}$ are lower by 9 and 10 percent, respectively. For most cases, correlations with A_p are considerably lower; however, significant correlations with A_p were found for some high-inclination, moderate-altitude orbits.

Results from this analysis enhance the understanding of the drag model and the accommodation of atmospheric density variations in the operational orbit determination support. The degree of correlation demonstrates the sensitivity of the orbit determination process to drag variations and to the input parameters that characterize aspects of the atmospheric density model. To this extent, the degree of correlation provides a measure of performance for methods of selecting or modeling the thermospheric densities using the solar $F_{10.7}$ and geomagnetic data as input to the process.

*This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, Contract NAS 5-31500.

1. INTRODUCTION

Operational orbit determination by the Flight Dynamics Division (FDD) of the Goddard Space Flight Center (GSFC) using the Goddard Trajectory Determination System (GTDS) provides routine estimates of atmospheric density. This paper uses the atmospheric drag data derived from the estimated atmospheric density scaling parameters for the following nine satellites: Dynamics Explorer (DE)-1, Earth Radiation Budget Satellite (ERBS), Landsat-4 and -5, Nimbus-7, National Oceanic and Atmospheric Administration (NOAA)-9 and -10, Solar Maximum Mission (SMM), and Solar Mesosphere Explorer (SME). The study period extends from January 1, 1988, to August 31, 1989, and covers the onset of solar cycle 22. Each orbit solution is a seven-parameter orbital state vector consisting of position and velocity vectors and an atmospheric density scaling parameter, ϱ_1. ϱ_1 is estimated to accommodate differences between modeled and actual atmospheric density and corresponding drag effects. Data for three of the nine satellites cover the early study period. Data for the other satellites became available when drag forces became large enough to provide reliable estimates of ϱ_1.

The combined data bases of ϱ_1 solutions are analyzed to evaluate correlations between density variations and changes in the following parameters, which are associated with atmospheric density modeling: 10.7-centimeter (2800 megahertz (MHz)) wavelength solar flux ($F_{10.7}$), geomagnetic index A_p, and two exospheric temperatures adapted from the Jacchia-Roberts (JR) 1971 atmospheric density model as used in GTDS (References 1 and 2). The first exospheric temperature, T_c, includes the contributions from daily $F_{10.7}$ and a centered 81-day mean, $\bar{F}_{10.7}$. The second adjusted exospheric temperature, T_∞, includes the contribution from the 3-hourly geomagnetic index K_p. Earlier work included only $F_{10.7}$ and A_p correlations with density data estimates for ERBS and SMM (Reference 3).

When data from several satellite orbits are compared, the effect of orbital geometry, altitude, and orbit solution accuracy on density correlations with solar-geomagnetic activity can be assessed. In particular, the results offer a method of assessment of the merits of using the Harris-Priester (HP) (References 4 through 6) or the JR atmospheric density model options available in GTDS without requiring extensive reprocessing of the orbit solutions using the JR model. Operational requirements to implement the JR model are now being assessed.

This paper is organized as follows. Section 2 discusses solar and geomagnetic activity, the HP and JR density models, and methods of density determination. Section 3 provides the results for each satellite, including an overview of the satellite orbit and satellite-specific operational orbit support. Section 4 provides conclusions and recommendations for further study. *NOTE*: Figures are included at the end of the text pages.

2. BACKGROUND AND PROCEDURES

The equivalent atmospheric densities are derived from the estimated ϱ_1. The absolute accuracy of this density is subject to errors in the spacecraft ballistic coefficient, but the variations in these densities relative to each other are expected to be accurate. The analysis is sensitive only to variations in the average ballistic coefficient over time intervals of 1 to 5 days, which are expected to be small.

ϱ_1 variations are caused by variations in atmospheric density driven by solar and geomagnetic activity in addition to tracking errors, geopotential modeling errors, spacecraft drag coefficient (C_D) variations, and effective cross-sectional area (A) errors. The degree to which the density variations are not correlated with the solar and geomagnetic effects is a measure of the influence of these additional factors.

A brief explanation of solar and geomagnetic parameters and their relationship to atmospheric density for the HP and JR models follows.

2.1 SOLAR-GEOMAGNETIC PHENOMENA

$F_{10.7}$ is a daily measure of solar activity and is used as an indicator of the intensity of extreme ultraviolet (EUV) radiation of the Sun, which heats the thermosphere. The degree of the correlation of $F_{10.7}$ to

EUV flux is high, except for the minimum phase of the 11-year cycle (Reference 7). The last solar minimum occurred in September 1986, when the monthly mean $F_{10.7}$ was 68.7 x 10^{-22} W m^{-2} Hz^{-1}. The $F_{10.7}$ values for this study were obtained from the National Geophysical Data Center of the Solar-Terrestrial Physics Division in Boulder, Colorado.

The 3-hourly geomagnetic index K_p is a quasi-logarithmic measurement of the geomagnetic field activity at geomagnetic latitude 50 degrees (deg) (Reference 8). The K_p values used in this study were obtained from the International Service of Geomagnetic Indices at the Institut fuer Geophysik, Goettingen, Federal Republic of Germany.

The "daily equivalent planetary amplitude," A_p, is derived by converting K_p values to a linear index and averaging over one day (Reference 9). The A_p values used in this study were provided by World Data Center A for Solar-Terrestrial Physics, National Oceanic and Atmospheric Administration E/GC2, 325 Broadway, Boulder, Colorado 80303.

The daily $F_{10.7}$ is characterized by variations with a period of the 27-day solar rotation and is referred to in the literature as the rotational component of the solar flux. The centered 81-day average, $\bar{F}_{10.7}$, is associated with clear-disk solar radiance (Reference 10). The geomagnetic indices are characterized by short, intense bursts at various intervals. A_p, $F_{10.7}$, and $\bar{F}_{10.7}$ for the 20-month period under study are shown in Figure 1.

2.2 THE HARRIS-PRIESTER DENSITY MODEL

The modified HP atmospheric density model is a set of 10 tables of atmospheric density versus altitude corresponding to ten $F_{10.7}$ levels. An atmospheric density scale factor (ϱ_1) is estimated in the differential correction (DC) solution to accommodate variations relative to the modeled drag using a selected HP table.

The acceleration due to drag, a_D, at a point in time, is given by

$$a_D = \frac{-C_D A v^2 \varrho_0 (1+\varrho_1)}{2m}$$

where m is the spacecraft mass, v is the spacecraft velocity, and ϱ_0 is the density computed from the HP tables. ϱ_0 is dependent upon the altitude, z, and the $F_{10.7}$ HP level, F_i. ϱ_0 is given by

$$\varrho_0 (z, F_i, \phi, n) = \varrho_{min} (z, F_i) + (\varrho_{max} (z, F_i) - \varrho_{min}(z, F_i)) \cos^n \left(\frac{\phi}{2}\right)$$

where ϕ is the angle between the spacecraft position and the apex of the bulge, n is an adjustable exponent, ϱ_{min} is the minimum density, and ϱ_{max} is the maximum density (References 11 and 12). Conventionally, n = 2 is used for equatorial orbits, and n = 6 is used for polar orbits. The average ϱ_0 is calculated by integrating over all ϕ.

$$\varrho_0 (z, F_i) = \frac{1}{\pi} \int_0^{\pi} \varrho_0 (z, F_i, \phi, n)\, d\phi$$

Using this ϱ_0 and the ϱ_1 estimated in the orbit solution, an equivalent density, $\varrho (z_J)$, can be determined by the following equation for each satellite orbit solution.

$$\varrho_0 (z_J) = \varrho_0 (z_J, F_i) [1 + \varrho_1 (F_i)]$$

The equivalent densities are calculated at a single HP tabulated altitude, z_J, for each spacecraft that has small altitude variations.

The variation in altitude for each spacecraft over the study period was much less than the atmospheric scale height, except for SMM and DE-1. To isolate changes in density that were not due to altitude, DE-1 and SMM densities were converted back to ϱ_1 (F_i), an equivalent ϱ_1, for HP tables 7 and 4, respectively. The following equation was used for SMM:

$$\varrho_1 (F_4) = \frac{\varrho_0 (z, F_i) [1 + \varrho_1 (F_i)]}{\varrho_0 (z, F_4)} - 1,$$

where ϱ_0 (z, F_i) is determined using exponential scale heights from HP tables.

2.3 JACCHIA-ROBERTS DENSITY MODEL AND EXOSPHERIC TEMPERATURES

The Jacchia-Roberts 1971 (JR71) atmospheric density model option in GTDS gives the density at a given altitude from the $F_{10.7}$, $\overline{F}_{10.7}$, and K_p input data. In the JR71 model, the nighttime minimum global exospheric temperature for zero geomagnetic activity (T_c) is computed from the daily $F_{10.7}$ and the 81-day mean $\overline{F}_{10.7}$, as follows (Reference 12):

$$T_c = 379° + 3.24° \overline{F}_{10.7} + 1.3° (F_{10.7} - \overline{F}_{10.7})$$

There is an approximate 1-day lag (Reference 13) between solar flux change and a resulting change in exospheric temperature.

For simplicity in this work, a geometrical factor, which varies T_c with spacecraft geodetic latitude and the solar declination, was not accounted for in determining exospheric temperature T_∞. The variation in the amplitude of this factor over a season is 30 percent of T_c, but it is averaged over many orbits and is not tested in the correlations that were much shorter than the seasonal variation.

The correction to the exospheric temperature for geomagnetic activity is

$$\Delta T_\infty = 28°.0 K_p + 0°.03 \exp (K_p)$$

The corrected exospheric temperature used here is

$$T_\infty = T_c + \Delta T_\infty$$

For a 6.7-hour lag in the response to K_p (Reference 12), T_c and T_∞ are illustrated in Figure 2.

3. DENSITY DATA AND RESULTS

Tables 1a and 1b provide background information on the satellites studied. Table 1a provides the average data arc length and the orbital parameters: altitude, eccentricity, and inclination. Table 1b provides the GTDS modeling characteristics: C_D, A, mass, diurnal bulge model, and the maximum order and degree of the Goddard Earth Model (GEM)-9 geopotential model matrix.

The standard solution tracking data arc length varies by several hours due to the granularity of the tracking schedule. The $F_{10.7}$, A_p, T_c, and T_∞ data were averaged for each satellite over the average tracking

Table 1a. Average Data Arc Lengths and Orbital Parameters

Spacecraft	Average Data Arc Length	Altitude (km)	Eccentricity	Inclination (deg)
DE-1	3 days, 4.40 hours	470 to 23,350	0.62	89.2
ERBS	5 days, 6.62 hours 4 days, 6.62 hours	600	0.0005	57.
Landsat-4	1 day, 3.09 hours	700	0.0003	98.2
Landsat-5	1 day, 2.54 hours	700	0.0003	98.2
Nimbus-7	3 days, 8.79 hours	950	0.0009	99.2
NOAA-9	7 days, 0.00 hour	850	0.0015	99.1
NOAA-10	7 days, 0.00 hour	813	0.0014	98.6
SMM	2 days, 7 hours, 1 day, 7 hours	485 to 390	0.0002	28.5
SME	2 days, 14.54 hours	503 to 482	0.0002	97.7

5841G(4)-2a

Table 1b. GTDS Modeling Characteristics

Spacecraft	C_D	A (m^2)	Mass (kg)	Bulge	GEM-9
DE-1	2.3	3.05	402.79005	$\cos^6$	16 x 16
ERBS	2.2	4.7	2116.	$\cos^2$	8 x 8
Landsat-4	2.2	12.2644	1932.2669	$\cos^6$	21 x 21
Landsat-5	2.2	12.664	1943.538	$\cos^6$	21 x 21
Nimbus-7	2.1	9.5597	938.03	$\cos^6$	8 x 8
NOAA-9	2.3	10.79	1029.3	$\cos^6$	8 x 8
NOAA-10	2.3	10.79	1029.3	$\cos^6$	8 x 8
SMM	2.2	17.5	2315.59	$\cos^2$	16 x 16
SME	2.3	1.129	415.5	$\cos^6$	21 x 21

5841G(4)-2b

arc to be consistent with the density parameters estimated for the tracking arc. Correlation analysis results for each spacecraft are presented in the following sections. The correlation coefficients (r) from linear regressions are summarized in Table 2.

3.1 DYNAMICS EXPLORER-1

DE-1 is unique among the satellites studied because of its high eccentricity orbit. The DE-1 orbit is affected by atmospheric drag primarily near perigee. Operational estimation of ϱ_1 began on November 12, 1988. The equivalent ϱ_1 values were calculated using the HP table 7 and the DE-1 altitude perigee. These values are plotted versus epoch in Figure 3a. The correlations of these data with

Table 2. Correlation Coefficients Analysis

Spacecraft	No. of Epochs	Density Correlation Coefficients			
		$F_{10.7}$	T_c	T_∞	A_p
DE-1	148	0.36	0.30	0.50	0.58
ERBS	172	0.87	0.88	0.90	0.41
Landsat-4	351	0.48	0.50	0.67	0.68
Landsat-5	344	0.51	0.55	0.70	0.68
Nimbus-7	132	0.47	0.39	0.57	0.72
NOAA-9	65	0.76	0.79	0.83	0.52
NOAA-10	65	0.75	0.80	0.84	0.55
SMM	412	0.83	0.85	0.88	0.40
SME	234	0.83	0.86	0.90	0.52

5841G(4)-1

$F_{10.7}$, T_c, T_∞, and A_p are shown in Figures 3b, 3c, 3d, and 3e, respectively. The highest equivalent ϱ_1 corresponds to an extreme geomagnetic storm on March 13, 1989. A large increase in T_∞ and T_c occurred in early December 1988 due to rapidly rising $F_{10.7}$, evidenced by clumping or gaps in ϱ_1 versus temperature in Figures 3c and 3d. The low correlation of ϱ_1 with solar and geophysical parameters demonstrates that the measurability of density is small for DE-1. At each perigee passage the satellite samples many different altitudes. During the study period the perigee height decreased from 570 to 470 km while the perifocal latitude decreased from 80-deg N to 20-deg N. Recent operational work has shown that a more consistent ϱ_1 is estimated using a 21 x 21 GEM-9 model. An improvement of 0.05 in the correlation coefficients for the $F_{10.7}$, T_c, and T_∞ relations occurred when an equivalent ϱ_1 was used in place of an equivalent density. The two densities are compared in Figure 3f.

3.2 EARTH RADIATION BUDGET SATELLITE

Operational estimates of ϱ_1 for ERBS have been made since launch in 1984. Data from launch through October 31, 1987, were analyzed in Reference 3. The equivalent densities for ERBS were calculated with the HP altitude of 600 km. These are plotted versus epoch in Figure 4a. The correlations of these data with $F_{10.7}$, T_c, T_∞, and A_p are illustrated in Figures 4b, 4c, 4d, and 4e, respectively. Except for r_{A_p}, these correlations are high, with $r_{T_\infty} = 0.90$. The low A_p correlation corresponds to the averaging reduction of those short-term phenomena by orbit estimation over long tracking data arcs. Corresponding Jacchia 1977 (J77) density-temperature models (Reference 14), with a nonlinear curvature similar to that in Figure 4d, have a linear correlation coefficient of 0.94. Figure 4d includes the J77 densities corresponding to the average exospheric temperatures over the ERBS solution arcs. A polynomial fit was used to determine the density-temperature relationship in the J77 model at ERBS's altitude. After converting each average temperature to a J77 density, the J77 densities were correlated with the equivalent densities. A zero-intercept linear regression of the two densities yielded a scale factor of 2.4 for the J77 densities to best match the equivalent densities. The scaled J77 density curve is plotted in Figure 4d.

Previous work for low levels of solar activity (October 1984 to October 1987) found $r_{F_{10.7}} = 0.24$ and $r_{Ap} = 0.56$, using values averaged over the solution days, without time lags. Current $F_{10.7}$ correlations are significantly higher. This is not surprising, since the previous work included data only during the solar minimum. The greatly increased density levels allow a far more reliable density estimation.

3.3 LANDSAT-4 AND -5

Operational estimates of ϱ_1 for Landsat-4 began on July 1, 1988. The equivalent densities were calculated using the HP altitude of 700 km. These are plotted versus epoch in Figure 5a. For the first few months, ϱ_1 determinations were made only near solar flux peaks. The correlations of density with $F_{10.7}$, T_c, T_∞, and A_p are illustrated in Figures 5b, 5c, 5d, and 5e. Figure 5d includes a J77 density curve, described above, with a scale factor of 4.70. The correlations of density with $F_{10.7}$, T_c, and T_∞ are lower for Landsat-4, with r_{T_∞} the highest at 0.67. The A_p correlation is 0.68, which reflects the high inclination orbit sensitivity that is enhanced by the shorter solution arc. Note the improvement in correlation with T_∞ for the highest density point, corresponding to the March 13, 1989, geomagnetic storm.

It has been demonstrated that errors in the geopotential model affect the determination of ϱ_1 for Landsat-4 (Reference 15). These may contribute to the low Landsat correlation coefficients. An attempt to remove these errors for Landsat-4 was made by averaging densities and corresponding solar-geophysical parameter values over a 5-day interval, which is near two beat periods of the Landsat orbital resonance with 15th-order geopotential harmonics. No change in correlation coefficients was observed.

The density and correlations plots for Landsat-5 are not pictured. They are nearly identical to Landsat-4. Landsat-5 equivalent densities are plotted versus Landsat-4 equivalent densities in Figure 5f.

3.4 NIMBUS-7

Operational estimates of ϱ_1 for Nimbus-7 began in October 1988. The equivalent density data for Nimbus-7 were calculated using the HP altitude of 950 km. These are plotted against epoch in Figure 6a. The correlations of these densities with $F_{10.7}$, T_c, T_∞, and A_p are illustrated in Figures 6b, 6c, 6d, and 6e, respectively. The correlations are low for Nimbus-7, with r_{A_p} the highest at 0.72. This highest r_{A_p} corresponds to the highest altitude, and high inclination. The low value of r_{T_∞} may be due to the truncation of the geopotential model but also indicates that density is less measurable at that altitude.

3.5 NOAA-9 AND -10

Operational estimates of ϱ_1 for NOAA-9 and -10 have been performed since June 1988. The equivalent densities for NOAA-9 and -10 were calculated using HP altitudes of 850 and 800 km, respectively. These are plotted versus epoch in Figures 7a. The correlations of these densities with $F_{10.7}$, T_c, T_∞, and A_p are illustrated in Figures 7b, 7c, 7d, and 7e. The correlations of density to $F_{10.7}$, T_c, and T_∞ for NOAA-9 and -10 are high, with r_{T_∞} = 0.83 and 0.84, respectively. This correlation is more pronounced than that for Landsat-4 and Landsat-5 described in a previous section. Landsat-4 and Landsat-5 are at lower altitudes and use shorter tracking data arcs. This result demonstrates that drag perturbations are better measured using long tracking data arcs. Lower r_{A_p} for these spacecraft are a result of an increased tracking arc length. NOAA-9 density is plotted versus NOAA-10 density in Figure 7f.

3.6 SOLAR MAXIMUM MISSION

SMM was launched February 14, 1980, and reentered on December 2, 1989. Daily orbit solutions were performed after January 27, 1989. Before that, orbit solutions were performed every other day. The equivalent ϱ_1 relative to HP table 4 (to remove altitude-dependent variations) are illustrated in Figure 8a. The correlations of ϱ_1 with $F_{10.7}$, T_c, T_∞, and A_p are illustrated in Figures 8b, 8c, 8d, and 8e, respectively. The correlations of density to $F_{10.7}$, T_c, and T_∞ are high for SMM, with r_{T_∞} = 0.88 being the highest. Thus, the acceleration due to drag is dominating the solutions, as expected. This lowest r_{A_p} corresponds to the lowest altitude and inclination of the group. Previous results from launch to October 1987 were comparable to current results, with $r_{F_{10.7}}$ = 0.86 and r_{A_p} = 0.25 (Reference 3).

3.7 SOLAR MESOSPHERE EXPLORER

Operational estimates of ϱ_1 were performed for SME until contact with the spacecraft was lost on April 14, 1989. The equivalent densities for SME were calculated using the HP altitude of 500 km.

These are plotted versus epoch in Figure 9a. The correlations of density to $F_{10.7}$, T_c, T_∞, and A_p are shown in Figures 9b, 9c, 9d, and 9e, respectively. The correlations of the data with $F_{10.7}$, T_c, and T_∞ are high for SME, with $r_{T_\infty} = 0.90$ being the highest. SME's low altitude and polar orbit thus cause more sensitivity to solar and geomagnetic activity, respectively. The low r_{A_p} corresponds to low altitude and a moderate tracking data arc length.

4. CONCLUSIONS AND RECOMMENDATIONS

The operational data base of estimated ϱ_1 values for nearly 2 years of orbit solutions has been presented. Correlations of estimated atmospheric density with solar and geomagnetic activity measurements have been evaluated. The highest correlations were with the exospheric temperature T_∞, adapted from the JR71 model. T_∞ includes a dependence on the daily $F_{10.7}$, 81-day mean $\overline{F}_{10.7}$, and on 3-hourly K_p values. Lower altitude spacecraft (ERBS, SME, and SMM) densities correlate best with T_∞, T_c, and $F_{10.7}$, due to the larger drag at lower altitudes. At higher altitudes, longer arc solutions had correlations nearly as high as the low-altitude solutions. The correlation of density with T_∞ in each case is discernibly higher than with T_c or $F_{10.7}$, showing that the K_p correction improves the correlations more than does the $\overline{F}_{10.7}$ component.

A_p affects estimated density more at high inclinations and high altitudes than it does at low inclinations and low altitudes. Shorter arcs are more sensitive to A_p. The degree of correlation with A_p is highest during severe geomagnetic storm conditions. A_p, which accounts for the largest deviations in the T_c and the $F_{10.7}$ correlations, generally has a low correlation with density, except for Nimbus-7 and the Landsats.

Lack of higher correlations for the Landsats may be a result of short arc lengths or orbit modeling errors due to resonance with geopotential harmonic coefficients of the 15th order and degree. The low correlations for DE-1 arise because its orbit solution is the least affected by atmospheric drag.

Jacchia-Roberts modeling of exospheric temperature T_∞ is a better monitor of variations in atmospheric density than T_c, $F_{10.7}$, or A_p. ϱ_1 estimates using the HP model reflect the additional physical processes that are part of the JR model but are not included in the HP model. The correlations of density with solar-geomagnetic activity shown in this paper can be used to improve the use of the HP model in the predictive mode.

This paper extends the work begun in Reference 3. A large amount of data for SME from launch in 1981 to January 1988 has not been used. Other data during the previous solar maximum also exists for Landsat-4 and Nimbus-7. ϱ_1 estimations for most satellites were suspended during the solar minimum. Data continue to be accumulated during solar cycle 22. We recommend that these data be included in similar future analysis.

The estimation of density from satellite data is a valuable extension of operational orbit determination. The correlations evaluated here provide only a survey of the information available. Higher resolution studies concentrating on measurements of the atmospheric response time to various solar-geophysical stimuli are possible. Archived tracking data can be used to obtain density estimates over shorter or longer time intervals.

ACKNOWLEDGMENTS

The authors wish to thank Dana Peters (CSC) for her valuable contribution to the generation of a portion of these data. We also wish to thank Osvaldo Cuevas (NASA/GSFC) and Avinash Beri (CSC) for their thoughtful reviews of the draft.

REFERENCES

1. L. G. Jacchia, *Revised Static Models of the Thermosphere and Exosphere with Empirical Temperature Profiles*, Smithsonian Astrophysical Observatory Special Report No. 332, Cambridge, Massachusetts, May 1971

2. E. R. Roberts, Jr., "An Analytic Model for Upper Atmosphere Densities Based Upon Jacchia's 1970 Models," *Celestial Mechanics*, vol. 4, December 1971, pp. 368-377

3. E. A. Smith, "Correlations Between Solar Activity and Operationally Determined Satellite Drag Variation Parameters," Paper No. 27, presented at the GSFC Flight Mechanics/Estimation Theory Symposium, Greenbelt, Maryland, May 1988

4. I. Harris and W. Priester, "Time Dependent Structure of the Upper Atmosphere," *Journal of the Atmospheric Sciences*, vol. 19, July 1962, pp. 286-301 (also NASA-TN-D-1443)

5. I. Harris and W. Priester, *Theoretical Models for the Solar Cycle Variation of the Upper Atmosphere*, Goddard Space Flight Center Report NASA-TN-D-1444, August 1962

6. I. Harris and W. Priester, "Atmospheric Structure and Its Variations in the Region From 120 to 80 km," *COSPAR International Reference Atmosphere (CIRA) 1965, Space Research IV*, North Holland Publishing Company, Amsterdam, 1965

7. W. K. Tobiska, C. A. Barth, and R. D. Culp, "Use of Solar Lyman-α and 1-8 Å X-Rays as EUV Flux Indices," Paper No. 88-4293, presented at AIAA/AAS Astrodynamics Conference, Minneapolis, Minnesota, August 1988

8. Solar-Geophysical Data, 515 Supplement, July 1987, pp. 30-31, U.S. Department of Commerce (Boulder, Colorado 80303)

9. G. Rostoker, "Geomagnetic Indices," *Reviews of Geophysics and Space Physics*, vol. 10, November 1972, pp. 935-950

10. L. G. Jacchia and J. W. Slowey, "A Study of the Variations in the Thermosphere Related to Solar Activity," *Space Research XIII*, Akademie-Verlag, Berlin, 1973, pp. 343-348

11. E. A. Smith, D. T. Ward, M. W. Schmitt, M. C. Phenneger, F. J. Vaughn, and M. L. Lupisella, "Lifetime Predictions for the Solar Maximum Mission (SMM) and San Marco Spacecraft," Paper No. 27, presented at the GSFC Flight Mechanics/Estimation Theory Symposium, Greenbelt, Maryland, May 1989

12. Computer Sciences Corporation, *Goddard Trajectory Determination System Mathematical Theory (Revision 1)*, FDD/552-89/0001 and CSC/TR-89/6001, eds. A. C. Long, J. O. Cappellari, Jr., C. E. Velez, and A. J. Fuchs, July 1989

13. L. G. Jacchia, J. W. Slowey, and I. G. Campbell, "Analysis of Solar Activity Effects on the Upper Atmosphere," *Planetary and Space Science*, vol. 21, November 1973, pp. 1835-1842

14. L. G. Jacchia, *Thermospheric Temperature, Density, and Composition: New Models*, Smithsonian Astrophysical Observatory Special Report No. 375, Cambridge, Massachusetts, May 1977

15. S. L. Hoge, D. O. Casteel, M. C. Phenneger, and E. A. Smith, "The Effects of Geopotential Resonance on Orbit Determination for Landsat-4," Paper No. 88-4296, presented at AIAA/AAS Astrodynamics Conference, Minneapolis, Minnesota, August 1988

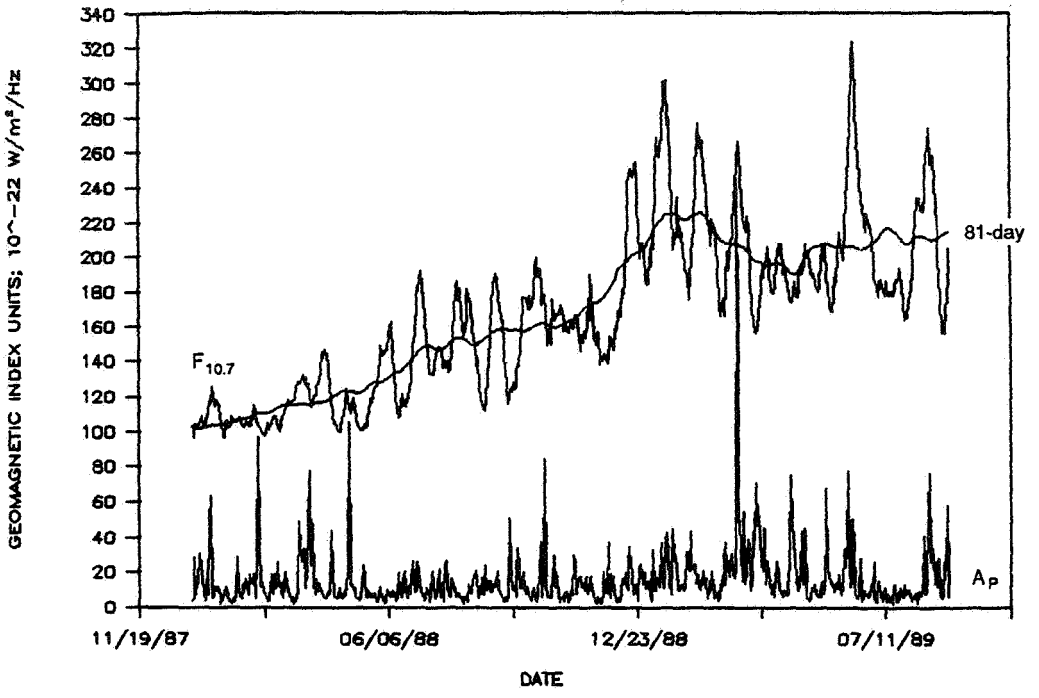


Figure 1. Daily Observed $F_{10.7}$, 81-Day Mean $\overline{F}_{10.7}$, and A_p

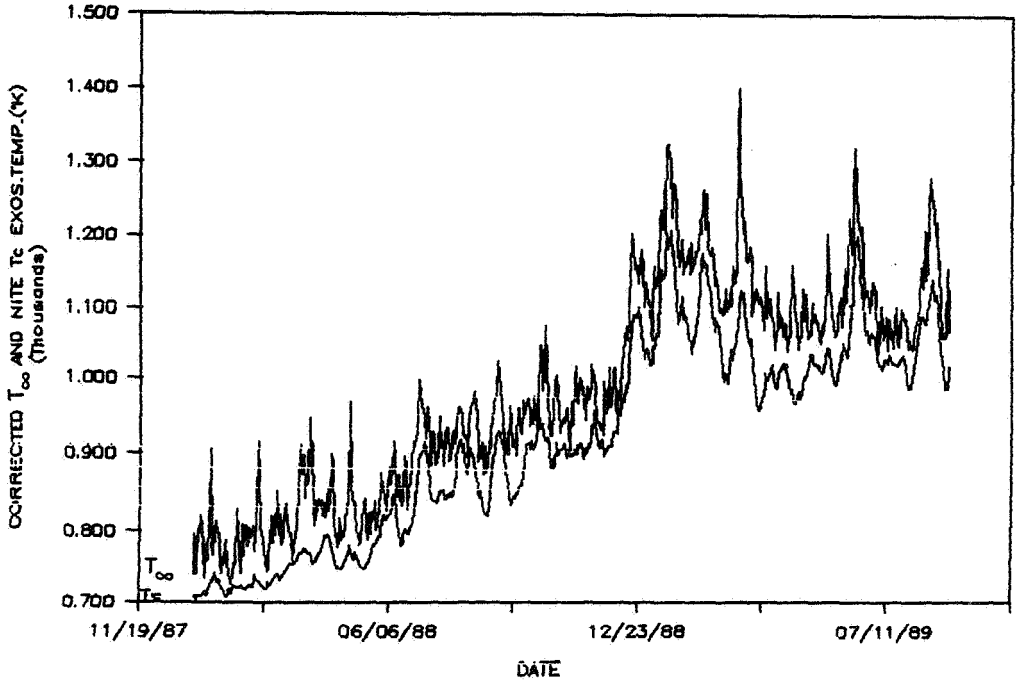


Figure 2. Daily Exospheric Temperatures, T_∞ and T_c

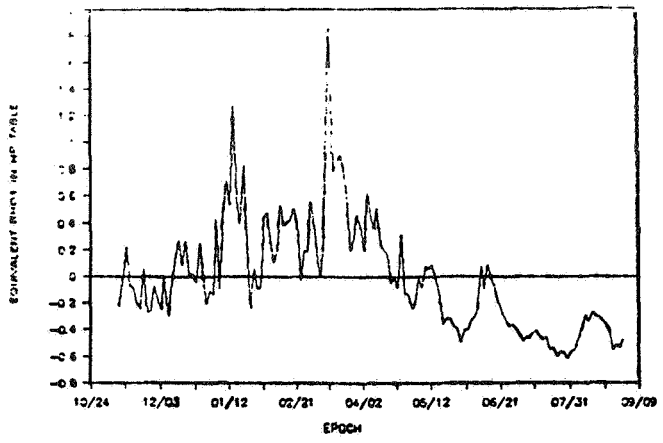


Figure 3a. DE-1 Equivalent ϱ_1 Values Versus Epoch

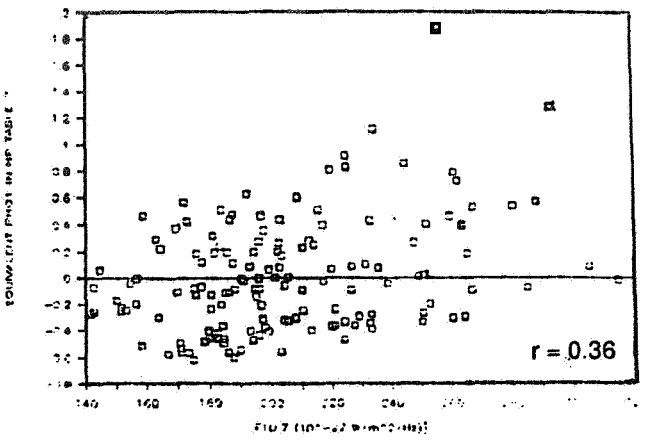


Figure 3b. DE-1 Equivalent ϱ_1 Values Versus $F_{10.7}$

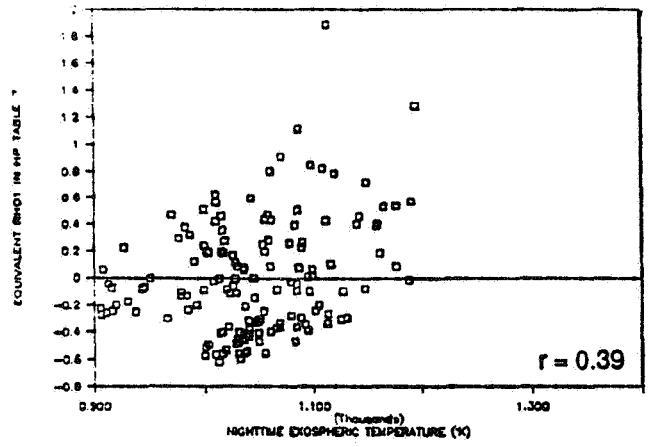


Figure 3c. DE-1 Equivalent ϱ_1 Values Versus T_c

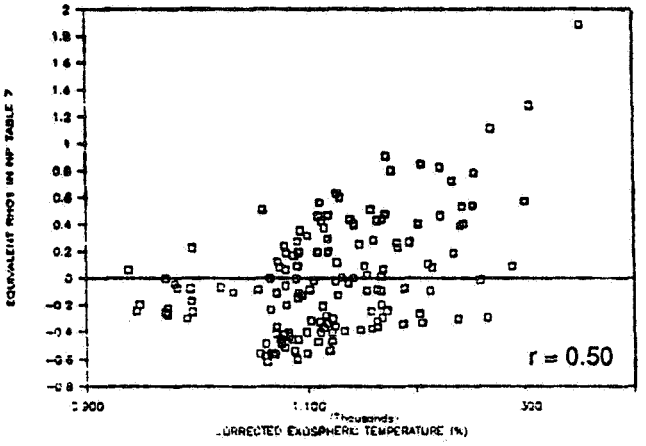


Figure 3d. DE-1 Equivalent ϱ_1 Values Versus T_∞

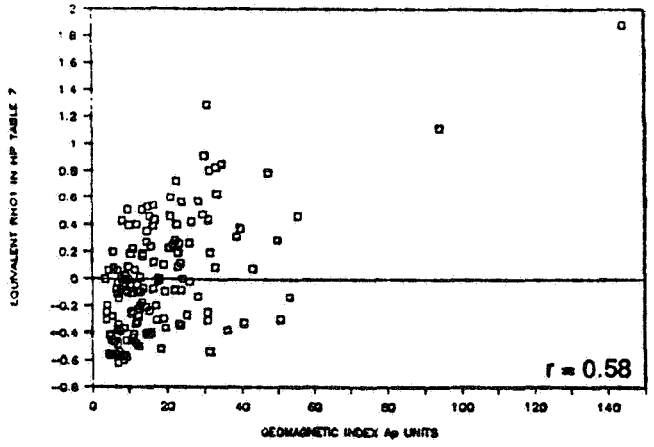


Figure 3e. DE-1 Equivalent ϱ_1 Values Versus A_p

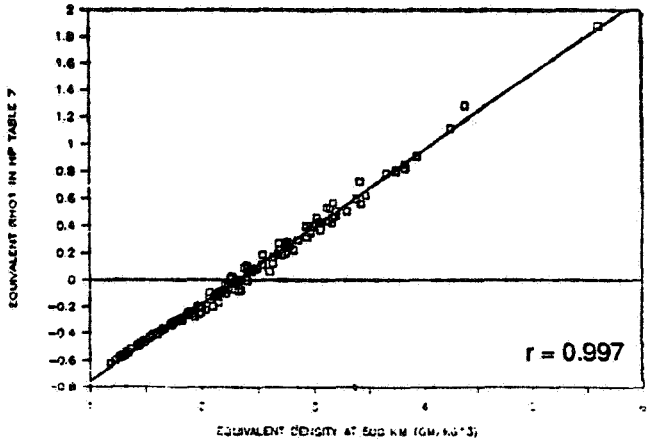


Figure 3f. DE-1 Equivalent ϱ_1 Values Versus Equivalent Density at 500 km

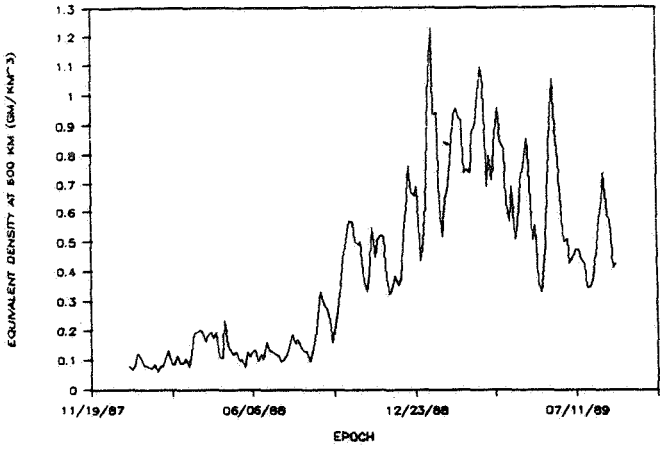


Figure 4a. ERBS Equivalent Density Versus Epoch

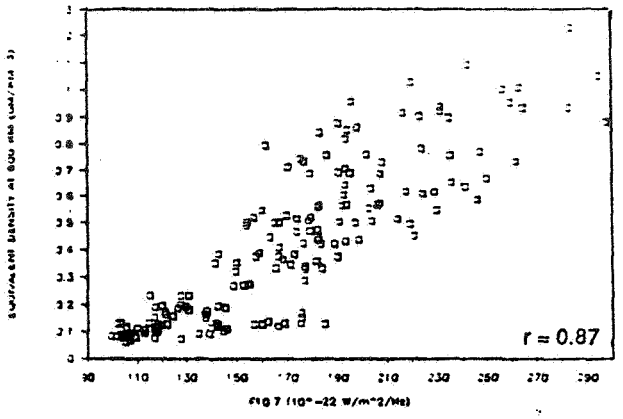


Figure 4b. ERBS Equivalent Density Versus $F_{10.7}$

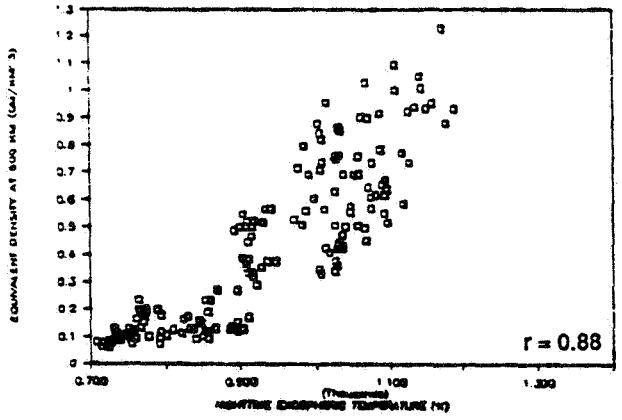


Figure 4c. ERBS Equivalent Density Versus T_c

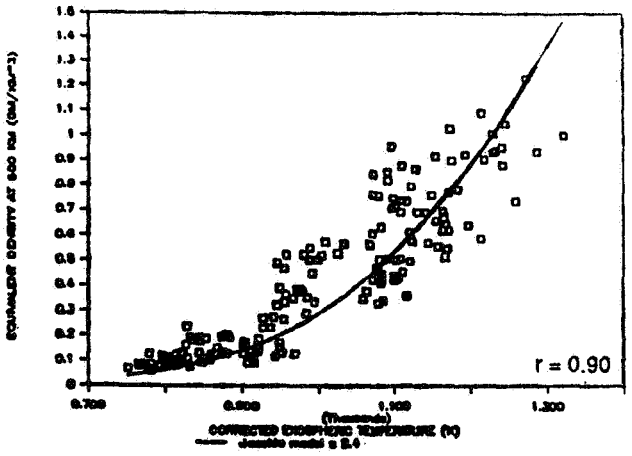


Figure 4d. ERBS Equivalent Density and Scaled Jacchia Model Versus T_∞

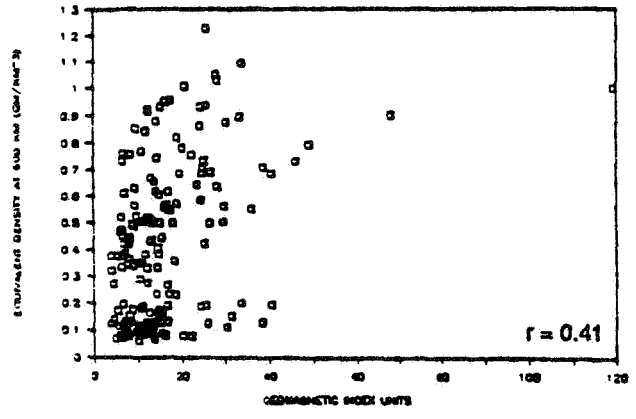


Figure 4e. ERBS Equivalent Density Versus A_p

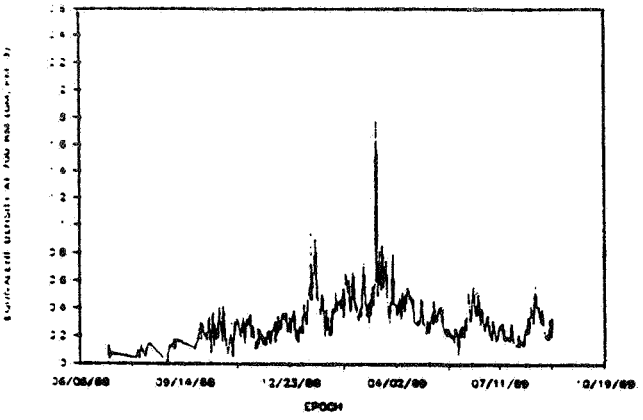


Figure 5a. Landsat-4 Equivalent Density Versus Epoch

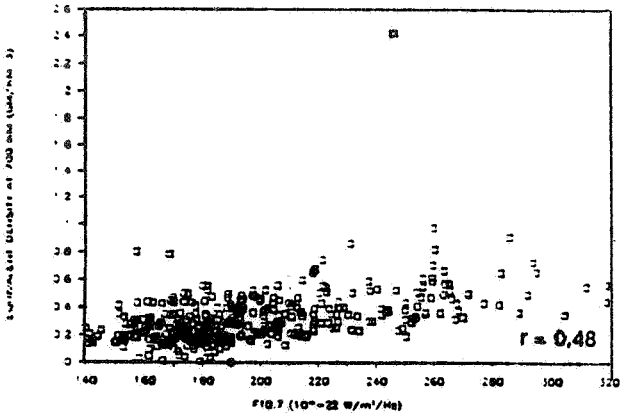


Figure 5b. Landsat-4 Equivalent Density Versus $F_{10.7}$

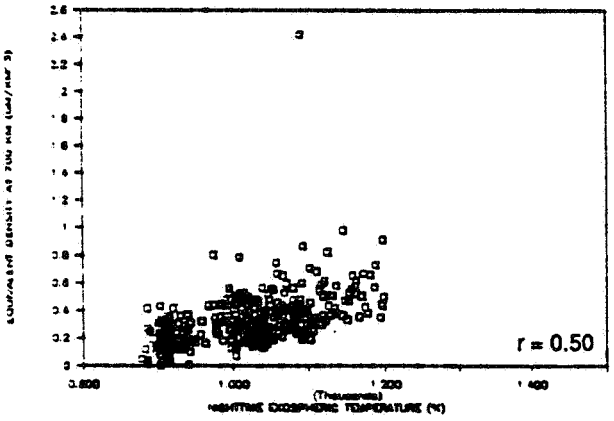


Figure 5c. Landsat-4 Equivalent Density Versus T_c

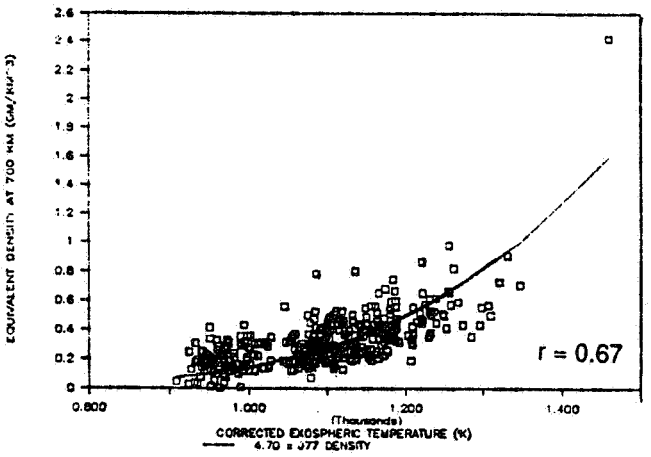


Figure 5d. Landsat-4 Equivalent Density and Scaled Jacchia Model Versus T_∞

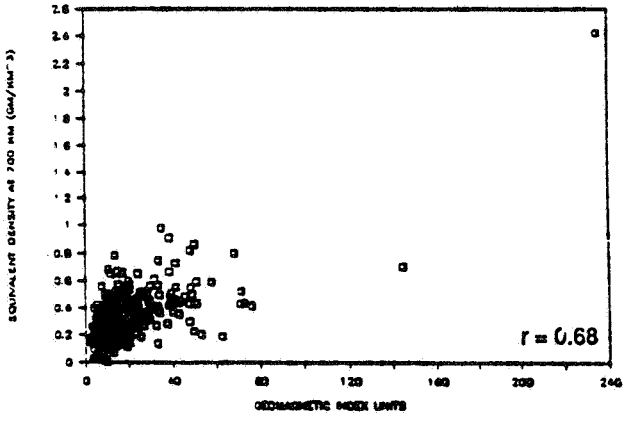


Figure 5e. Landsat-4 Equivalent Density Versus A_p

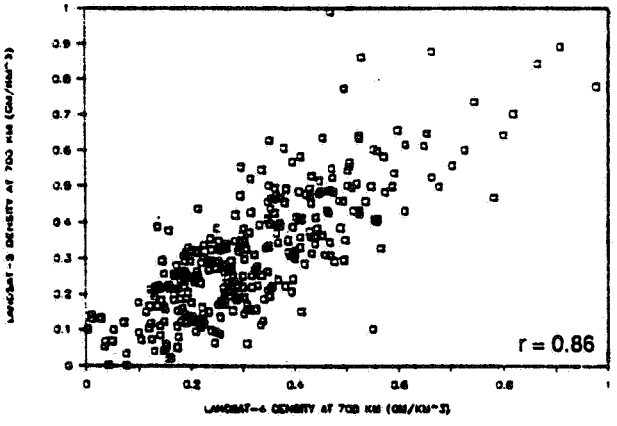


Figure 5f. Landsat-5 Equivalent Density Versus Landsat-4 Equivalent Density

C-3

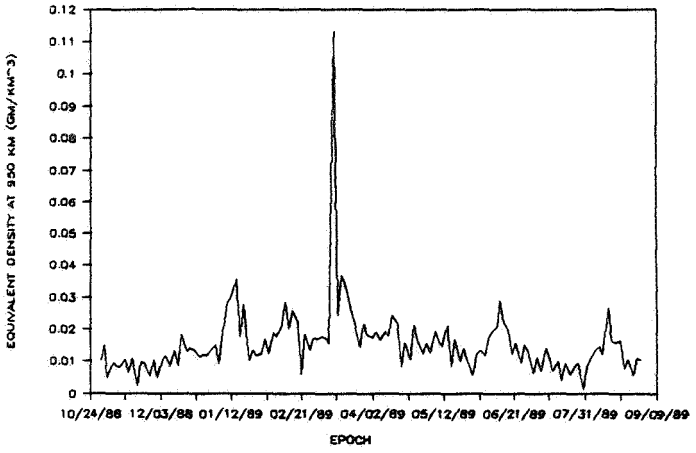


Figure 6a. Nimbus-7 Equivalent Density Versus Epoch

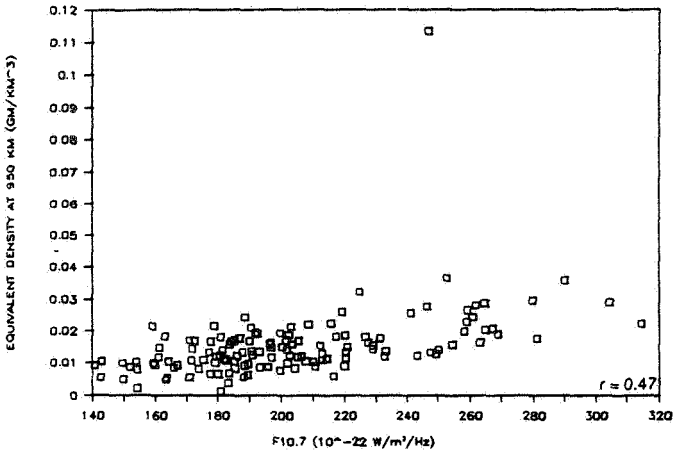


Figure 6b. Nimbus-7 Equivalent Density Versus $F_{10.7}$

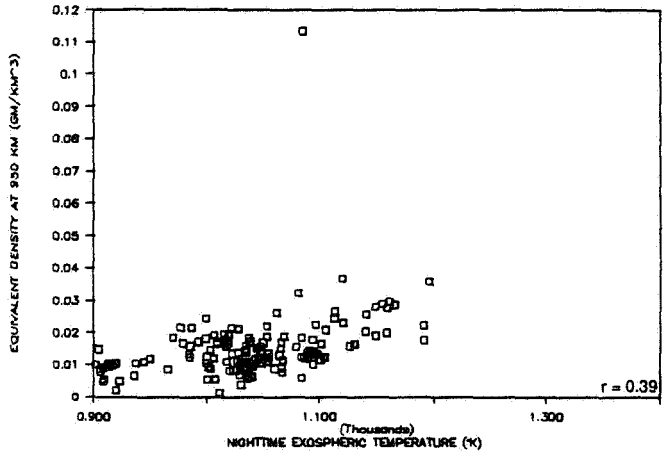


Figure 6c. Nimbus-7 Equivalent Density Versus T_c

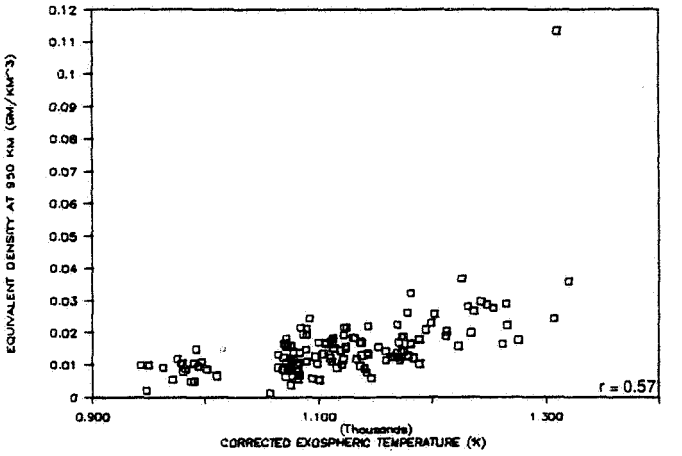


Figure 6d. Nimbus-7 Equivalent Density Versus T_∞

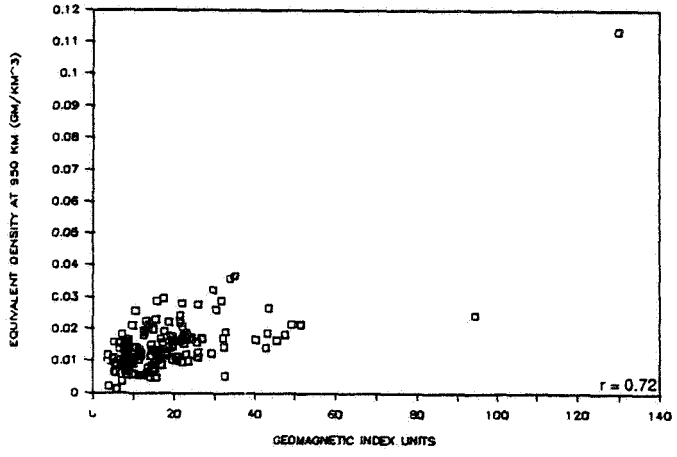


Figure 6e. Nimbus-7 Equivalent Density Versus A_p

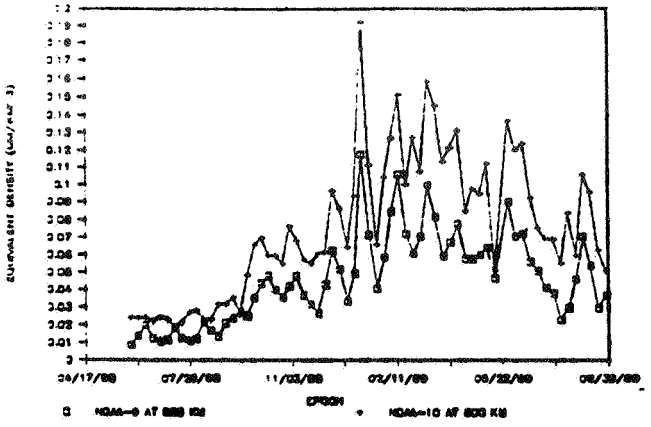


Figure 7a. NOAA-9, -10 Equivalent Density Versus Epoch

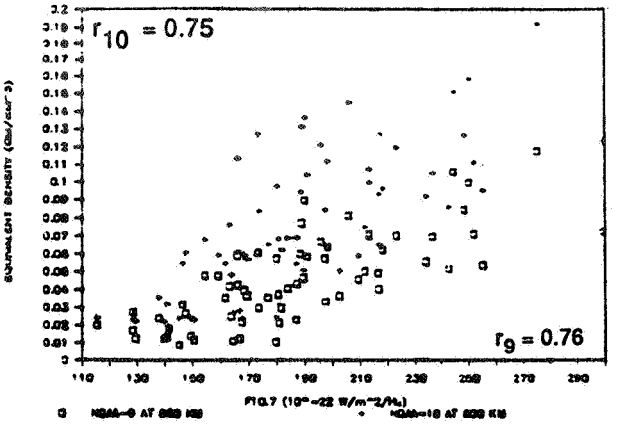


Figure 7b. NOAA-9, -10 Equivalent Density Versus $F_{10.7}$

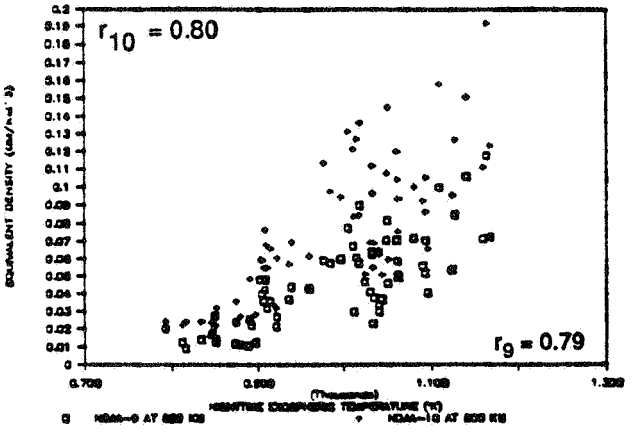


Figure 7c. NOAA-9, -10 Equivalent Density Versus T_c

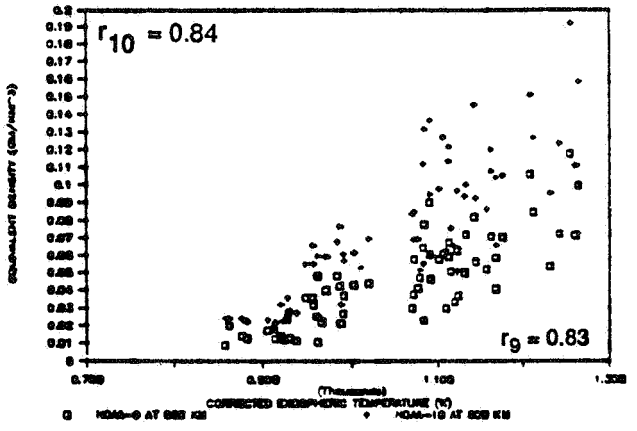


Figure 7d. NOAA-9, -10 Equivalent Density Versus T_∞

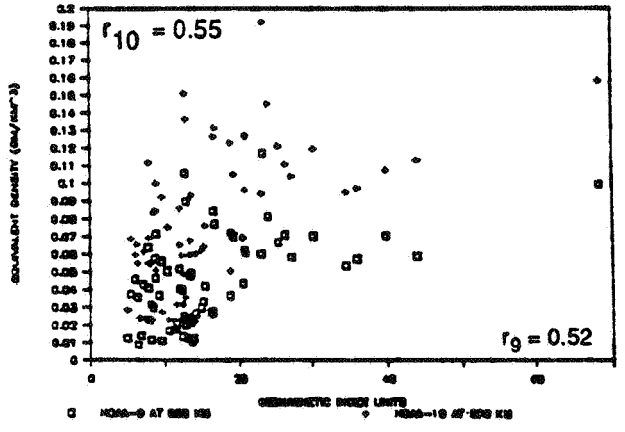


Figure 7e. NOAA-9, -10 Equivalent Density Versus A_p

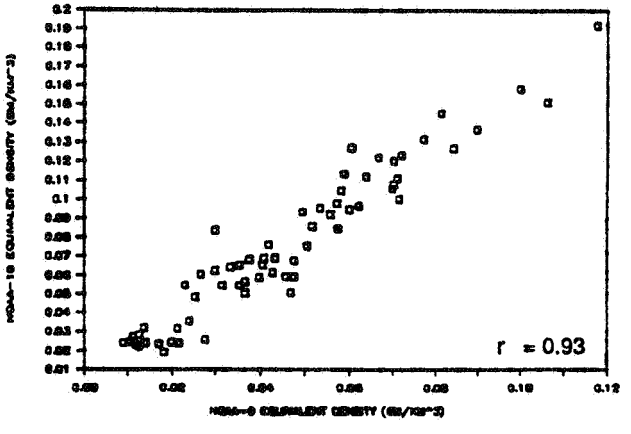


Figure 7f. NOAA-10 Equivalent Density Versus NOAA-9 Equivalent Density

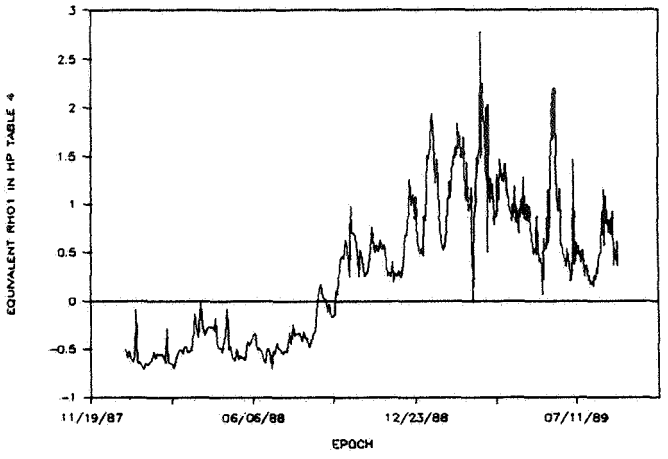


Figure 8a. SMM Equivalent ϱ_1 Values at HP 4 Versus Epoch

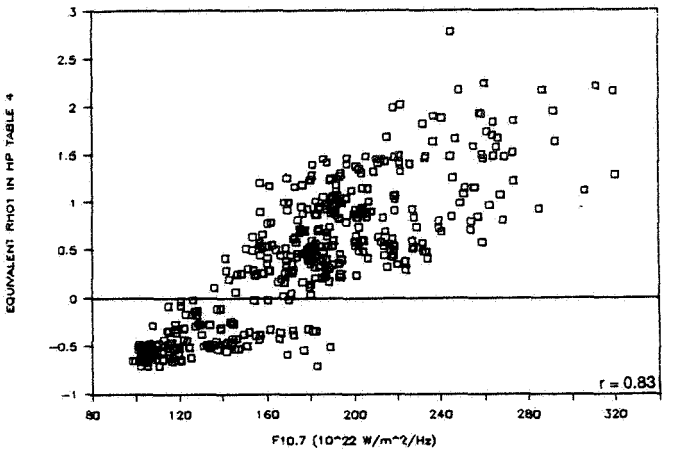


Figure 8b. SMM Equivalent ϱ_1 Values Versus $F_{10.7}$

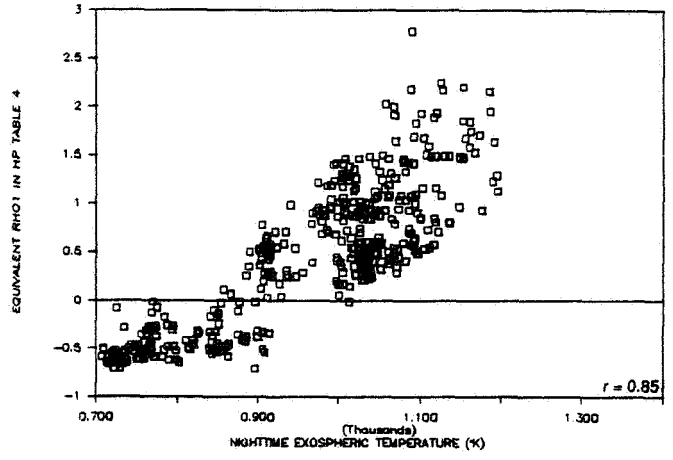


Figure 8c. SMM Equivalent ϱ_1 Values Versus T_c

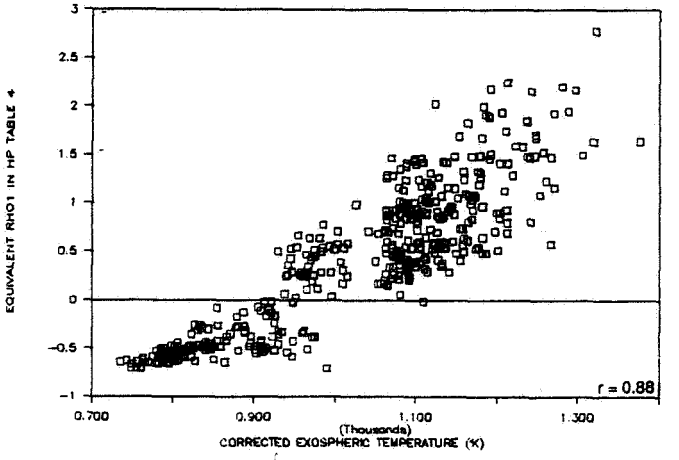


Figure 8d. SMM Equivalent ϱ_1 Values Versus T_∞

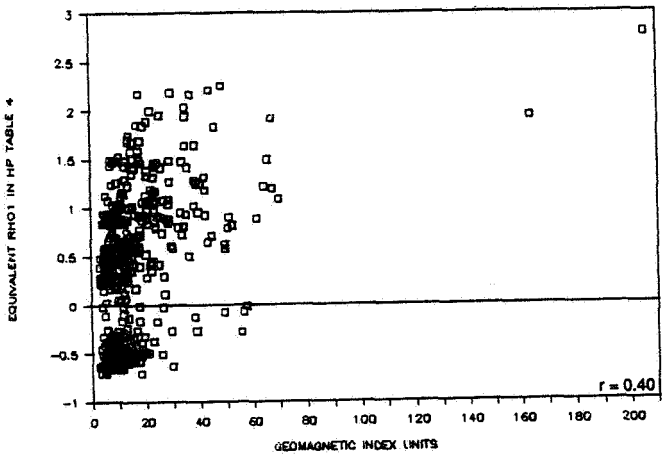


Figure 8e. SMM Equivalent ϱ_1 Values Versus A_p

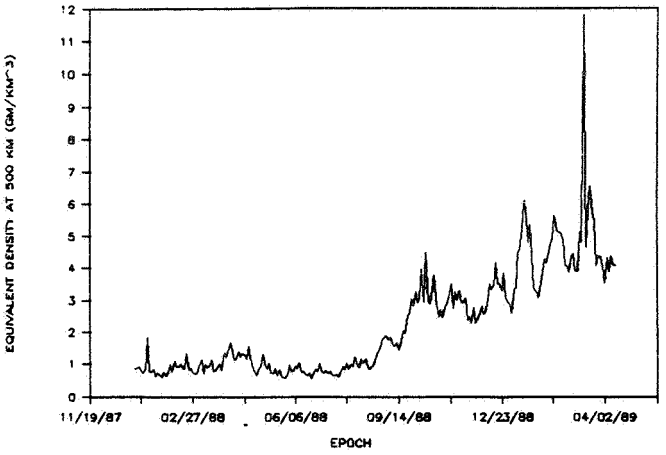


Figure 9a. SME Equivalent Density Versus Epoch

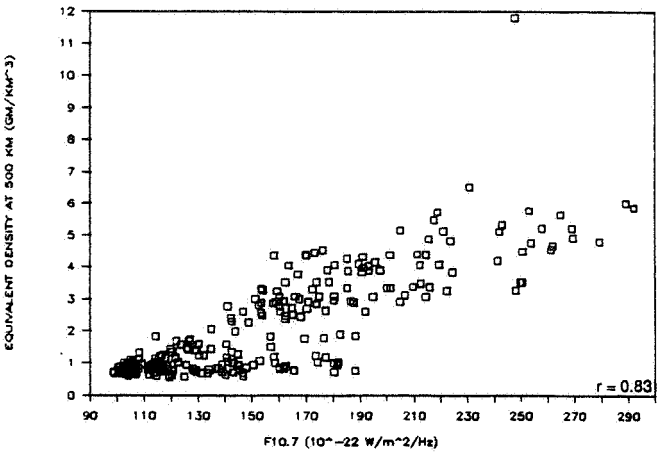


Figure 9b. SME Equivalent Density Versus $F_{10.7}$

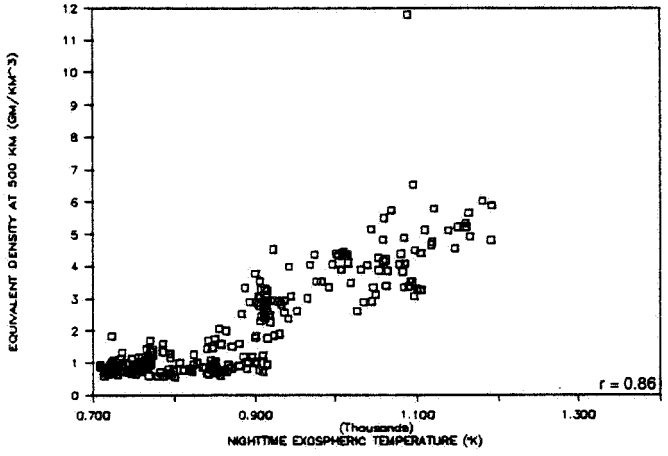


Figure 9c. SME Equivalent Density Versus T_c

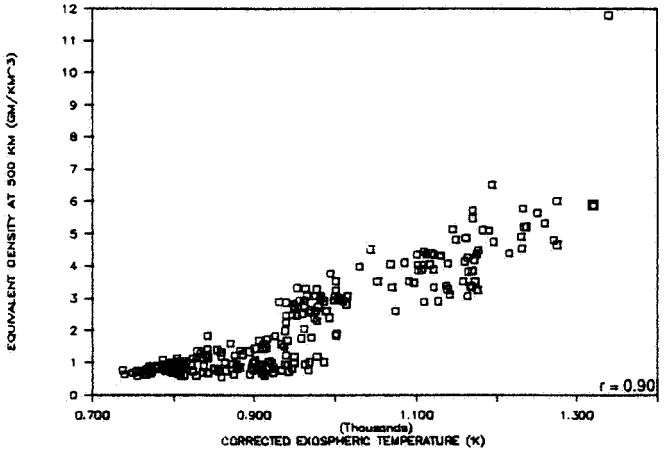


Figure 9d. SME Equivalent Density Versus T_∞

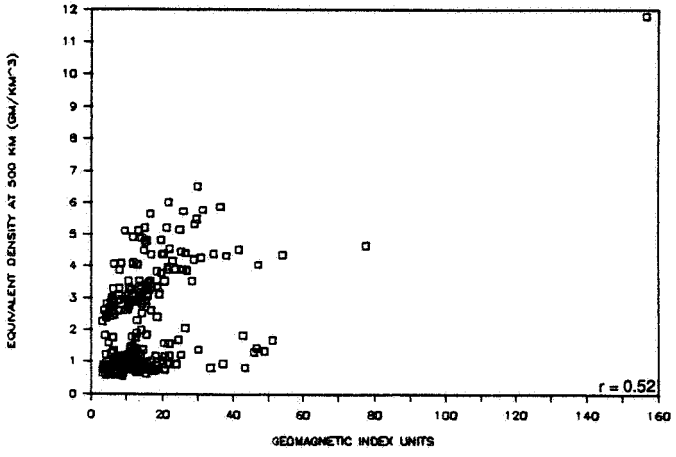


Figure 9e. SME Equivalent Density Versus A_p

N91-17084

"Procedures for Minimizing the Effects of High Solar Activity on Satellite Tracking and Ephemeris Generation"

by

Gordon D. Bredvik

ABSTRACT

We are currently experiencing a period of high solar radiation combined with wide short-term fluctuations in the radiation. The short-term fluctuations, especially when combined with highly energetic solar flares, can adversely affect the mission of U.S. Space Command's Space Surveillance Center (SSC) which catalogs and tracks the satellites in orbit around the earth.

Rapidly increasing levels of solar electromagnetic and/or particle radiation (solar wind) causes atmospheric warming, which, in turn, causes the upper-most portions of the atmosphere to expand outward, into the regime of low altitude satellites. The increased drag on satellites from this expansion can cause large, unmodeled, in-track displacements, thus undermining the SSC's ability to track and predict satellite position.

On 13 March 1989, high solar radiation levels, combined with a high-energy solar flare, caused an exceptional amount of short-term atmospheric warming. The SSC temporarily lost track of over 1300 low altitude satellites--nearly half of the low altitude satellite population. Observational data on satellites that became lost during the days following the 13 March "solar event" was analyzed and compared with the satellites' last element set prior to the event (referred to as a geomagnetic storm because of the large increase in magnetic flux in the upper atmosphere). The analysis led to a set of procedures for reducing the impact of future geomagnetic storms. These procedures adjust selected software limit parameters in the differential correction of element sets and in the observation association process and must be manually initiated at the onset of a geomagnetic storm. Sensor tasking procedures must be adjusted to ensure that a minimum of four observations per day are received for low altitude satellites. These procedures have been implemented and, thus far, appear to be successful in minimizing the effect of subsequent geomagnetic storms on satellite tracking and ephemeris computation.

Introduction

On 13 March 1989, one of the stations which report three-hourly values of geomagnetic flux (Ap), Fredricksburg, Virginia, reported a level of flux averaged over all eight three-hourly measurements which was the highest recorded value since 1960 (Reference 3). This major geomagnetic storm was believed to be linked to a very

strong solar flare observed on the surface of the sun three days earlier, on 10 March, which was pointed almost directly at Earth. Also on 13 March, another strong flare occurred; the x-ray radiation from this flare, classified as an x-level, or highest energy-level flare, undoubtedly contributed to the intensity of the geomagnetic storm already in progress. Additional x-level flares occurring on the 14th, 16th, and 17th of March extended the duration and effects of the storm, which included dramatic displays of the aurora borealis as far south as the Caribbean (Reference 1). Other, much publicized effects of the storm included disruption of shortwave radio communications, power outages such as the one which blacked out parts of Montreal and the province of Quebec for as long as nine hours, and the spurious opening and closing of automatic garage doors (Reference 2). The storm had an even more dramatic effect on US SPACECOM's Space Surveillance Center (SSC), which lost track of over 1300 low altitude satellites--over half of the low altitude satellite population. Several days of intensive around-the-clock manual analysis effort was required to "catch-up" with the lost satellites and reduce the "lost list" to a marginal level. Several more weeks of manual effort was required to reduce the list back to nominal levels.

Rapid changes in solar radiation, such as occurred on 13-17 March, 1989, produce large unmodeled drag effects on low altitude satellites and can defeat automatic observation processing and element set maintenance. Prediction accuracy is thereby degraded, thus causing problems in identifying and tracking low altitude satellites.

Analysis of satellite observations and the SSC's observation processing during and after the March Event led to a set of software procedures for minimizing the effects of future such events (large changes in solar radiation) on the SSC. A summary of the analysis and a description of the procedures themselves follows, preceded by a brief discussion of the effects of changes in solar radiation on the orbits of low altitude satellites.

Effects of Solar Radiation on Earth's Atmosphere and Low Altitude Satellites

Solar radiation warms the Earth's upper (tenuous) atmosphere through two primary effects: photoionization, caused by electromagnetic radiation, and ionization caused by collisions between solar wind particles (mostly electrons and protons), and air molecules. Increases in solar radiation increase atmospheric warming, which in turn causes the tenuous atmosphere to expand outward. This expansion increases atmospheric density in the realm of low altitude satellites, producing increased drag. Conversely, reductions in solar radiation cause the tenuous atmosphere to contract, thus reducing drag on low altitude satellites.

The SSC measures the amount of drag recently encountered by low altitude satellites and assumes this value remains constant as it predicts future satellite positions. (There are exceptions to this rule, such as when high interest satellites are maintained using special perturbations theory. For these special cases, measurements of solar radiation are incorporated in ephemeris prediction). If changes in drag are gradual and drag measurements are taken frequently enough, the assumption of constant drag provides satisfactory accuracy. However, when large, rapid changes in solar radiation produce similarly large, rapid changes in atmospheric drag, measurement of drag can lag significantly behind what is actually being experienced, and orbit prediction accuracy can deteriorate.

There are both long and short-term variations in solar radiation. The 11-year solar cycle is an example of long-term variation. A solar flare, which can last from minutes to hours, is an example of a short-term variation. Another example of short term variation is the fluctuation in total solar radiation that occurs during the peaks in the eleven-year cycle; these fluctuations can run their course in as little as a few days. The SSC's method of satellite drag prediction easily accommodates long-term fluctuations in solar radiation. However, solar conditions such as those which occurred 10-17 March, 1989, in which the solar wind from a major flare arrived at the earth at the same time that a peak in the short-term fluctuation in total radiation was occurring, can cause drag to change significantly in a matter of half a day or less. For example, a spike in Ap, such as occurred in 10-17 March 1989, can cause a 45 nm displacement in a near earth satellite's predicted position in just 12 hours (perigee = 185 km). It was this rapid increase in drag on low altitude satellites (generally, satellites with periods less than 110 minutes) which caused the SSC to temporarily lose track of over 1300 satellites.

Overview of the SSC's Satellite Observation Processing and Satellite Element Set Maintenance Segments

The SSC maintains the element sets of nearly 7000 satellites in a file called the SATF. The element sets are updated each time observations are received through a process called differential correction, similar to a Kalman filter or, in effect, a seven dimensional least squares fit. A complete differential correction (DC) is not performed each time an observation is received; instead, a simpler "sequential" DC is performed until a preset period of time expires, at which time a full DC is performed using all of the observations received during a period of time referred to as the Length of Update Interval (LUPI). The LUPI varies from 5 to 14 days for low altitude satellites, depending on the apogee and perigee heights of their orbits. The sequential DC permits only relatively small changes to the satellite's element set but has the advantage that it is much faster than a full DC, since it does not attempt to do a "least squares" fit of all LUPI

observations (it performs a simple "update" of the covariance matrix) and thus requires only one "pass". Once the update interval for a given satellite expires, or the sequential DC fails, or the element set accuracy declines past a given point, a full DC is performed for that satellite's element set. A full DC requires a minimum of seven observations.

Element sets from the SATF are periodically transmitted to spacetrack sensors in the Space Surveillance Network. The sensors use the latest element set received from the SSC to predict look angles for satellites for which they are tasked to provide observations. The sensors compare their observations against the predicted positions of the satellites using the latest element set received from the SSC. If the comparison is within established association criteria, the sensor will tag the observation with the requested satellite number and send it to the SSC as a routine observation. If it does not meet the association criteria, the sensor will tag it as an Unknown Object (UO) and send it to the SSC tagged with a 9XXXX number in place of the satellite number. The processing of UO observations within the SSC is much more tedious and time consuming because they have to be compared against every satellite in the SATF file.

Deterioration in the Observation Processing and Element Set Maintenance Segments during the March 1989 Solar Event

Prior to the Solar Event of March 13-17, 1989, Air Force Space Command (AFSPACECOM) had anticipated that problems might be encountered during the upcoming peak in the solar cycle and had tasked Kaman Sciences (on contract to maintain the SSC software) to examine ways to minimize the effects of changing solar radiation on the SSC's mission of satellite tracking and ephemeris prediction. Thus, at Kaman's request, satellite observation data taken before, during, and after the March Solar Event was saved for analysis. Using this data, Kaman recreated the March Solar Event scenario in AFSPACECOM's Off-Site Test Facility using computers and software similar to the SSC's. The analysis showed that deterioration occurred in three chronological phases.

Phase 1. During the initial phase, low altitude satellites began reflecting the effects of the increased drag, caused by the solar flare-induced geomagnetic storm, through relatively large differences between their observed and predicted positions. Although still within SSC and sensor association limits, the magnitude of the difference caused the satellites to fail sequential DC updates or fail the SSC's internal accuracy check; they were then scheduled for a full DC update. During normal periods, this process is able to "catch" and update delinquent element sets before they degrade further. However, during this solar event, the full DC typically either rejected the latest observations because they were too far from "nominal", or, when too many recent observations were

available to ignore, the full DC simply failed. In either case, the element set was not updated with the new observations. When the full DC failed, it failed either because it could not accept the large change in drag and mean motion suggested by the newest observations, or the full DC could not meet convergence criteria, or both. The convergence criteria could not be met because the disparity between the old (before flare) and new (post flare) observations prevented an acceptable "fit".

Phase 1 - Suggested Fixes. To correct the problems in this phase, the following steps suggested themselves.

- Reduce LUPI. If the Length of Update Interval (LUPI) were reduced to the point where most or all of the observations to be used in the DC update were new (post flare), then the DC would not be permitted to ignore the post flare observations. Further, convergence criteria could be achieved since the preponderance of the observations were post flare. Testing showed that convergence could be achieved with a LUPI of three days in nearly every case. Any further reduction greatly increased the likelihood that less than the minimum number of observations (seven) would be available for the full DC.

- Increase Parameter Change Limits for Full DC. To prevent full DCs from failing because drag and mean motion changes exceeded normal limits, these limits were increased. Typically, the effects of the increase or decrease in satellite drag caused by changes in solar radiation are in-track. For example, the effect of increased drag is to drop the satellite to a lower orbit, thus reducing the semi-major axis of the orbit and increasing mean motion. Other orbit parameters will remain largely unchanged. The orbit plane will remain the same, and only very slight changes in altitude and eccentricity will occur. To the spacetrack sensor, the orbit will appear the same, except the satellite will be ahead of its predicted position (case of increased drag), or behind its predicted position (case of decreased drag).

  Thus, during periods of high solar activity, the full DC should be permitted to accept wider changes in drag and the parameter which describes in-track motion, mean motion (n, revolutions per day). Testing indicated that the limit to changes in drag should be increased by a factor of ten and the limit for changes in mean motion increased by a factor of six.

- Increase Observation Flow Rate. If the daily observation flow rate is not increased when the LUPI is decreased to three days, the number of observations available for the full DC will be drastically reduced, close to, or even below the minimum of seven. To prevent the full DC from failing because of a lack of the minimum number of observations, and to maintain an acceptable level of accuracy in the full DC, sensor tasking must be increased to ensure that a minimum of four observations per day are received for all low altitude satellites. If a still higher number of observations were received, accuracy could approach that of normal solar conditions with normal LUPI values.

Phase 2. During Phase 2, the satellite element sets were so inaccurate that the SSC could no longer associate routine observations with their element sets. During this phase, many of the sensors were still able to associate the observations with the drag-displaced satellites because the sensors have considerably wider association criteria than the SSC. In this phase, the SSC typically detags properly tagged routine observations and places them in an Unassociated Observations File as unknown observations.

Phase 2 - Suggested Fix: Increase In-Track Multiplier. This problem can be corrected by increasing the association criteria of the SSC to equal that of the sensors in the in-track direction. Analysis showed that if the in-track association multiplier were doubled, i.e. increased from 3 to 6 for low altitude satellites, the SSC's in-track association criteria would approximate that of the Eglin phased array sensor. (Association criteria is not the same for every spacetrack sensor. Eglin was selected because, as a dedicated spacetrack sensor, it provides more observations than any other phased array sensor).

Phase 3. During Phase 3, the element sets are so old that neither the SSC nor the spacetrack sensors can associate the observations with the drag-displaced satellites. Thus, all of the observations on these satellites are tagged as Unknown Objects by the sensors. When they arrive at the SSC, they eventually wind up in the Unassociated Observations File. At this point, trained Orbital Analysts must manually retrieve the observations from this file and attempt to associate them with element sets in the SATF using various software tools currently available. This process is very tedious and time consuming and greatly increases CPU usage. However, this intensive manual interaction was necessary 24 hours a day for several days after the March Solar Event. Once this phase is reached, there is no known method for enhancing the manual recovery processes already known to SSC analysts and crew personnel. The objective of the fixes recommended in the two

previous phases is to prevent Phase 3 from ever being reached.

Summary of SSC Performance Parameters During the March 1989 Solar Event

During the three phases described above, post-event analysis revealed that the percentage of observations that were tagged as Unknown Objects increased from 18% to over 40%; satellite catalog accuracy dropped drastically; the success rate for full DCs dropped sharply; total observation flow rate increased dramatically; CPU usage rose to the point of saturation; and, as was previously mentioned, the satellite "lost list" increased to over 1300 satellites. The normal level for this list is below 100 satellites.

Summary of Recommended Software Procedures

The following procedures were recommended for implementation in the SSC. They apply to low altitude satellites (non-payloads) with periods less than 110 minutes:

1. Increase spacetrack sensor tasking to ensure a minimum flow of four observations per day. Maintain this tasking until solar maximum subsides in late 1994.

2. When Ap rises over 100, or a sharp change in F_{10} (solar radiation measured at 10.7 cm wavelength, in units of 10^{-22} watts/M^2/Hertz) occurs, implement the following steps:

   a) Decrease LUPI to three days;

   b) Increase the in-track multiplier for each satellite to 6;

   c) Increase DC change limit multiplier for drag to 10;

   d) Increase the DC change limit multiplier for mean motion to 6.

   Return the above parameters to normal approximately three days after solar activity returns to normal.

Conclusion

The software procedures described in this paper have been adopted by USSPACECOM's Space Surveillance Center (SSC) as standard operating procedures during periods of high solar activity since late June, 1989. Their objectives are to maintain the satellite "lost list", the Unknown Object (UO) rate, and CPU (Central Processing Unit) usage at near normal levels during such periods. Thus far, these objectives have been achieved.

REFERENCES

1. Marcia Bartusiak, "The Sunspot Syndrome"; Discover, November 1989, p. 45.

2. Leon Jaroff, "Fury on the Sun"; Time, July 3, 1989, p. 46.

3. "Preliminary Report and Forecast of Solar Geophysical Data", published weekly by the Joint NOAA-USAF Space Environment Center, SESC PRF 707, 21 March 1989.

1990 FLIGHT MECHANICS/ESTIMATION THEORY SYMPOSIUM
Goddard Space Flight Center
May 22-24, 1990

SPACE STATION FREEDOM ALTITUDE STRATEGY

Brian M. McDonald

Scott B. Teplitz

McDonnell Douglas Space Systems Company - Engineering Services
Houston, Texas

ABSTRACT

The Space Station Freedom (SSF) altitude strategy provides guidelines and assumptions to determine an altitude profile for Freedom. The process for determining an altitude profile incorporates several factors such as where the Space Shuttle will rendezvous with the SSF, when reboosts must occur, and what atmospheric conditions exist causing decay.

The altitude strategy has an influence on all areas of SSF development and mission planning. The altitude strategy directly affects the micro-gravity environment for experiments, propulsion and control system sizing, and Space Shuttle delivery manifests. Indirectly the altitude strategy influences almost every system and operation within the Space Station Program.

Evolution of the SSF altitude strategy has been a very dynamic process over the past few years. Each altitude strategy in turn has emphasized a different consideration, examples include a constant Space Shuttle rendezvous altitude for mission planning simplicity, or constant micro-gravity levels with its inherent emphasis on payloads, or lifetime altitudes to provide a safety buffer to loss of control conditions.

Currently a new altitude strategy is in development. This altitude strategy will emphasize Space Shuttle delivery optimization. Since propellant is counted against Spcae Shuttle payload-to-orbit capacity, lowering the rendezvous altitude will not always increase the **net** payload-to-orbit, since more propellant would be required for reboost. This altitude strategy will also consider altitude biases to account for Space Shuttle launch slips and an unexpected worsening of atmospheric conditions. Safety concerns will define a lower operational altitude limit, while radiation levels will define upper altitude constraints.

This paper will discuss the evolution of past and current SSF altitude strategies and the development of a new altitude strategy which focuses on operational issues as opposed to design.

I INTRODUCTION

Productive utilization of the Space Station Freedom (SSF) will depend on a careful blend of many operational factors. One of the most influential of these factors is to determine what operational altitudes SSF should fly. This paper discusses the current program altitude strategy used for system design and introduces modifications to this strategy intended to transition altitude related considerations into the operational era.

The organization of the remaining portion of this paper is as follows: Section II familiarizes the reader with the background necessary for a basic understanding of SSF altitude strategies. Section III presents a historical perspective of SSF altitude strategy evolution. The primary drivers as well as advantages and disadvantages of each strategy are discussed. Section IV presents the current program altitude strategy used to design and size systems. Section V discusses considerations for an operational altitude strategy. Only through careful evaluation of such considerations can an altitude strategy be developed which optimizes the key performance parameters. Finally, Section VI presents conclusions regarding the operational altitude strategy and future altitude work.

II BACKGROUND

The term ***altitude strategy*** is used throughout this paper and must be defined at this point. An altitude strategy refers to a set of guidelines and assumptions necessary to determine where SSF will operate. Altitude planners will use the altitude strategy to generate a set of lower (Space Shuttle/SSF rendezvous) and upper (reboost) altitudes. Specifically, the guidelines presented in the altitude strategy will provide a methodology for computing both lower and upper altitudes while the assumptions provide the necessary conditions to perform the analysis. From the altitude strategy, mission planners can estimate Space Shuttle delivery capability for manifest planning and reboost requirements for propulsion system sizing and resupply.

Currently, the lower (rendezvous) altitude is constrained by lifetime to a loss of control altitude. At the loss of control altitude (assumed to be 150 nmi [278 km]), the atmospheric torques would quickly overwhelm the SSF control system, making rescue impossible and catastrophic re-entry inevitable. This lifetime altitude was chosen to give the SSF Program adequate response time (90 days) in case of a total propulsion system failure. Thus, the lowest allowable altitude at which SSF may operate is defined as 90 days of decay to 150 nmi and is herein referred to as the **lifetime altitude**. System designers use the lifetime altitude as a design point since it represents the highest atmospheric densities the SSF will encounter. During the operational era, SSF altitude planners will choose rendezvous altitudes based on numerous factors. These factors include SSF safety, life cycle costs, delivery system utilization, radiation limits, mission planning, orbital debris density, mission requirements, and launch window considerations. Even after consideration of all previously mentioned factors, SSF altitude planners may still want to further bias the planned SSF rendezvous altitudes from the lifetime altitude to account for unpredictable, yet expected, deltas (e.g., Space Shuttle launch slips or atmospheric worsening).

The upper (reboost) altitudes are determined from the rendezvous altitudes and the Space Shuttle flight schedule. The SSF will reboost to an altitude such that at the end of the flight interval, SSF will have decayed down to the chosen rendezvous altitude by the next planned Space Shuttle visit. Figure 1 depicts a segment of an altitude profile. It is assumed SSF will reboost as soon as operationally possible after Space Shuttle departure to ensure Space Shuttle/SSF rendezvous at the lowest point possible, thus maximizing Space Shuttle delivery capability.

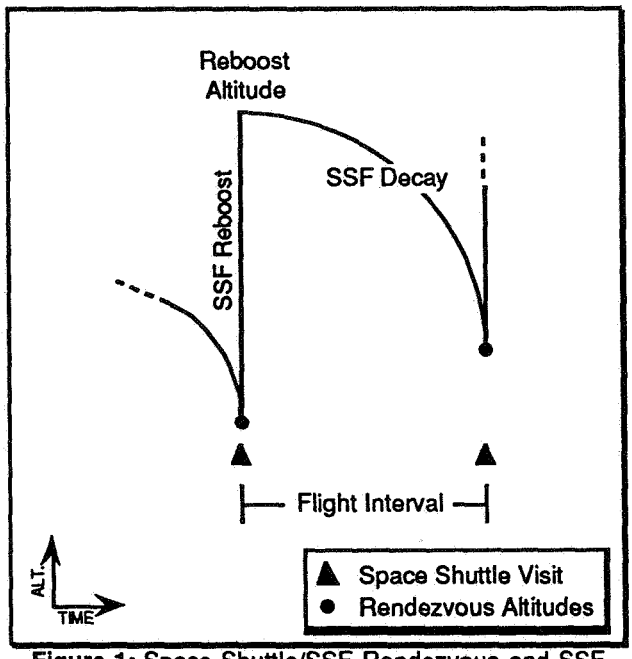


Figure 1: Space Shuttle/SSF Rendezvous and SSF Reboost Profile

Determining both the rendezvous and reboost SSF operating altitudes is dependent on SSF rate of decay for a specified period of time, the Space Shuttle flight interval. The rate of decay is primarily tied to two parameters: the *atmospheric density* and the *ballistic number* (BN, characterizing a vehicle's resistance to orbital decay).

Atmospheric Density and Solar Cycle Predictions

Density is the key atmospheric parameter used by trajectory analysis programs and represents the greatest uncertainty for altitude planners. The rate of vehicle altitude decay is proportional to the atmospheric density and inversely proportional to the BN:

$$\frac{dh}{dt} \approx -\frac{C\rho(r+h)^{1/2}}{BN} \quad (1)$$

where h is the SSF altitude, ρ is the atmospheric density as a function of h, r is the mean radius of the earth, and C is a constant. Atmospheric density calculations are based on the energy output from the sun, which varies over an eleven year solar cycle. During peak solar energy output, the Earth's atmosphere expands outward, similar to a balloon when heated. Likewise, during minimum solar energy output, the atmosphere contracts. The net result is that any vehicle maintaining a relatively constant altitude experiences widely varying density levels during a solar cycle.

There are numerous methodologies available for determining atmospheric density. The model accepted by the SSF Program is the **Jacchia 1970 atmospheric model** (Reference 1). This model primarily uses two solar parameters to calculate density, *solar flux ($F_{10.7}$)* and *geomagnetic index (A_p)*. Given the variation of both measurements over a solar cycle, the Jacchia atmosphere model will predict the atmospheric density for any given date and orbital position (altitude, latitude, and longitude).

Predictions for both $F_{10.7}$ and A_p are provided by the Mission Analysis Division at the Marshall Space Flight Center (MSFC) (Reference 2). Predictions for both mean (statistically, the actual value should be below the predicted value 50% of the time) and +2σ (statistically, the actual value should be below the predicted value 97.7% of the time) atmospheres are provided by MSFC and are shown in Figure 2. Generally, the +2σ atmospheric predictions are considered conservative, used mainly during system sizing and critical operational periods such as assembly. The mean atmospheric predictions are used when estimating resupply/return requirements during nominal SSF operations after assembly complete (AC). Solar cycle predictions have been normalized to an 11-year cycle. Actual past solar cycles have ranged from 9 to 13 years. Because actual $F_{10.7}$ and A_p values

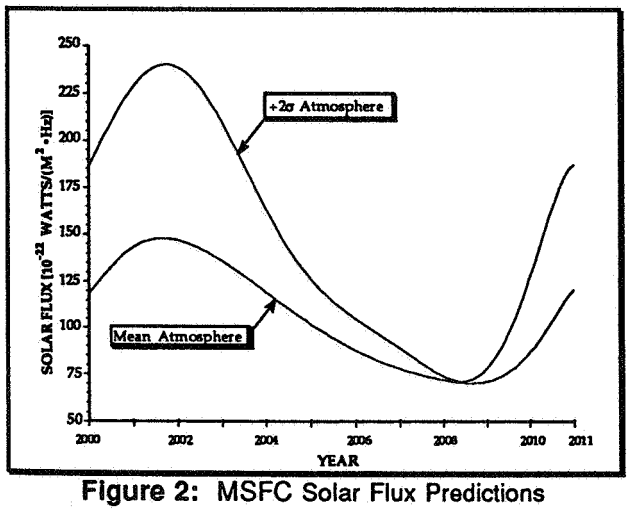


Figure 2: MSFC Solar Flux Predictions

may be significantly different than predicted, the operational altitude strategy must specify how and when SSF reacts to changes in atmospheric conditions.

Ballistic Number Estimations

While the BN is a key parameter in determining SSF orbital decay, SSF configuration experts are still uncertain as to how accurate BN predictions can be at this stage of SSF development. BN is calculated using the following relationship:

$$BN = \frac{\text{SSF Weight}}{C_d \cdot \text{Area}} \quad (2)$$

where Area equals the area exposed in the direction of motion($+x_{LVLH}$), and C_d equals 2.3 (a typical drag coefficient for orbiting spacecraft).

The exposed area in the $+x_{LVLH}$ direction varies over an orbit due to articulating elements such as the solar arrays and thermal radiators. The exposed area also varies with SSF attitude. While torque equilibrium attitudes are maintained, the exposed area varies over an orbit since these attitudes are adjusted to account for atmospheric density changes. Therefore, even the best predictions for BN are an average for one orbit. During assembly, the BN varies significantly as SSF elements are added to the growing configuration. This fact produces uncertainty in determining SSF orbit lifetimes and precise rendezvous altitudes.

Altitude Strategies

There are four basic approaches to defining rendezvous altitudes as part of an altitude strategy. Each approach is centered around an operational preference considered to be of paramount importance (i.e., planning simplicity, disturbance levels, safety, or life cycle costs). Development of an operational altitude strategy will consider both the virtues and failings of each approach.

1. ***Constant altitude*** maintenance requires varying the magnitude of the reboost with changing density levels over the course of a solar cycle. A constant rendezvous altitude has an obvious benefit, i.e., mission planning simplicity. Design of standard Space Shuttle rendezvous profiles and long range payload-to-orbit estimates for manifest planning are definite advantages. However, the rendezvous altitude selected must not violate lifetime considerations at any time during the solar cycle. This forces the altitude selection to be based on predicted conditions at the solar cycle peak. The solar cycle peak represents a relatively small segment of the entire solar cycle, lasting only 6 to 18 months. At off-peak times during the solar cycle, the rendezvous altitudes are considerably higher than dictated by lifetime considerations, thus representing a Space Shuttle delivery penalty. Therefore, a constant rendezvous altitude profile trades Space Shuttle payload-to-orbit capability for operational planning simplicity.

2. ***Constant micro-gravity altitude*** maintenance limits the Space Shuttle/SSF operational altitudes to a maximum micro-gravity (μg) level. Since the maximum density encountered is at the lowest point in SSF trajectory, and atmospheric acceleration is a direct function of atmospheric density, a specified μg level defines the rendezvous altitudes. This approach takes advantage of the varying atmospheric density levels over a solar cycle by lowering rendezvous altitudes when the atmosphere contracts. In this way, Space Shuttle delivery capability can be appreciably increased compared to the constant altitude strategy.

   For the constant μg altitude strategy, each Space Shuttle/SSF rendezvous occurs at the same atmospheric μg level or decay rate. Therefore, the decay rate varies little over the solar cycle. Since both the rendezvous and reboost altitudes are tied to decay, any change in parameters which influences decay, such as the BN or atmospheric conditions, will cause the entire altitude profile to be biased, but result in relatively constant SSF propellant requirements. This aspect of the constant μg altitude strategy simplifies the design of the propulsion system and propellant resupply planning since requirements are not affected by major changes in SSF configuration or atmospheric predictions.

   This approach has the apparent benefit of providing users with a maximum expected μg environment during nominal operations (except during planned perturbations such as Space Shuttle docking and SSF reboost). Although this strategy was accepted by the SSF Program for many years, the basic premise is very misleading. The specified μg limits the aerodynamic torques only. Gravity gradient torques are considerably higher, as much as an order of magnitude within the laboratory modules. Gravity gradient torques vary with distance from the SSF center of gravity (CG). The farther from the CG, the greater the gravity gradient torque. Only at the CG are the gravity gradient torques equal to zero and the aerodynamic torques alone determine the overall SSF μg environment. Since few, if any, experiments could be located at the CG, users should not assume maximum disturbance levels are limited by altitude.

3. ***Constant lifetime altitude*** maintenance sets rendezvous altitudes at the minimum allowable lifetime level. This approach attempts to maximize the Space Shuttle payload-to-orbit capability by rendezvousing as low as possible. There are two major drawbacks to this approach. First, there is no altitude margin for unplanned or unexpected events such as a Space Shuttle launch slip or atmospheric worsening. Second, any change to SSF BN or atmospheric predictions will significantly change previous estimates for Space Shuttle payload-to-orbit and propellant requirements. The rendezvous altitudes are directly linked to the predictions for BN and atmospheric conditions.

   This approach is useful, however, in defining system sizing requirements since the minimum lifetime altitudes represent the lowest allowable operating altitudes and thus the largest system requirements. This lifetime altitude strategy is discussed in detail in Section IV and currently is incorporated in the *Space Station Projects Description and Requirements Document*, JSC 31000 (Reference 3), and is a change request to the *Space Station Program Definition and Requirements Document* (PDRD), JSC 30000 (Reference 4).

4. ***Optimal altitude*** maintenance sets the rendezvous altitude at a point which maximizes **net** payload-to-orbit (total Space Shuttle delivery capacity minus SSF reboost propellant requirements). Both the reboost propellant usage and Space Shuttle delivery capability are directly related to altitude. The lower the Space Shuttle rendezvous with SSF, the more Space Shuttle can deliver to orbit (a rule of thumb is an additional 100 lbm/nmi). However, the lower SSF operates, the more propellant required for reboost since the atmosphere is more dense, thereby causing greater decay. The altitude which maximizes Space Shuttle **net** payload-to-orbit is called the **optimal (optimum) altitude** (shown in Figure 3). The optimal altitude defines the altitude at which flying lower would cause more additional propellant to be used than gained in Space Shuttle payload-to-orbit, and flying higher would cause more Space Shuttle payload-to-orbit lost than would be saved in reduced propellant needs.

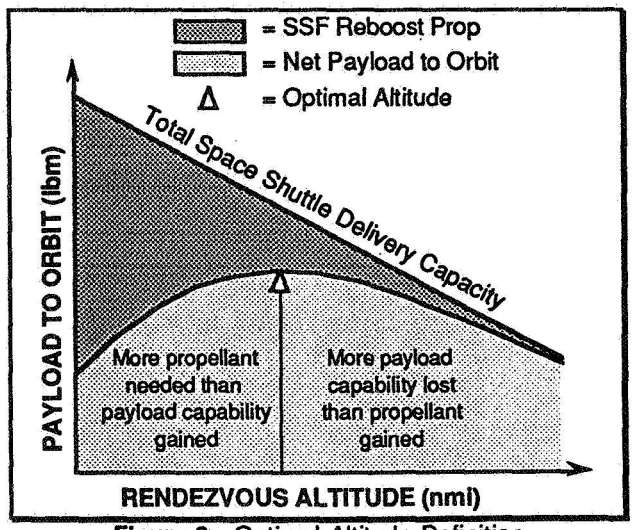


Figure 3: Optimal Altitude Definition

An altitude strategy based on optimal altitudes has the advantage of being relatively insensitive to changes in configuration and atmospheric predictions. Although these changes will cause the optimal point to move, the resultant loss in

net payload-to-orbit increases at a surprisingly slow rate as actual altitudes diverge from the optimal point. This approach will be used as a basis for an operational altitude strategy discussed in Section V.

III HISTORICAL VIEW

The SSF altitude strategy has evolved with program maturity over the past several years. Beginning simply with a constant altitude strategy early in Phase B, the altitude strategy has evolved to its present state where it is in transition from a system design emphasis to an operational emphasis. Although this process of change appears simple enough, it has been a long and arduous road. As the SSF Program evolved, various altitude related issues were rearranged in relative importance. Each priority change represented a new altitude philosophy or strategy. To date, altitude strategies have been used to identify operational envelopes for system design. Although these altitude strategies have been used as a basis for operational cost studies, it is understood that these altitude strategies have not adequately addressed operational issues which will ultimately drive the operational altitudes. A brief description of each altitude strategy as they evolved from early Phase B concept studies to the present is provided below.

Constant Rendezvous Altitude

270 nmi (500 km) Constant Altitude

Early Phase B rendezvous altitudes for the SSF were set at a constant 270 nmi. The 270 nmi altitude served to minimize the drag and thus the propellant requirements for resupply. This was the highest the SSF could fly and maintain safe levels of crew radiation exposure. Although the 270 nmi altitude was chosen without explicit concern for the SSF lifetime, it did provide sufficient safety at the peak of the solar cycle in terms of a catastrophic re-entry into the atmosphere. Maintaining a constant lower altitude meant that the reboost sizes needed to be varied throughout the solar cycle in order to decay to the same altitude for the next Space Shuttle rendezvous. Figure 4 clearly shows the variation of the reboost sizes with the changes in the solar activity.

At this time in program history, the standards used to determine acceptable crew radiation exposure levels were re-evaluated. Radiation standards established by the Occupational Safety and Health Administration (OSHA) were adopted in lieu of more liberal NASA standards. As a result, many of the reboost altitudes violated OSHA radiation exposure levels for both the eyes and skin. This forced a program requirement which set the upper bound for SSF operations at 270 nmi. To accommodate this new requirement rendezvous altitudes were lowered to 250 nmi (463 km) which ensured that operational altitudes would remain within the 270 nmi radiation limit.

250 nmi (463 km) Constant Altitude

Resulting from the concerns for crew radiation exposure levels, the rendezvous altitudes were lowered to 250 nmi. This strategy resulted in an altitude profile which at no time violated radiation limits for the crew. The constant 250 nmi constant altitude strategy (Figure 5) was accepted by the Space Station Program later in Phase B (circa 1984).

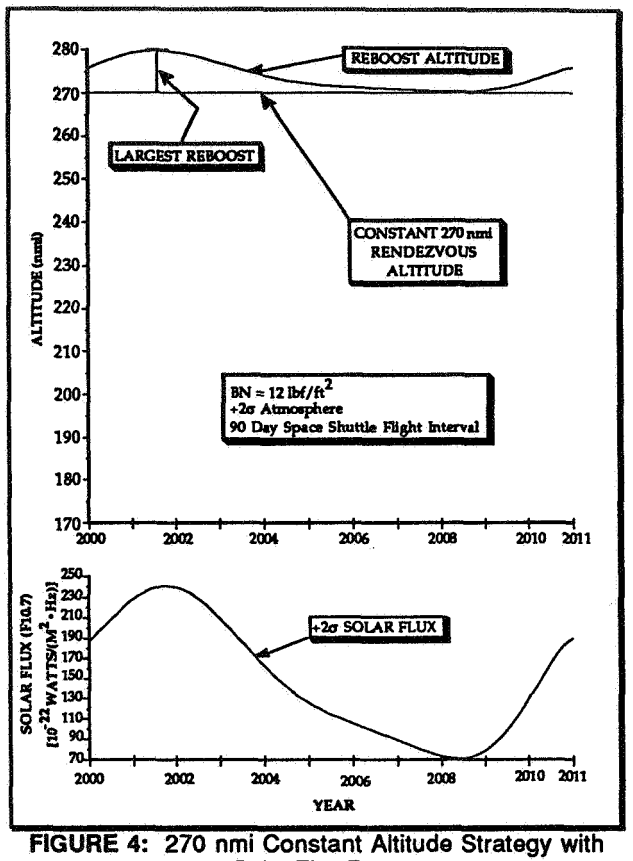


FIGURE 4: 270 nmi Constant Altitude Strategy with Solar Flux Data

Advantage of a Constant Altitude Strategy

A constant altitude strategy provides mission planners with a relatively constant target altitude and thus constant payload-to-orbit capability. Standardized Space Shuttle/SSF rendezvous profiles are also possible with a constant altitude strategy, thus simplifying such planning.

Disadvantages of a Constant Altitude Strategy

Maintaining a constant altitude profile has several disadvantages that severely impact the program. Although Space Shuttle payload-to-orbit remains constant over a solar cycle, large variations in the reboost complicate the overall Space Shuttle manifest planning. Another disadvantage is that a system designed using a constant altitude strategy severely limits operational flexibility. Several space station systems are sized based on constraints set by the

altitude strategy and Space Shuttle flight schedule. The operational altitude profile cannot be lower than the system design profile (constant rendezvous altitudes) since the systems can be sized to operate at higher altitudes. An obvious example of this would be the propulsion system; higher altitudes generally require smaller reboosts. This would result in propellant tanks inadequately sized to operate at lower altitudes.

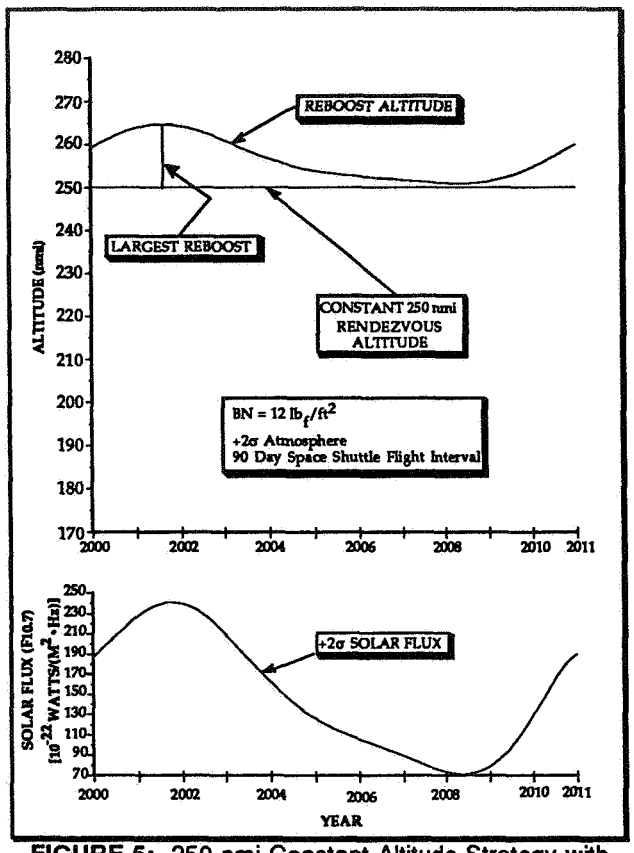


FIGURE 5: 250 nmi Constant Altitude Strategy with Solar Flux Data

Near the end of Phase B conceptual studies, assembly concerns began to surface. Additional assembly flights or additional payload-to-orbit are required to make assembly even possible. Since the constant altitude strategy was determined by choosing an altitude at the worst point in the solar cycle, altitudes at off-peak times could be lowered to gain significant Space Shuttle delivery performance while maintaining acceptable orbit lifetimes. At this point, payload-to-orbit was a paramount concern, and a new altitude strategy had to be developed in which the rendezvous altitude varied with the solar cycle.

Constant Micro-Gravitational Level

The constant μg altitude strategy was developed to take advantage of solar cycle changes by defining the rendezvous altitudes at a specified constant μg. The acceleration experienced by SSF is a function of the atmospheric density; however, the atmospheric density varies over the solar cycle. For this reason, and also so that SSF users could have an input in determining the altitude profile, a constant μg strategy became the accepted altitude strategy for SSF. The μg level represents the maximum drag acceleration SSF will experience due to atmosphere effects. This is because the SSF/Space Shuttle rendezvous occurs at the lowest point in SSF trajectory (reboost occurs shortly after Space Shuttle departs).

Constant .3 μg Level

The PDRD (JSC 30000) baselined the constant μg level at .3 μg (.3x10^{-6} g) later in Phase B, around 1986. Variations in the lower altitude caused by the solar cycle can be seen in Figure 6. The .3 μg altitude strategy introduced a concept referred to as the minimum controllable altitude (the point at which SSF was deemed uncontrollable and catastrophic re-entry was inevitable). However, this altitude could not be defined because the SSF Program was unable to agree upon the conditions which would describe this point. As it turned out, several of the lower altitudes defined by this strategy did not meet acceptable orbit lifetime limits. Therefore, to alleviate this concern, the μg level was lowered from .3 to .2 μg. This change increased the safety margin or lifetime, decreased the size of the reboosts since the new altitude profile was higher in the atmosphere, decreased the atmosphere disturbance level, and also decreased Space Shuttle delivery capability.

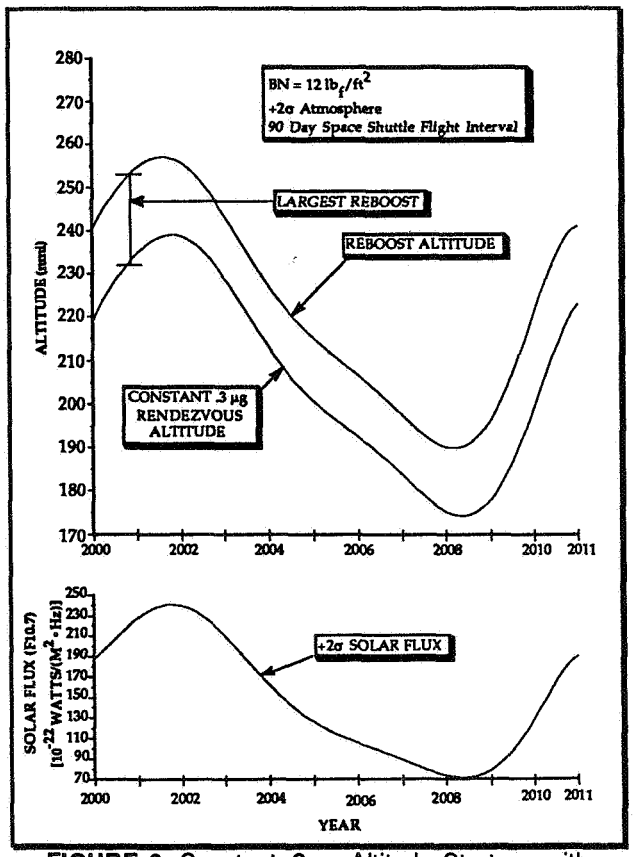


FIGURE 6: Constant .3 μg Altitude Strategy with Solar Flux Data

Constant .2 μg Level

JSC 31000 (Reference 5) baselined .2 μg as the new constant μg level. This change to a .2 μg level was made for several reasons. The .3 μg level placed SSF at altitudes that were determined to have unacceptable orbit lifetimes of less than 50 days to 150 nmi. At .3 μg, the SSF was flying at altitudes that were so low that the configurations necessary to maintain control were determined to be unacceptable. Figure 7 shows the altitude profile for a .2 μg altitude strategy. While this new strategy placed SSF above the ill-defined minimum controllable altitude and had acceptable flight configurations, it had its problems as well. As the BN of SSF got worse due to a more realistic analysis of its configuration, the .2 μg strategy put SSF at altitudes that violated radiation requirements. The strategists also realized that the emphasis of an altitude strategy was really directed towards system sizing. Therefore, a strategy was needed that forced the designs to meet the maximum requirements on the station's systems. That concept evolved into the lifetime variable altitude strategy.

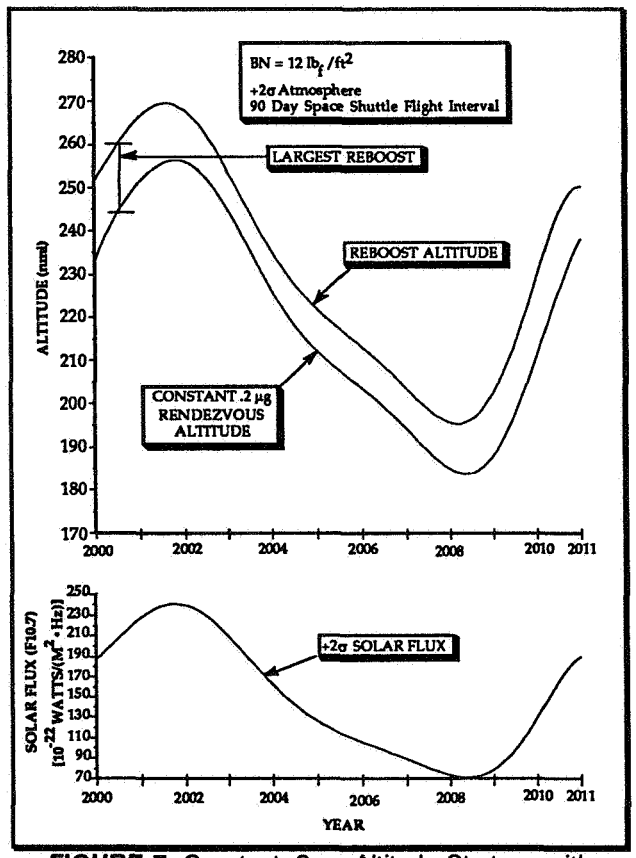


FIGURE 7: Constant .2 μg Altitude Strategy with Solar Flux Data

Advantages of a Constant μg Level Strategy

By taking advantage of the changing energy output of the sun, a constant μg strategy allows for increased payload-to-orbit during times of lower solar activity. This is a significant savings in terms of payload capability not only during assembly but also in terms of life cycle costs throughout the 30-year operational lifetime of SSF. Another advantage to this strategy is that it provides a guaranteed μg environment for payloads and experiments. This is a benefit to gravity sensitive experiments. However, as mentioned earlier, this premise can be misleading since the μg level will be at the guaranteed level only at the exact SSF CG.

Disadvantages of a Constant μg Level Strategy

The main disadvantage of the constant μg strategy is that the lifetime from the rendezvous altitudes varies significantly over a complete solar cycle. In light of the Challenger (51-L) accident, a lifetime reference has assumed greater importance for safety considerations. The need for defining a minimum operational period that SSF must be able to survive without a Space Shuttle visit was identified. Using a .3 μg level, the lifetime varied from 50 to 160 days of orbital lifetime to 150 nmi. A .2 μg level varied from 90 to 330 days of orbital lifetime to 150 nmi over a solar cycle. Although the .2 μg strategy altitudes were an improvement over the .3 μg level's lifetime, it was determined that several .2 μg altitudes had unacceptable orbit lifetime margins and thus the lifetime variable altitude strategy was developed.

Lifetime Altitude

It was determined that the altitude strategy needed at this point in the program must emphasize system design for SSF, thus defining the *minimum design* altitude. The strategies discussed thus far all recognized the concept of minimum orbit lifetime. Although the μg strategies were aware of a minimum controllable altitude, orbital lifetime to this altitude was the underlying concern. How much time does SSF have before it decays and enters the earth's atmosphere? What time frame would a contingency scenario require to rescue SSF after a catastrophic failure? This idea of the time necessary to save SSF resulted in an altitude strategy that defined the minimum design altitude (operational altitude limit) and based it on orbit lifetime. The strategy calls for SSF to be able to survive a total failure of its propulsion system for a period of at least 90 days without a loss of attitude control. It is felt that a rescue and repair mission for saving SSF could be made ready and launched within this amount of time. Therefore, the minimum design altitudes are defined as the altitudes which give SSF 90 days of lifetime to 150 nmi. The 150 nmi loss of control altitude was determined to be the altitude where the aerodynamic torques would overwhelm the attitude control system and a loss of control would result. The altitudes resulting from the lifetime altitude strategy can be seen in Figure 8.

Presently, the lifetime strategy is incorporated in the latest revision of JSC 31000 (Reference 3). At the time of this writing, a change request submitted in 1988 to the program (Reference 6) is awaiting final approval before acceptance into the *Space Station Program Definition and Requirements Document* (PDRD) (Reference 4), JSC 30000. However, the idea of

lifetime is now widely accepted throughout the SSF Program.

Advantages of the Lifetime Variable Altitude Strategy

The lifetime altitude strategy incorporates and corrects several of the ideas and problems identified in previous strategies. While the .3 μg strategy *mentioned* a minimum controllable altitude, the lifetime altitude strategy *defines* one. As in the μg altitude strategy, lifetime altitudes vary with the solar cycle, taking advantage of the changing energy output of the sun. Lifetime altitudes represent the lowest operating altitudes allowable, thus providing two very important parameters: maximum payload-to-orbit and system design requirements. As mentioned earlier, the minimum altitude represents the maximum requirement imposed on several SSF systems. Sizing to these altitudes provides for the most flexible operation of those systems.

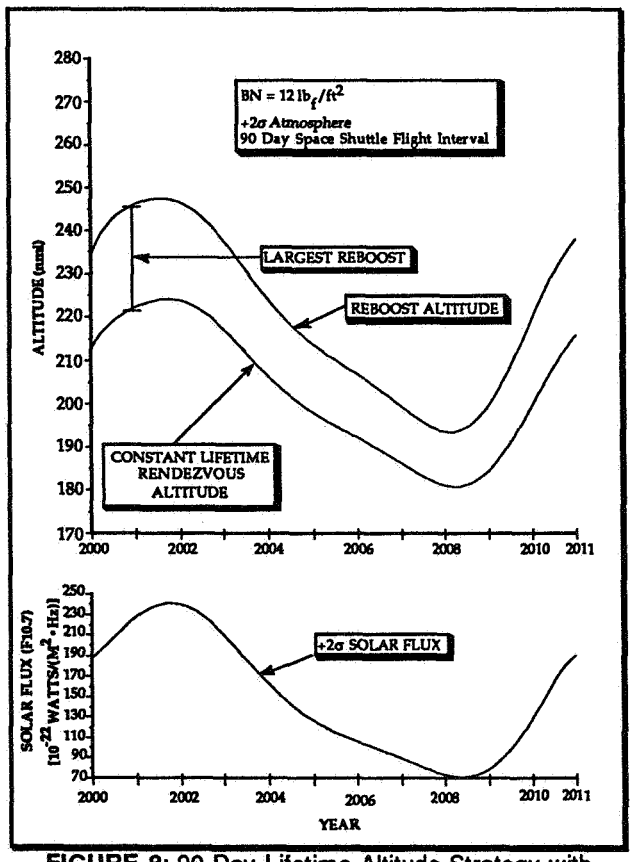


FIGURE 8: 90 Day Lifetime Altitude Strategy with Solar Flux Data

Disadvantages of the Lifetime Strategy

Operationally, lifetime altitudes provide a reference to safety. However, they are difficult to plan, since they are based on SSF characteristics and solar flux predictions which both may be updated as the actual flight date approaches. Additionally, the varying μg level experienced by flying lifetime altitudes presents difficulties to μg sensitive experimenters in both planning and data reduction.

IV CURRENT STRATEGY

The main purpose of the current lifetime altitude strategy was to identify and provide an altitude strategy that emphasized system design while providing adequate safety margins for SSF and its crew. The current altitude strategy defines an operational altitude envelope. Since this strategy reflects operational limits, systems are designed to the extremes of this range. The current lifetime altitude strategy levies requirements on the SSF Program and has been submitted for final approval into the SSF PDRD, JSC 30000.

> *SSF shall orbit between a minimum operational altitude, defined by operational lifetime to 150 nmi, and a maximum operational altitude, defined by radiation limits.* (Reference 6, Sec. 3.1.1.1)

Assembly

The minimum operational altitude is further divided into the two identifiable regions of SSF operations, assembly and post-AC.

> *The minimum operational altitude for assembly is defined as the altitude that provides 180 days of orbital decay to 150 nmi prior to a verified, dual fault tolerant reboost system, and 90 days of decay to 150 nmi thereafter. The decay shall be calculated using +2σ solar flux predictions.* (Reference 6, Sec 3.1.1.1.1)

Dual fault tolerance ensures that SSF has adequate lifetime prior to its propulsion system being completed. Once the system is dual fault tolerant, SSF can maintain the minimum altitude of 90 days to 150 nmi. While the present baseline assembly altitudes attempted to satisfy these minimum operational altitude requirements, they were chosen with more emphasis placed on planning simplicity. The present baseline assembly sequence is contained in the *Space Station Stage Summary Databook 12/15/89* (Reference 7). The Databook defines 220 nmi (407 km) as the rendezvous altitudes for flights 1 through 5 (MB-1 - MB-5) and 190 nmi (352 km) for flights 6 through 29 (MB-6 - L-11). These *defined* assembly rendezvous altitudes do not take advantage of the variations in the atmosphere but act as placeholders to simplify long range Space Shuttle manifest planning.

Figure 9 shows the Databook defined rendezvous altitudes and the minimum design rendezvous altitudes for the baseline assembly sequence. Recent studies have shown the 220 nmi altitudes to be conservative compared to the required 180 days to 150 nmi lifetime altitudes (~15 nmi, or 350 extra days of lifetime). However, the 220 nmi altitudes are being used as the *design-to* altitudes for Space Shuttle planning since historically it is easier to give capability back to the

program than to take it away. As the SSF design and launch schedule mature, these altitudes may be lowered to provide increased payload-to-orbit. Several lifetime violations occur at the transition from 220 nmi to 190 nmi altitudes. These violations stem from dramatic changes in SSF configuration (solar photovoltaic (PV] arrays are delivered and deployed on flight 6 [MB-6]). Lifetime violations also occur later in the sequence at flights 20 (MB-13) through flight 29 (L-11). These violations occur as a direct result of increasing solar activity. Figure 9 also depicts the solar flux values predicted for during assembly. Significant resistance to change the assembly altitudes exists in the Space Station program. However, these lifetime violations are significant and need to be addressed. The authors recommended to the Mission Planning and Analysis Division (MPAD NASA) at the Johnson Space Center (JSC), and to the Assembly Planning Review (APR), that the minimum lifetime altitudes be used as the planning altitudes for Space Shuttle manifest planning on those flights with lifetime violations. The long range planning altitudes and manifests must be reworked to correct these problems.

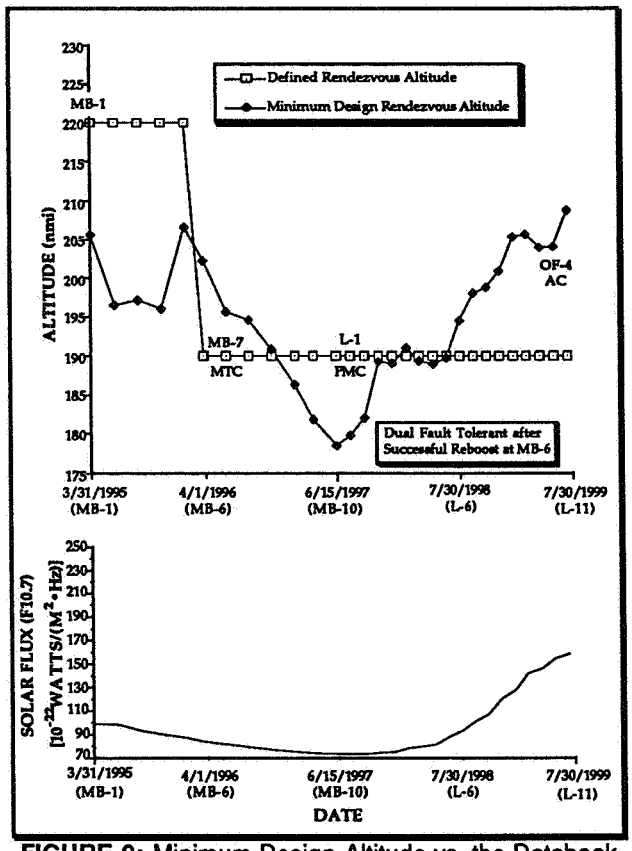


FIGURE 9: Minimum Design Altitude vs. the Databook Defined Altitudes with Solar Flux Data

Additional requirements must be met by the assembly altitudes that were not considered in the previously mentioned study:

> *SSF shall be capable of maintaining the minimum operational altitude under contingency atmosphere conditions. Contingency atmosphere conditions are induced increases in solar flux values due to the effects of a +2 sigma solar flux prediction combined with a 2 year first element launch slip and a 9 year solar cycle 22.* (Reference 6, Sec 3.1.1.1.1)

Analysis of the defined assembly altitudes with a 2-year first element launch (FEL) slip in conjunction with +2σ solar flux predictions show the 220 nmi altitude to have approximately the same conservative margin, although significantly greater lifetime violations occur at flight 8 (OF-1) through flight 29 (L-11). (See Figure 10.) Analysis results show that there are 11 additional lifetime violations when compared to the results of just the +2σ atmosphere study. These additional lifetime violations arise from moving the assembly sequence forward two years into a region of higher solar flux values. The shift in the solar cycle is clearly shown in Figure 10 when compared with that in Figure 9.

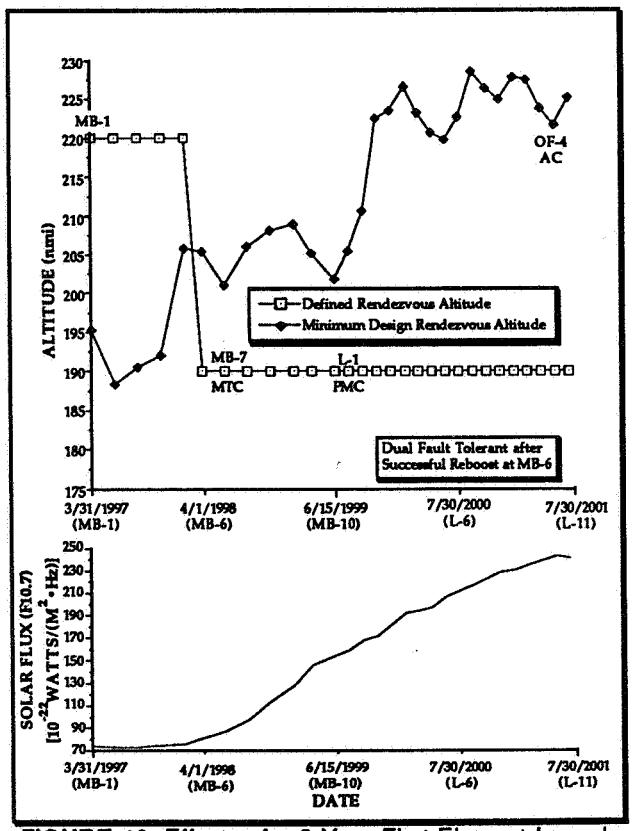


FIGURE 10: Effects of a 2-Year First Element Launch Slip on the Minimum Design Altitudes Showing the Solar Cycle Shift

The requirement for a 9-year solar cycle has not yet been analyzed at the time of this publication. However, this only increases the solar flux values towards the end of assembly since the solar minimum would shift back toward the FEL point. A 9-year solar cycle analysis would result in lifetime violations of greater

magnitude due to the increased solar flux values. It needs to be determined if the two year FEL slip and 9-year solar cycle are necessary biases, and, if so, to fly the minimum design altitudes in place of defined altitudes for each assembly flight which violate the lifetime requirement.

Assembly Complete

The altitude requirements for assembly complete (AC) are slightly different than those for assembly.

> *The minimum operational altitude for assembly complete shall be the altitude that provides 90 days of orbital decay to 150 nmi using the most current solar flux values.* (Reference 6, Sec 3.1.1.1.2)

The lifetime altitude defines the minimum altitude that the SSF can fly during post-AC. Presently, the post-AC rendezvous altitudes are set at this minimum operational altitude. The SSF will rendezvous at this altitude and SSF will reboost as soon as operationally possible after the Space Shuttle departure. SSF will reboost to an upper altitude such that it decays to the next scheduled Space Shuttle rendezvous altitude.

System Design

The present altitude strategy has been a major driver for determining system design related requirements. The current lifetime altitude strategy represents the absolute minimum allowable altitudes for SSF operations. This will cause SSF to operate in a more dense region of the atmosphere, resulting in maximum reboosts and aerodynamic torques. Designing systems to meet the needs of SSF at this minimum altitude will ensure the most flexible system design capable of surviving a variety of real world operational contingencies.

Several influences on SSF system design are incorporated into the lifetime strategy. For example, the rendezvous altitudes follow the solar activity. This takes advantage of the solar cycle not only in terms of increasing Space Shuttle payload-to-orbit, but also in varying the SSF reboost. The current altitude strategy is employed to support design studies and size the propulsion system, specifically, the size of the on-board propellant storage tanks. In a recent trade study it was determined that the propellant tanks could be sized to reduce life cycle costs. This would be accomplished by using the largest tanks that would fit in the cargo bay of the Space Shuttle as well as fulfill the requirements for the largest reboost. The maximum reboost occurs during the rise in the solar cycle curve, six to eight months prior to peak solar activity. The propulsion system has been designed to the maximum reboost of a $+2\sigma$ atmosphere at SSF maturity, since SSF mass has a direct influence on the amount of propellant required for reboost.

Propellant specific impulse (I_{sp}) also has an influence on propulsion system design; the lower the I_{sp}, the more propellant required for a particular reboost. As a result of the SSF *Scrub* reconfiguration effort (1989), mono-methyl hydrazine propellant (I_{sp} = 230 sec) propulsion system was baselined for SSF. This results in larger propellant requirements and higher life cycle costs than for the previously baselined hydrogen/oxygen propellant (I_{sp} = 370 sec for H_2/O_2).

The BN affects system sizing as well. The smaller the BN, the greater the orbital decay, and thus larger reboosts are required to decay to the same point.

Other systems are indirectly influenced by the altitude selection. The electrical power system batteries are sized to provide power during orbital nighttime, a function of altitude. Some operations are also influenced by altitude: extra-vehicular activity (EVA) planning flexibility goes down as the South Atlantic Anomaly grows in size (which grows larger with increasing altitude). Many payloads are sensitive to the μg environment induced by altitude choice. Finally, altitude selection plays a key role in the utilization efficiency of the delivery system (i.e., Space Shuttle). Decreasing Space Shuttle payload-to-orbit capability by rendezvousing higher may cause carriers to be manifested at less than 100% capacity. This indirectly affects all SSF operations since Space Shuttle will deliver all resupply requirements for SSF operations.

Operationally, the SSF must <u>never</u> violate the lifetime altitude limit. While SSF systems have been designed to operate at or above this altitude limit, the lifetime altitude represents the point where a deviation from nominal operations must occur. An operational altitude strategy will need to provide an altitude safety margin based on possible operational deviations. These deviations could include such scenarios as a Space Shuttle launch slip, missed rendezvous, atmosphere worsening, or solar cycle phase shifting. While the lifetime altitude strategy provides design-to altitudes for system sizing with inherent concerns for SSF safety, an operational altitude strategy must provide an additional lifetime margin to allow for the unpredictable yet expected real world occurrences. The operational altitude strategy must also incorporate several other operational considerations as well. These considerations as well as the factors that influence the lifetime buffer will be discussed in the next section.

V OPERATIONAL ALTITUDE STRATEGY

The emphasis of an Operational Altitude Strategy (OAS) is overall operations cost, whereas the emphasis of the current lifetime altitude strategy is system design. Development of an OAS must consider all aspects of SSF operations:

- On-orbit operations
  - aerodynamic disturbance levels
  - radiation exposure levels
  - contamination
  - orbital debris density
  - safety (lifetime)
  - satellite servicing

- Logistics system operations
  - delivery system (Space Shuttle) utilization
  - logistics elements utilization

- Space Shuttle mission planning
  - launch windows
  - flight profile/rendezvous
  - planning simplicity
  - payload-to-orbit capability

Developing an altitude strategy based on any one of the influences listed above would result in off-nominal altitudes for all the other influences. Therefore, one influence must be considered predominant and altitudes biased off this solution to accommodate the remaining influences. Program requirements dictate that the operational altitudes may be chosen anywhere between an upper and lower altitude limit. The upper altitude bound is defined by radiation concerns and the lower limit is defined by orbit lifetime. Flying SSF as low as possible would maximize the Space Shuttle delivery capability. However, this does not necessarily represent the most efficient use of the delivery system.

Given acceptable safety levels, delivery system utilization efficiency (in the opinion of the authors) should be considered the primary performance indicator of any OAS. The Department of Defense estimates that 35-40% of the total operational cost for the military is logistics. Delivery of resupply from the ground to SSF is a significant part of the overall logistics cost and represents an area in which operations cost could be significantly reduced through careful application of an OAS.

Optimal Altitudes

Eventually, all operational influences will be accommodated through biasing **optimal** altitudes to the extent required. Initially however, optimal altitudes must be investigated and understood.

In order to reduce resupply costs, optimal altitudes must maximize Space Shuttle delivery capability. As part of resupply, propellant must be delivered. However, propellant usage is tied directly to altitude selection. Increasing rendezvous altitudes results in decreased propellant requirements (reboosts are smaller since the SSF operates in a less dense region of the atmosphere), but at the same time, the Space Shuttle delivery capabilities are also reduced. The optimal altitude is therefore a balance between SSF propellant usage and Space Shuttle delivery capability and can be defined as maximizing the **net** Space Shuttle delivery capability on a flight by flight basis (**net** = total Space Shuttle delivery capability - SSF reboost propellant requirements).

The reboost propellant requirements must include support hardware (tankage and attachments) as part of the **net** payload-to-orbit determination. Final propellant selection will greatly influence the support mass required. For example, a hydrogen/oxygen propulsion system has relatively small support mass requirements since propellant is delivered as water in either a simple water tank on a fluids carrier or scavenged from the Space Shuttle fuel cell tanks (no support mass requirement). On the other hand, hydrazine has large support mass requirements. Since hydrazine is a very volatile substance, on-orbit disconnections of hydrazine fluid lines are restricted. An entire Propellant Module (PM) (consisting of propellant, propellant tanks, reaction control system [RCS] thrusters and structure) must be exchanged each time propellant is delivered.

Propellant deliveries do not necessarily occur at regular intervals. Generating optimal altitudes based on specific flight manifests of PM delivery results in a very jagged rendezvous altitude profile and may force either undesirably large reboosts or rendezvous altitudes which would require a SSF deboost to achieve. For this reason, the required support mass is evenly distributed over all flights for purposes of determining a smoother optimal altitude profile. In order to accomplish this distribution, each pound of propellant needed to perform a reboost will require some amount of support mass. The ratio of support mass to propellant is termed the mass fraction and is different for each type of propellant or PM. The currently designed PM requires approximately 0.7 lbm of supporting hardware for each pound of hydrazine delivered.

Optimal Altitude Influences

Optimal altitudes are driven by propellant requirements which in turn are driven by the SSF BN, SSF mass, Space Shuttle flight interval, I_{sp}, solar flux ($F_{10.7}$) predictions, and propellant mass fraction. In general, larger propellant requirements result in higher optimal altitudes.

Currently, working values for each of these influences are baselined within the SSF Program. However, the actual value may turn out to be considerably different. An understanding of how each of these influences drives the optimal altitudes and how sensitive optimal altitudes are to these influences is essential for OAS development.

SSF Ballistic Number

BN characterizes the aerodynamic configuration and weight of the SSF. Low BNs result in high decay rates, while high BNs imply low decay rates. Lower BNs result in higher propellant usage since SSF decay is greater. Figure 11 shows the optimal altitude sensitivity to BN changes. This figure plots altitude (x-axis) vs Space Shuttle delivery capability. The slanted line at the top represents the total Space Shuttle lift capacity. As the rendezvous altitude increases, the total Space Shuttle delivery capacity decreases. The curved lines indicate the **net** payload-to-orbit for three values of BN. The optimal altitude occurs at the highest point on each curve (the maximum **net** payload-to-orbit) and is represented with a Δ on the graph. Since **net** payload-to-orbit is defined as the total Space Shuttle delivery capacity minus the SSF reboost propellant requirements, the vertical distance from the total Space Shuttle lift capacity line to the **net** payload-to-orbit line

represents reboost propellant requirements (includes both propellant and supporting hardware). The figure shows how the optimal altitudes increase with a decrease in the BN. Additionally, rendezvous altitudes derived from the .2 μg, .3 μg and lifetime altitude strategies are also presented for comparison. On each of the parametric plots which follow, the standard case is identified by a bold line. The conditions for this standard case are:

BN =	12 lb_f/ft^2
SSF Mass =	500000 lbm
Flight Interval =	90 days
I_{sp}=	230 sec
Solar Flux =	Maximum (+2σ peak)

and can be used as reference between the various influences.

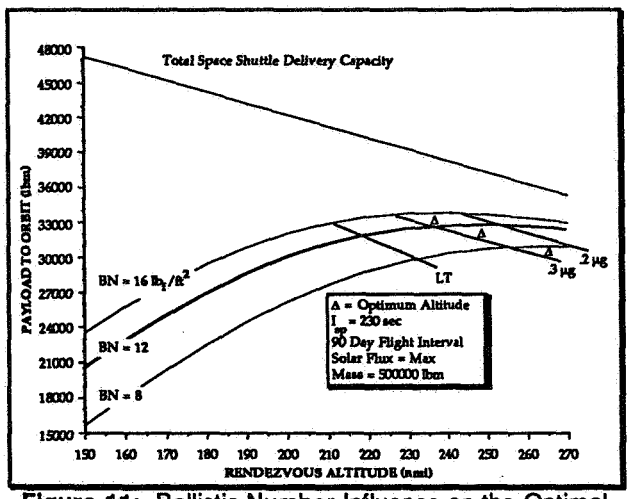


Figure 11: Ballistic Number Influence on the Optimal Altitude

SSF Mass

The more massive the SSF, the more propellant required for reboost. Although continuous low thrust burns will actually be used to reboost SSF, the propellant required to reboost SSF impulsively (infinite thrust in an instant of time) is very close (< 1%) to the propellant required using a continuous low thrust burn and is given by the following relationship:

$$\text{Prop} = \text{SSF Mass} * G_\theta * (1 - \exp(-\Delta V / (G_\theta * Isp))) \quad (3)$$

where G_θ is the acceleration of gravity at the earth's surface, ΔV is the velocity change required to achieve a circular target orbit based on the height of the reboost, and I_{sp} is the propellant specific impulse.

Figure 12 shows the optimal altitude sensitivity to SSF mass. The optimal altitude increases with mass since propellant requirements are proportional to SSF mass. Altitudes based on the .2 μg, .3 μg and lifetime altitude strategies are not affected by SSF mass changes.

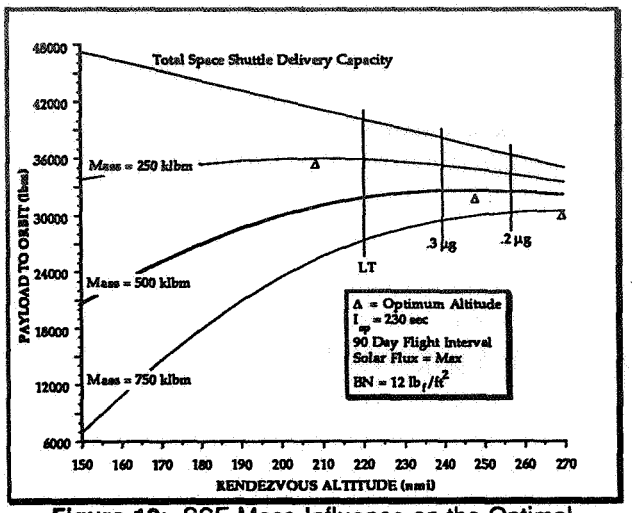


Figure 12: SSF Mass Influence on the Optimal Altitude

Space Shuttle Flight Interval

Operationally, SSF is assumed to reboost soon after Space Shuttle departure. The size of the reboost is determined such that the SSF orbit will decay down to the appropriate rendezvous altitude by the next Space Shuttle visit. A longer flight interval results in a longer decay time and consequently larger reboosts. Figure 13 shows the optimal altitude sensitivity to Space Shuttle flight interval. The optimal altitude increases with flight interval. Rendezvous altitudes based on the .2 μg, .3 μg and lifetime altitude strategies are not affected by flight interval changes.

Multiple reboosts between Space Shuttle visits tend to decrease the size of each reboost, yet increase the total propellant requirements for the interval. This is a direct result of SSF spending more time at lower altitudes where the atmosphere is more dense. Multiple reboosts within a flight interval will increase the optimal altitudes as well as the operational work loads for the SSF crew and mission support teams. Therefore, a single reboost is assumed between each rendezvous.

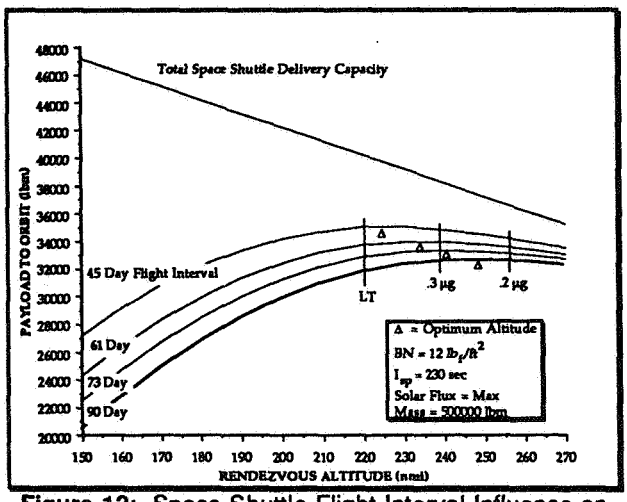


Figure 13: Space Shuttle Flight Interval Influence on the Optimal Altitude

SSF Specific Impulse

The I_{sp} characterizes the efficiency of a propellant. Less propellant is required to achieve the same reboost for propellants having higher I_{sp}. Therefore, lower I_{sp} propellants increase the optimal altitude since propellant requirements are higher. For example, hydrogen/oxygen propellant has an I_{sp} of 370 sec, and hydrazine has an I_{sp} of 230 sec. This translates into thousands of additional pounds of hydrazine per year as compared to water. As expected, the optimal altitude for an I_{sp} of 230 sec is considerably higher than for an I_{sp} of 370 sec. Reboost propellant for a given I_{sp} is shown in equation (3).

Figure 14 shows the optimal altitude sensitivity to I_{sp}. The optimal altitude increases as I_{sp} decreases. Altitudes based on the .2 μg, .3 μg and lifetime altitude strategies are not affected by I_{sp} changes.

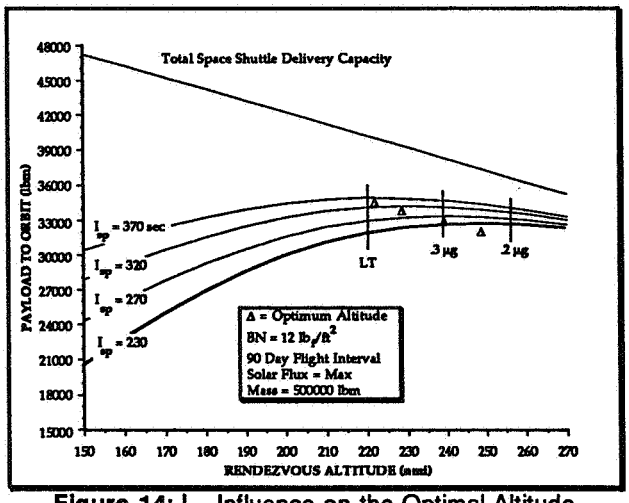


Figure 14: I_{sp} Influence on the Optimal Altitude

Solar Flux ($F_{10.7}$)

Atmospheric density is derived from $F_{10.7}$ values. Larger $F_{10.7}$ values increase the derived density and consequently increase the rate of orbital decay. $F_{10.7}$ values change with date and atmospheric predictions (+2σ or mean). If all influences discussed above were to remain constant, optimal altitudes would vary with the solar cycle.

Figure 15 shows the optimal altitude sensitivity to $F_{10.7}$. The optimal altitude increases with increasing $F_{10.7}$. Altitudes based on the .2 μg, .3 μg and lifetime altitude strategies also vary with $F_{10.7}$ changes.

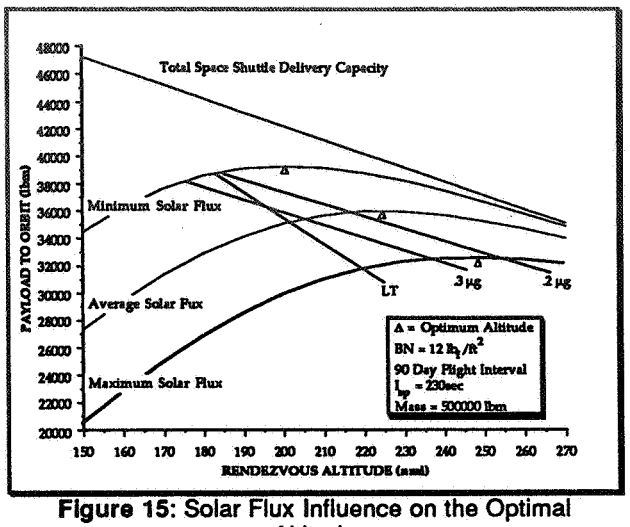


Figure 15: Solar Flux Influence on the Optimal Altitude

As shown in Figures 11 - 15, the **net** payload-to-orbit curves are relatively flat near the optimal points. This indicates that altitude biases due to other operational considerations could be accommodated without significantly impacting **net** payload-to-orbit. In fact, assuming average values for each influence, the rendezvous altitude could vary by +/- 10 nmi with a very small variance in the **net** payload-to-orbit. Within this 20 nmi band around the optimal altitude, less than a 500 lbm **net** payload loss is experienced. Given the choice of flying 10 nmi higher or 10 nmi lower, rendezvous altitudes would be adjusted upward, if possible, to increase SSF lifetime.

Reaction to Unpredictable Events

An OAS must specify (or levy requirements on SSF operational concepts) when and how the SSF reacts to real time changes which drive altitude selection, for example, Space Shuttle launch delays. Operationally, SSF lifetime should never drop below the minimum lifetime level of 90 days to 150 nmi. Depending on the reasons for the slip, SSF may react before ever reaching this limit.

Additionally, real time changes in the solar flux could result in current altitudes with undesirable lifetimes or which are significantly off-optimal altitudes. Current thinking suggests that the operational altitudes will be determined using **mean** solar flux predictions. As can be seen in Table 1, recently observed solar flux data are appreciably greater than the mean predictions for those dates (Reference 8). Again, how and when should SSF react to a situation which significantly deviates from the predicted conditions?

TABLE 1: Predicted vs. Observed Solar Flux Values

DATE	MEAN	OBSERVED
Dec. '88	113.2	199.8
Jan. '89	118.4	235.4
Feb '89	120.7	222.4
Mar. '89	123.1	205.1
Apr. '89	125.9	189.6
May '89	129.2	190.1
June '89	132.4	239.6
July '89	135.2	181.9
Aug. '89	137.5	217.1
Sep. '89	139.3	225.9
Oct. '89	140.8	208.7
Nov. '89	142.5	235.1
Dec. '89	144.1	213.0
Jan. '90	145.3	210.1
Feb. '90	146.4	178.3

What circumstances (SSF system failures, Space Shuttle launch slips, atmosphere worsening, etc.) will cause SSF to adjust altitude and depart from nominally planned profiles? How should SSF maintain altitude under contingency conditions? Such issues must be addressed in an OAS.

VI CONCLUSIONS

Defining an operational altitude strategy poses significant challenges by attempting to incorporate and blend numerous operational considerations. Only recently has the SSF Program reached a maturity level to where altitude strategy planners could begin assimilating and integrating many of the influences and considerations necessary to develop an operational altitude strategy. In addition to providing the guidelines for computing nominal operational altitudes for long range and near *real time* planning, an operational altitude strategy must also be adaptive, providing procedural road maps as to how and when SSF must react to real time off-nominal conditions, such as, Space Shuttle launch slips and unexpected deviations in atmospheric parameters.

It is important to note that any altitude strategy ultimately proposed will inherently favor certain operational aspects over others. This paper has not attempted to produce a final altitude strategy to be used for the 30-year operational lifetime of SSF; rather it attempts to identify and put into perspective the associated issues and influences which will drive the development of a final operational altitude strategy.

Realistically, this process will take years of discussion and prioritization by system, element, and operations areas before mission planners are "smart" enough to implement any strategy. Unfortunately, such a strategy could be of great use in the near term to efficiently design logistics elements and estimate user and SSF core support requirements. The authors of this paper feel that ***optimal*** altitudes are an excellent first cut at an operational altitude strategy providing considerable flexibility to accommodate future operational considerations as they become pertinent.

VII ACKNOWLEDGEMENTS

The authors would like to recognize and express their appreciation to Lauri Norton of MPAD at Johnson Space Center for her input and guidance on this paper as well as all of their SSF altitude related endeavors.

VIII REFERENCES

1) Jacchia, L. G., *New Static Models of Thermosphere and Exosphere with Empirical Temperature Profiles.* Smithsonian Astrophysical Observatory Special Report 313. May 6, 1970.

2) *Solar Activity Inputs for Upper Atmospheric Models Used in Programs to Estimate Spacecraft Orbital Lifetime,* EL21 (88-41). Marshall Space Flight Center. . October 19, 1988.

3) *JSC 31000, Space Station Projects Description and Requirements Document,* Revision E, Volume 3, "Project Design and Development Requirements", Section 3.1.1. November 1, 1989.

4) *JSC 30000, Space Station Program Definition and Requirements Document.* October 15, 1985.

5) *JSC 31000,* Revision C, Section 3, "System Requirements", Paragraph 2.1.1, "Orbit", pp. 3-1 - 3-2. March 6, 1987.

6) *JSC 30000,* Section 3.1.1, "Orbit". Change Request #BB-000-521. 1988.

7) *Space Station Stage Summary Databook.* Space Station Freedom Program Office. December 15, 1989.

8) *Solar Activity Inputs for Upper Atmospheric Models Used in Programs to Estimate Spacecraft Orbital Lifetime,* EO41 (90-17). Marshall Space Flight Center.. April 5, 1990.

FLIGHT MECHANICS/ESTIMATION THEORY SYMPOSIUM

MAY 22-24, 1990

SESSION 3

THE ACCURACY OF DYNAMIC ATTITUDE PROPAGATION*

E. Harvie and D. Chu
Computer Sciences Corporation (CSC)

M. Woodard
Goddard Space Flight Center

ABSTRACT

Propagating attitude by integrating Euler's equation for rigid body motion has long been suggested for the Earth Radiation Budget Satellite (ERBS) but until now has not been implemented. Because of limited Sun visibility, propagation is necessary for yaw determination. With the deterioration of the gyros, dynamic propagation has become more attractive. Angular rates are derived from integrating Euler's equation with a stepsize of 1 second, using torques computed from telemetered control system data. The environmental torque model was quite basic. It included gravity gradient and unshadowed aerodynamic torques. Knowledge of control torques is critical to the accuracy of dynamic modeling. Due to their coarseness and sparsity, control actuator telemetry were smoothed before integration.

The dynamic model was incorporated into existing ERBS attitude determination software. Modeled rates were then used for attitude propagation in the standard ERBS fine-attitude algorithm. In spite of the simplicity of the approach, the dynamically propagated attitude matched the attitude propagated with good gyros well for roll and yaw but diverged up to 3 degrees for pitch because of the very low resolution in pitch momentum wheel telemetry. When control anomalies significantly perturb the nominal attitude, the effect of telemetry granularity is reduced and the dynamically propagated attitudes are accurate on all three axes.

*This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, Contract NAS 5-31500.

INTRODUCTION

This paper is based on work performed in the Flight Dynamics Division by the Earth Radiation Budget Satellite (ERBS) attitude support team. It describes the development and evaluation of a model of the ERBS attitude dynamics, based on direct integration of Euler's equation of motion, using telemetered control system information and estimated disturbance torques. Knowledge of the ERBS dynamics is important for accurately propagating the attitude over an entire orbit between occurrences of Sun data. The ERBS fine attitude determination system (FADS) was designed to use onboard inertial reference unit (IRU) measurements of spacecraft motion as a model for attitude motion. With the deterioration of the IRUs, an alternative method for deriving spacecraft body angular rates is necessary to utilize the existing FADS software and retain the fine attitude determination capability. Comparison of FADS attitudes propagated using rates from this alternative method to those using valid IRU rates provides a good assessment of the accuracy of the dynamic model.

The dynamic model presented here exploits available telemetered attitude control system data for computation of the control torques. However, since the control system engineering data were intended for spacecraft health and safety monitoring and trend analysis, the data have coarse digitization and are less frequent than is necessary for an accurate dynamics determination. Extensive preprocessing of the data is, therefore, required to approximate real control actuator behavior. The low quality of the telemetered control system data dominates the effects on solution accuracy of unmodeled disturbance torques, numerical integration error, and uncertainties in spacecraft mass properties.

The dynamic model is integrated into the Data Adjuster subsystem of the ERBS Attitude Determination System (ADS) and takes advantage of existing subroutines for data processing. This method of implementing the model is extremely efficient by eliminating redundant software development and allows the autonomy of the ADS to be preserved. Previous attempts at dynamic modeling have coupled the solution of the dynamics with complex state estimators, which solved for unmodeled parameters and propagated the attitude in one algorithm (References 1 through 3). In this work, every effort was made to keep both the mathematical detail and the solution method as simple as possible. The model demonstrates the feasibility of applying a simplified dynamics determination to real spacecraft data.

ERBS OVERVIEW

The ERBS was launched in October 1984. It is in a 600-kilometer, near circular orbit of 57-degree (deg) inclination. The ERBS is an angular momentum biased spacecraft, with attitude referenced to a geodetic coordinate system; yaw is defined about the local unit nadir vector (+z body axis); pitch is defined about negative orbit normal (+y body axis); and roll is defined about the remaining right-handed orthogonal unit vector. Pitch is controlled to within 1 deg of geodetic null with a y-axis angular momentum wheel using an analog control loop. Roll is controlled to within 1 deg of null by a pair of y-axis mounted electromagnets. Yaw is controlled to within 2 deg of null through roll/yaw kinematic coupling with electromagnets and active yaw angular momentum control by a pair of differentially driven ITHACO Scanwheels, mounted in the y-z plane. The Infrared (IR) Scanwheels, which measure geodetic pitch and roll, also serve as the sole input to the pitch and roll/yaw control loops. Figure 1 illustrates the ERBS body coordinate system and attitude actuator hardware. Table 1 summarizes the ERBS orbit and attitude characteristics. The remaining components of the Magnetic Control System (MCS) include a three-axis Schoenstedt magnetometer for electromagnet control input and attitude determination, one roll axis and one yaw axis electromagnet dipole torque rods for pitch momentum management, and the associated control electronics. Besides the Scanwheels and the magnetometer, the remaining attitude determination sensors are a pair of ADCOLE two-axis fine Sun sensors and two redundant three-axis Northrop IRUs (Reference 4).

ERBS ADS OVERVIEW

The ERBS Attitude Determination System computes single-frame coarse attitudes to within 5 deg, fine attitudes to within 0.25 deg, attitude rates to within 0.005 deg/second, and monitors control system

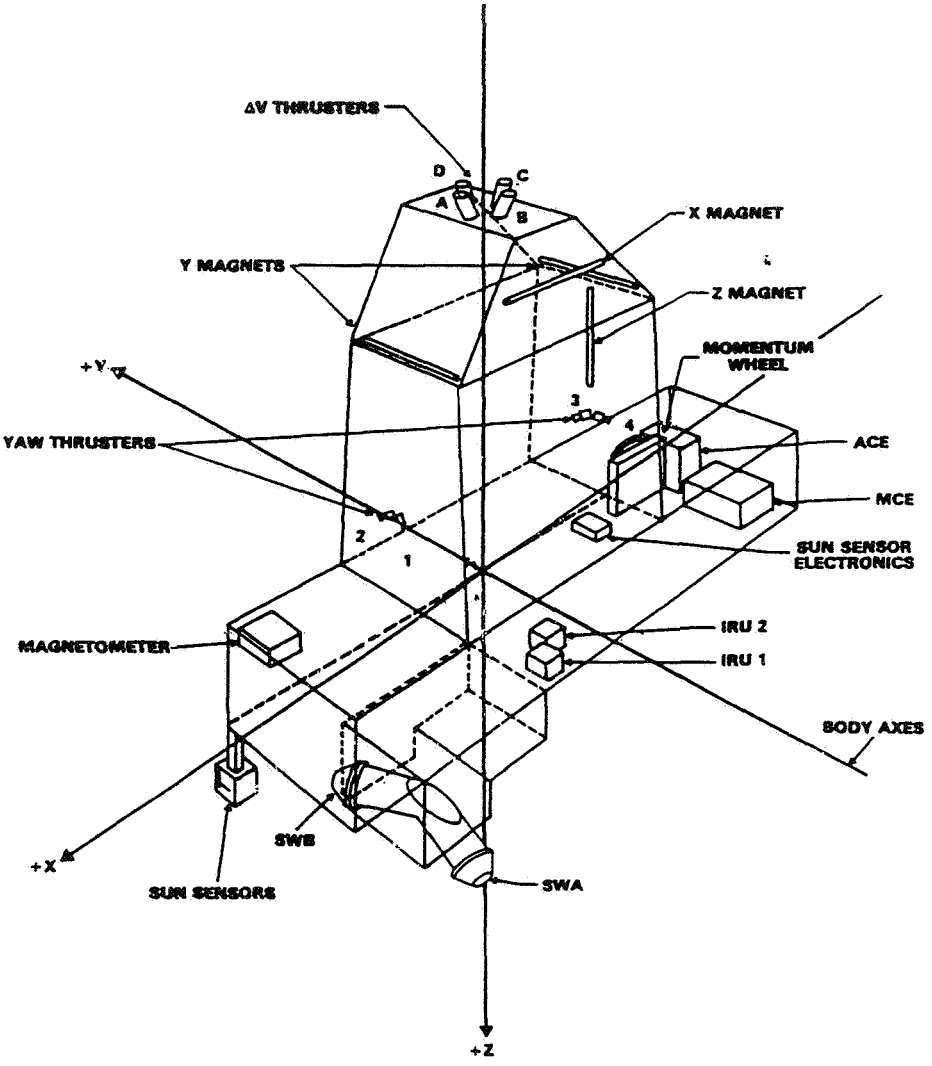


Figure 1. ERBS Attitude Hardware and Body Coordinate System

performance (Reference 5). Attitude telemetry data are unpacked and converted to engineering units in the Telemetry Processor. The Data Adjuster subsystem applies sensor biases and misalignment corrections, generates reference vectors from ephemerides, and smooths raw data on option. The Coarse Attitude Determination System (CADS) computes single-frame attitudes using the Quaternion Estimator (QUEST) algorithm. QUEST requires no dynamics information for attitude propagation and has inherent accuracy much greater than the specified accuracy of 5 deg when accurate observations (Sun data) are available. The FADS uses IRU measurements as a mechanical motion model for attitude propagation to model sensor data. It then applies a batch least-squares optimization of observed to modeled sensor data to estimate a state vector of epochal attitude angles and gyro calibration parameters. The FADS is very sensitive to the quality of the input rate data.

Table 1. ERBS Attitude and Orbit Characteristics

ORBIT:

SEMIMAJOR AXIS: 6891 km

INCLINATION: 57 deg

ECCENTRICITY: 0.0014 (NEAR-FROZEN ORBIT)

ATTITUDE PARAMETERS:

ANGULAR MOMENTUM BIASED, EARTH ORIENTED, 1 REVOLUTION PER ORBIT

NOMINAL GEODETIC PITCH AND ROLL = 0.0 deg

NOMINAL YAW = 0.0 OR 180.0 deg FOR SOLAR ARRAY ILLUMINATION

SPACECRAFT PRINCIPAL MOMENTS OF INERTIA: I_x = 2860 kg-m^2
I_y = 4590 kg-m^2
I_z = 3040 kg-m^2

ATTITUDE SENSORS:

TWO ADCOLE FINE SUN SENSORS 64x64 deg 0.004 deg (l.s.b.) @ 1 Hz

TWO ITHACO SCANWHEEL IR SENSORS 0.025 deg (l.s.b.) @ 1 Hz

ONE SCHOENSTEDT THREE-AXIS FLUXGATE MAGNETOMETER 4.68 mg (l.s.b.) @ 1/8 Hz

TWO IRUs WITH THREE NORTHROP RATE GYROS 0.001 deg/sec (l.s.b.) @ 1 Hz

MAGNETIC CONTROL SYSTEM ACTUATORS:

ONE PITCH MOMENTUM WHEEL 50 rpm (l.s.b.) @ 1/8 Hz, I_{MW} = 0.0881/kg-m^2

TWO ITHACO SCANWHEELS 12 rpm (l.s.b.) @ 1/8 Hz, I_{SW} = 0.0271 kg-m^2

ONE ROLL AXIS AND ONE YAW AXIS, 50-AMPERE TURN METER SQUARED (ATm^2) MAGNETIC DIPOLE TORQUE RODS FOR PITCH MOMENTUM CONTROL
4.7 ATm^2 (l.s.b.) @ 1/16 Hz

TWO PITCH AXIS 50 ATm^2 DIPOLE TORQUE RODS FOR ROLL CONTROL
9.4 ATm^2 (l.s.b.) @ 1/16 Hz

5848G(B)-1

MATHEMATICAL BACKGROUND

Euler's equation for rigid body rotation, including the contribution of internal wheel angular momenta, $\vec{h}(t)$, is written as

$$[I]\frac{d\vec{\omega}(t)}{dt} = \vec{N}(t) - \frac{d\vec{h}(t)}{dt} - \vec{\omega}(t) \times \{[I]\vec{\omega}(t) + \vec{h}(t)\} \qquad (1)$$

where $\vec{\omega}(t)$ is the body angular velocity vector, [I] is the body moment-of-inertia tensor, and $\vec{N}(t)$ is the sum total external torque acting on the body. This equation is solved for the angular velocity state, $\vec{\omega}(t)$ in body coordinates. For the ERBS, $\vec{h}(t)$ is the sum of the Scanwheel momenta ($\vec{h}_{SW_A}$ and $\vec{h}_{SW_B}$) and momentum wheel angular momentum ($\vec{h}_{MW}$) resolved in body coordinates:

$$\vec{h}(t) = \vec{h}_{MW} + \vec{h}_{SW_A} + \vec{h}_{SW_B} \qquad (2)$$

The Scanwheels are differentially driven ($\omega_{SW_A} + \omega_{SW_B}$ = constant) so that they produce no y-axis component of torque. The momentum wheel exerts very strong control over the pitch axis. The $d\vec{h}(t)/dt$ term in Equation (1) is approximated by the difference in wheel speed over the sampling interval.

The sum total external torque, $\vec{N}(t)$ in Equation (1), is approximated as the contributions from nonwheel control torques (magnetic dipole torques, $\vec{N}_{MAG}$) and environmental torques (gravity gradient, $\vec{N}_{GG}$, and aerodynamic, $\vec{N}_{AERO}$)

$$\vec{N}(t) = \vec{N}_{MAG} + \vec{N}_{GG} + \vec{N}_{AERO} \tag{3}$$

Solar radiation pressure is not included; it is expected to be one order of magnitude less than gravity gradient or aerodynamic forces, and its variability and attitude dependence over an orbit would unnecessarily complicate the environmental torque model. Also excluded is the residual dipole moment torque that is easily modeled as a constant torque bias, but it was never determined for the ERBS mission (Reference 6). The disturbance torques arising from science instrument activity are also neglected, and an estimate of their magnitude is not available. Both the residual dipole moment and the science instrument torques are assumed to be small. The magnitude of the gravity gradient torque is a maximum of 1×10^{-4} N-m in roll for a 1-deg offset from null attitude and two orders of magnitude less in pitch and yaw. The aerodynamic torque is important only for pitch and can be as high as 1×10^{-4} N-m.

It should be noted that the gravity gradient and aerodynamic torques are low frequency phenomena that are observable only over long timespans (on the order of one orbit), while the magnetic dipole and wheel control torques are high frequency components of the spacecraft dynamics.

Equation (1) is integrated as a system of three coupled equations for each component of $\vec{\omega}(t)$ using a second-order Runge-Kutta method, with a fixed stepsize corresponding to the data discretization interval of 1 sec. The stepsize is made as small as possible to keep the frequency response of the modeled torques high relative to the actual dynamical response of the spacecraft. The coarseness of the telemetry data and uncertainty in disturbance torques and mass properties preclude the need for a higher order numerical integration method.

The integration for each component of $\vec{\omega}(t)$ is begun at the nominal values of zero for roll and yaw, and the one revolution per orbit rate of 0.062 deg/sec for pitch. It is possible to start the integration from an anomalous state relative to the orbital frame only if *a priori* estimates of the body rates are available.

ERBS TELEMETRY DATA

The data required to compute the state of the ERBS dynamics consist of Scanwheel and momentum wheel speeds, magnetic dipole moments, sensed magnetic field, and IR scanner pitch and roll angles for use in the gravity gradient and aerodynamic torque expressions. In addition, Sun sensor data are used for attitude determination, and the IRU rates provide the reference for evaluation of the dynamic model.

Reconstruction of the ERBS attitude dynamics from control system telemetry is complicated by the coarseness and sparsity of the telemetry data from the principal control actuators. Ideally, the telemetered data frequency should be high relative to the characteristic frequency of the actuators. The resolution in magnitude should also be high relative to typical trends in the data. The pitch momentum wheel telemetry satisfies neither of these criteria.

The ERBS pitch axis control is dominated by the momentum wheel. Momentum wheel telemetry conversion is nonlinear and the least significant bit at nominal speed (approximately 2500 revolutions per minute (rpm)) is approximately 50 rpm over an 8-sec interval. This corresponds to a resolvable change in wheel angular momentum of about 0.46 N-m-sec over 8 sec. Since the resolution in onboard deliverable momentum wheel torque is 4×10^{-4} N-m (Reference 4), observability of the momentum wheel control torque in the downlink telemetry is very poor and limits accurate modeling of the pitch-axis dynamics. The maximum deliverable torque of the momentum wheel is 0.054 N-m so that more than one time interval is necessary to resolve even the maximum rate of change in wheel speed.

Yaw axis control is dominated by the differentially driven Scanwheels. Resolution in the Scanwheel speed is much better, about 12 rpm at nominal speed (approximately 2000 rpm). Because the Scanwheel

moments of inertia are smaller than the momentum wheel moment and because the component of Scanwheel yaw control is small (sine 10 deg), observability of the yaw axis control torque from telemetry is good.

Magnetic dipoles are used for roll axis control and for pitch angular momentum management. Although the resolution in the magnitude of the dipole moments is good, the period of the telemetered data is 16 seconds. Due to the relatively weak contribution of the dipoles to the ERBS dynamics (maximum 4 x 10^{-3} N-m in roll), the lack of time resolution in the dipole telemetry is less important. Also, the digitization of the magnetometer data contributes a smaller error to the dipole torques.

DATA CONDITIONING TECHNIQUES

To accurately represent the ERBS dynamics, the control system telemetry data require extensive preprocessing. Interpolation between low-frequency data points is necessary for the control torques to be evaluated at each integration point. As a benefit of a properly chosen smoothing or interpolation method, increased accuracy may be synthesized from coarse data. The behavior of different smoothing and interpolation techniques varies with data characteristics.

Incorporated in the ERBS ADS is a polynomial smoothing algorithm that fits a Chebyshev polynomial of user-specified order to a set of data (Reference 4). The smoother provides interpolation between low-frequency data points. For the Scanwheel data, the Chebyshev polynomial smoothing algorithm may be tuned to approximate actual Scanwheel activity. The fit is limited by array size to a maximum of 600 sec through one cycle of the Data Adjuster. Problems can arise at the endpoints of high order fits. Since a high-order fit is necessary to match peak-to-peak trends, a modification to the algorithm allows extrapolated beginning and endpoint divergences to be replaced with the constant values from the first and last raw data points, respectively. Figure 2 illustrates an 18th-order Chebyshev fit to raw Scanwheel data.

The coarseness of the raw momentum wheel data, resembling a series of step discontinuities about a mean, exacerbates the representation of realistic wheel behavior using a simple smoother. The raw momentum wheel data of Figure 3 most likely results from transitions near a telemetry digitization boundary. An 18th-order Chebyshev fit to the data is represented by the smooth line. A high-order polynomial follows the step discontinuities too closely while causing boundary matching problems for multiple Data Adjuster cycles. For a low-order polynomial, this problem is reduced, but detail is lost. Trends are represented by two or more steps in the same direction. Matching the step values exactly considerably overestimates the change in momentum wheel speed and drives unrealistic momentum into the model of pitch axis rates.

A more realistic fit to the momentum wheel data is obtained by preprocessing the raw data in a simple running average algorithm to reduce discontinuities. A Chebyshev polynomial is then fit to the averaged data to interpolate between the 8-sec data points. The polynomial fit also permits additional tuning to mimic real wheel behavior. Figure 4 illustrates averaged and smoothed momentum wheel data. Problems still exist in extracting actual wheel behavior from the coarse data.

Good results are achieved using the Chebyshev fit for the low-frequency magnetometer data because the telemetry digitization is small relative to the magnitudes of the trends in the data. Figure 5 illustrates a 12th-order fit to magnetometer data. Typical magnetic dipole moment telemetry is represented by the points in Figure 6. Polynomial smoothing of dipole data is inappropriate because of the large discontinuities. The dipole moments are interpreted as constant valued between 16-sec data points. This assumption is made because the chance of the subcommutator interrogating the dipole at a particular value is proportional to the amount of time the dipole moment maintains that value. The solid line in Figure 6 represents the interpolated dipole moment values. Error imparted to the dynamic model by the low frequency of dipole telemetry is acceptable because of the small magnitude of the dipole control torques.

RESULTS

Two criteria are used to assess the accuracy of the modeled dynamics: comparison of the modeled rates to valid IRU rates from the same data span and comparison of FADS attitudes propagated by modeled

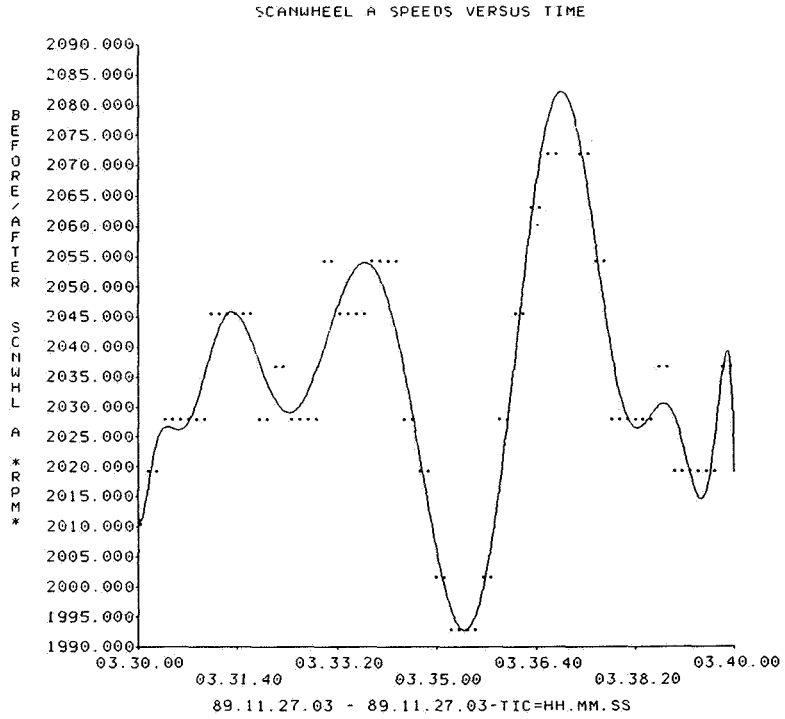


Figure 2. Raw Scanwheel Speed Data (Points) and 18th-Order Chebyshev Polynomial Fit (Smooth Line)

rates with FADS attitudes propagated by valid IRU rates. For the evaluation, telemetry data were chosen from periods when all three IRU axes were functioning and gyro noise was minimal. Full orbit timespans of data were processed with the model. Figure 7 shows the results of one run with the modeled rates represented by the smooth line together with raw IRU telemetry. The IRU data are considered an absolute reference. Root mean square noise in the IRU telemetry for this period is about three digitization steps (0.003 deg/sec).

Agreement between the modeled and measured rates is good for roll and yaw, in most cases within 0.002 deg/sec of the mean of the IRU signal, but poor for pitch. Although some trends are seen in the modeled data, much of the pitch dynamics information is not resolvable in the large momentum wheel telemetry digitization. The constant offset between the modeled and measured rates in Figure 7 is due to the arbitrary integration starting point and may be corrected with an appropriate initial value corresponding to the constant of integration. The FADS is able to estimate these values as rate biases.

The most significant test of this method of dynamic modeling is its ability to accurately propagate attitudes over a full orbit. The FADS was used to produce one-orbit attitudes using modeled rates in place of measured IRU rates. Figures 8 through 13 show the FADS roll, pitch, and yaw derived from modeled rates of Figure 7 together with the reference FADS angles derived from IRU measured rates. In this example, only the epochal angles and rate biases were included in the estimated state. The modeled roll and yaw match the reference to within 0.2 deg. Pitch, however, diverges up to 3 deg.

To demonstrate the effect of environmental torques on the model, the data of Figure 7 were reprocessed with no environmental torques included. Figure 14 shows the results of this run. No change in the roll or yaw rates was observed, but lack of the constant component of the aerodynamic torque causes the pitch axis rate to gradually diverge. Although it is difficult to resolve the environmental torques because of the

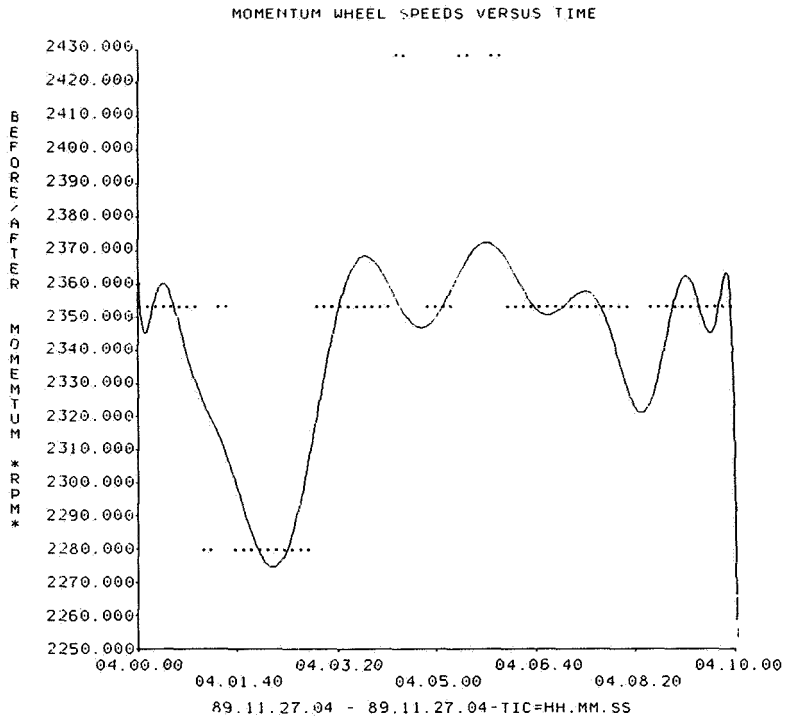


Figure 3. Raw Momentum Wheel Speed Data (Points) and 18th-Order Chebyshev Polynomial Fit (Smooth Line)

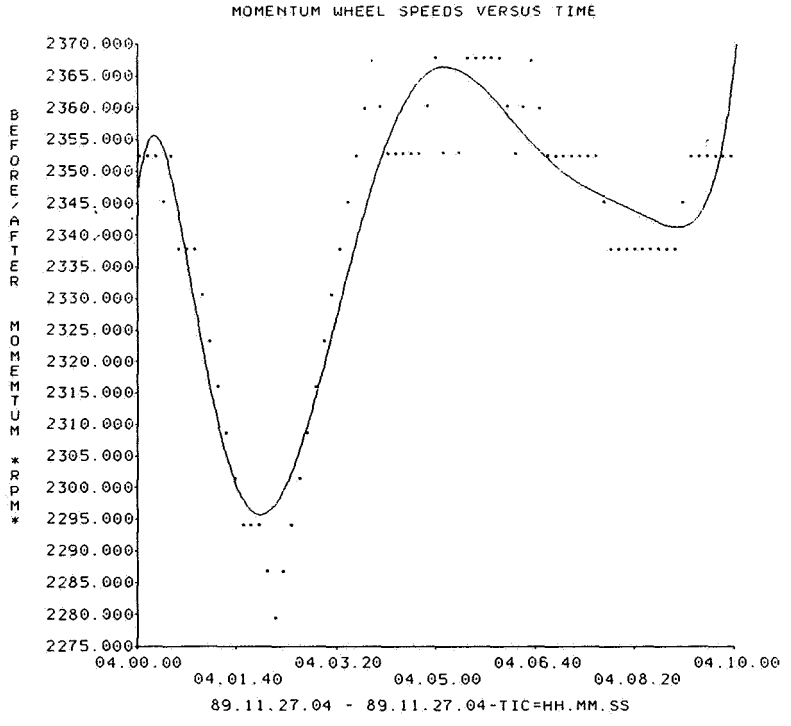


Figure 4. Running Averaged Momentum Wheel Speed Data (Points) and Eighth-Order Chebyshev Polynomial Fit to the Averaged Data

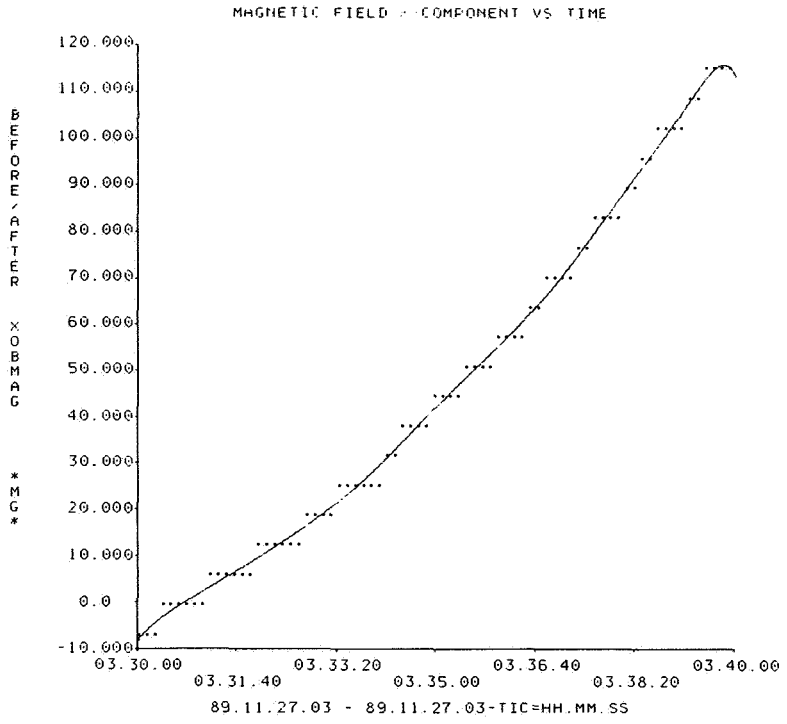


Figure 5. Twelfth-Order Chebyshev Polynomial Fit (Smooth Line) to Raw Magnetometer Data (Points)

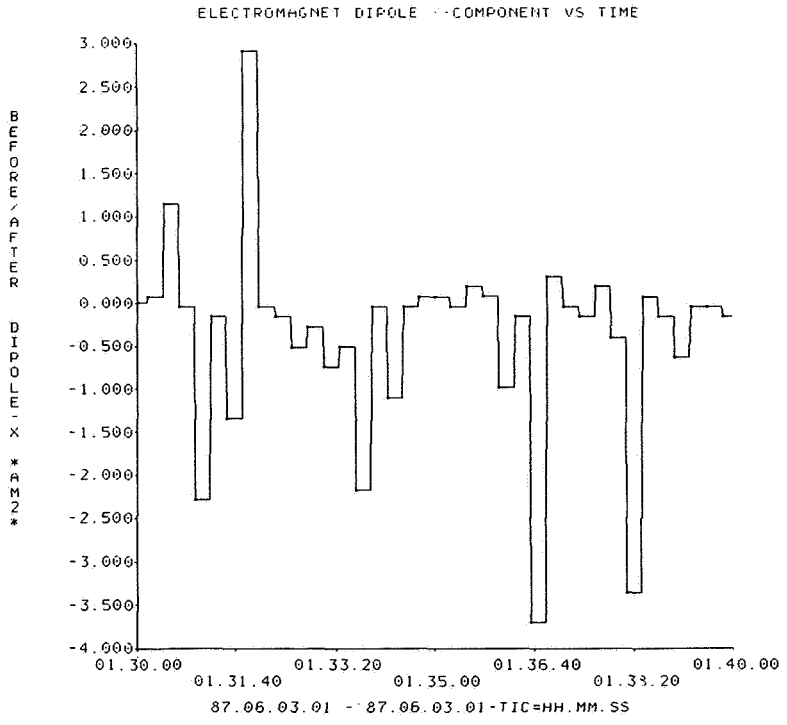


Figure 6. Raw Magnetic Dipole Moment Data at 16-sec Intervals (Points) and Interpolated Values (Solid Lines)

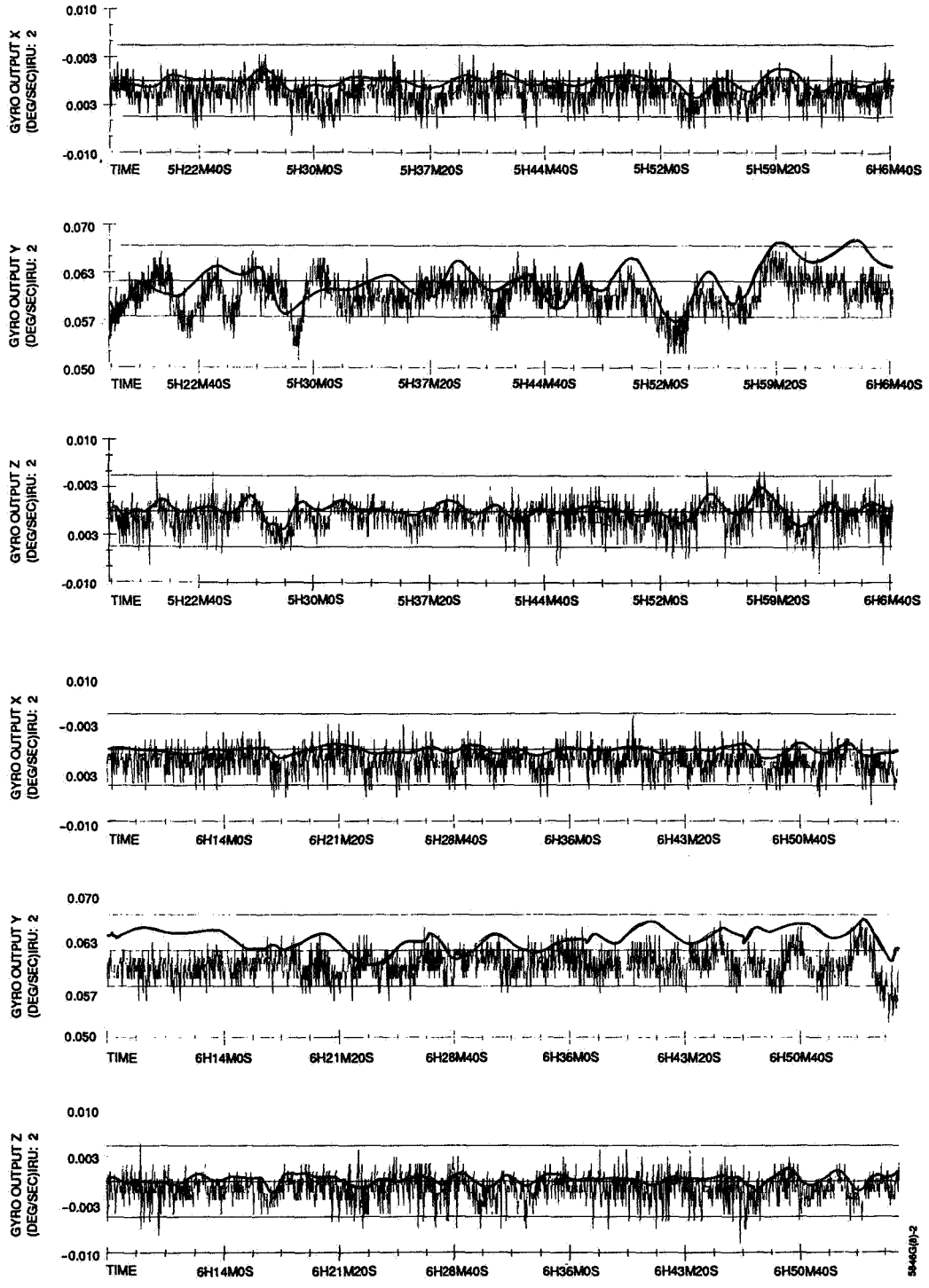


Figure 7. Comparison of Modeled Rates (Smooth Line), Including Environmental Torques, To Measured IRU Rates for One Orbit

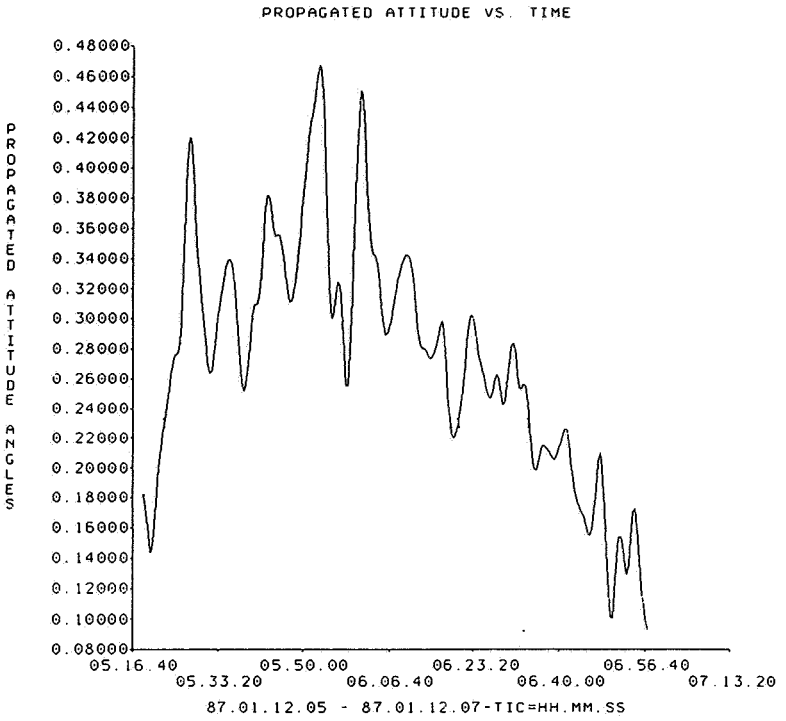


Figure 8. FADS Roll Using Modeled Rate Data

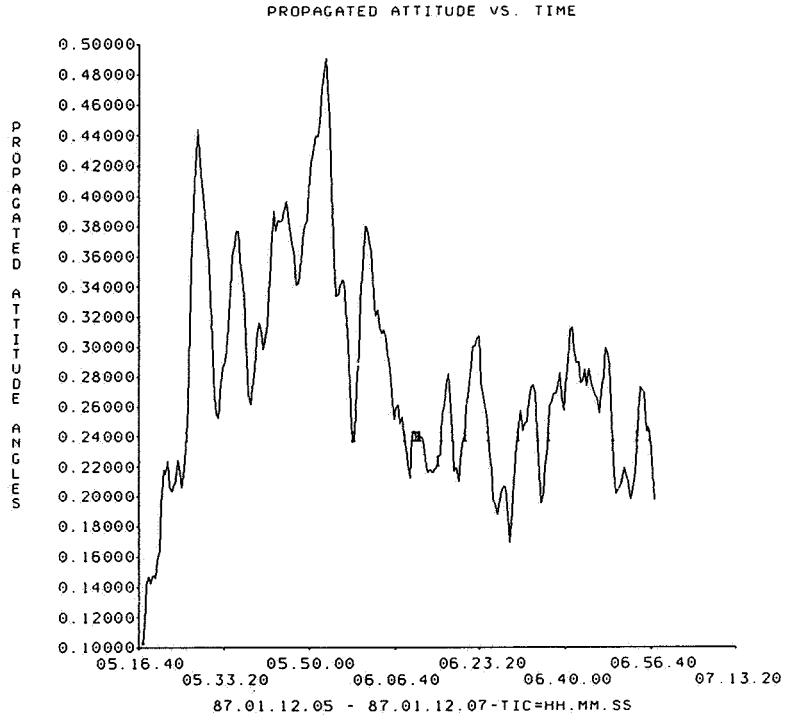


Figure 9. FADS Reference Roll Using IRU Data

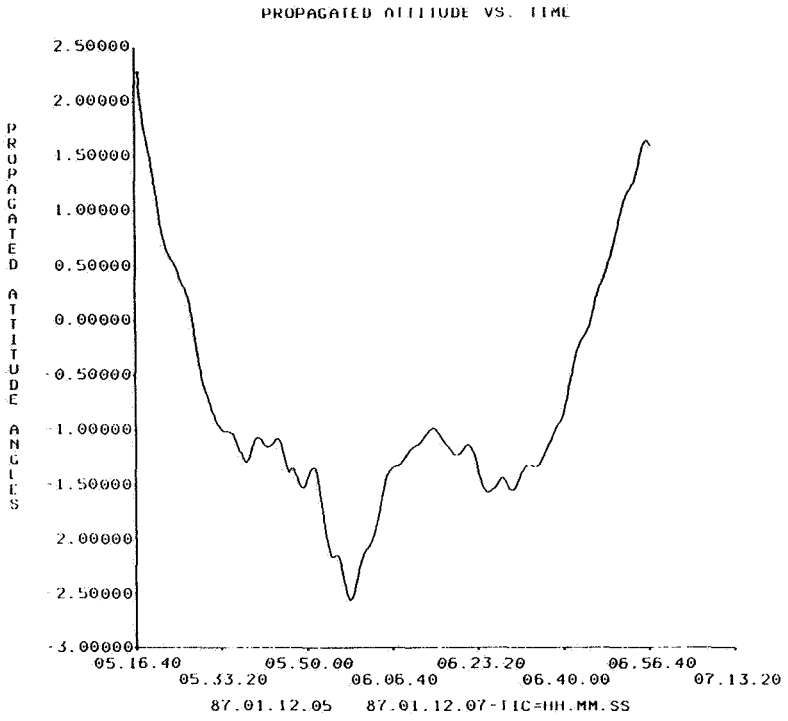


Figure 10. FADS Pitch Using Modeled Rate Data

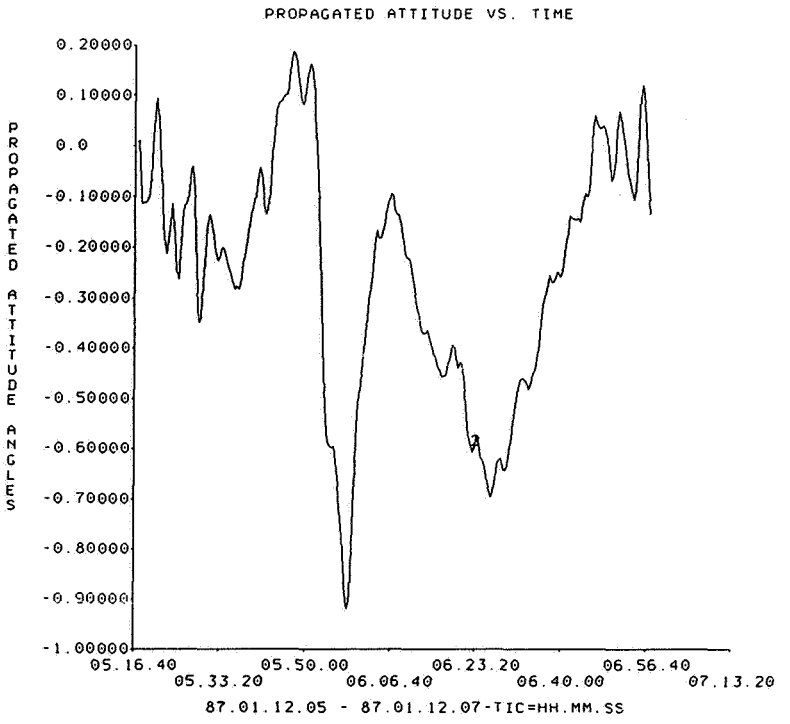


Figure 11. FADS Reference Pitch Using IRU Data

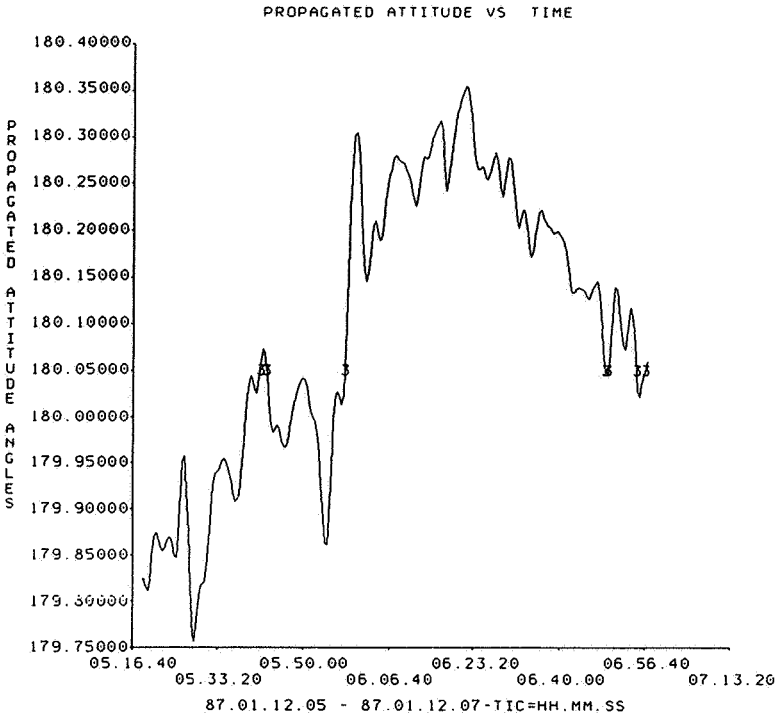


Figure 12. FADS Yaw Using Modeled Rate Data

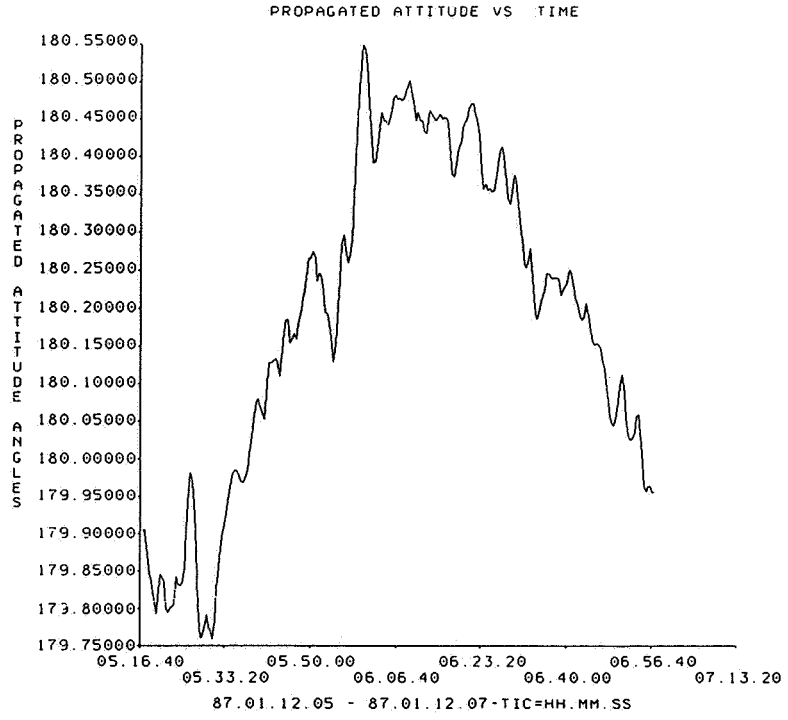


Figure 13. FADS Reference Yaw Using IRU Data

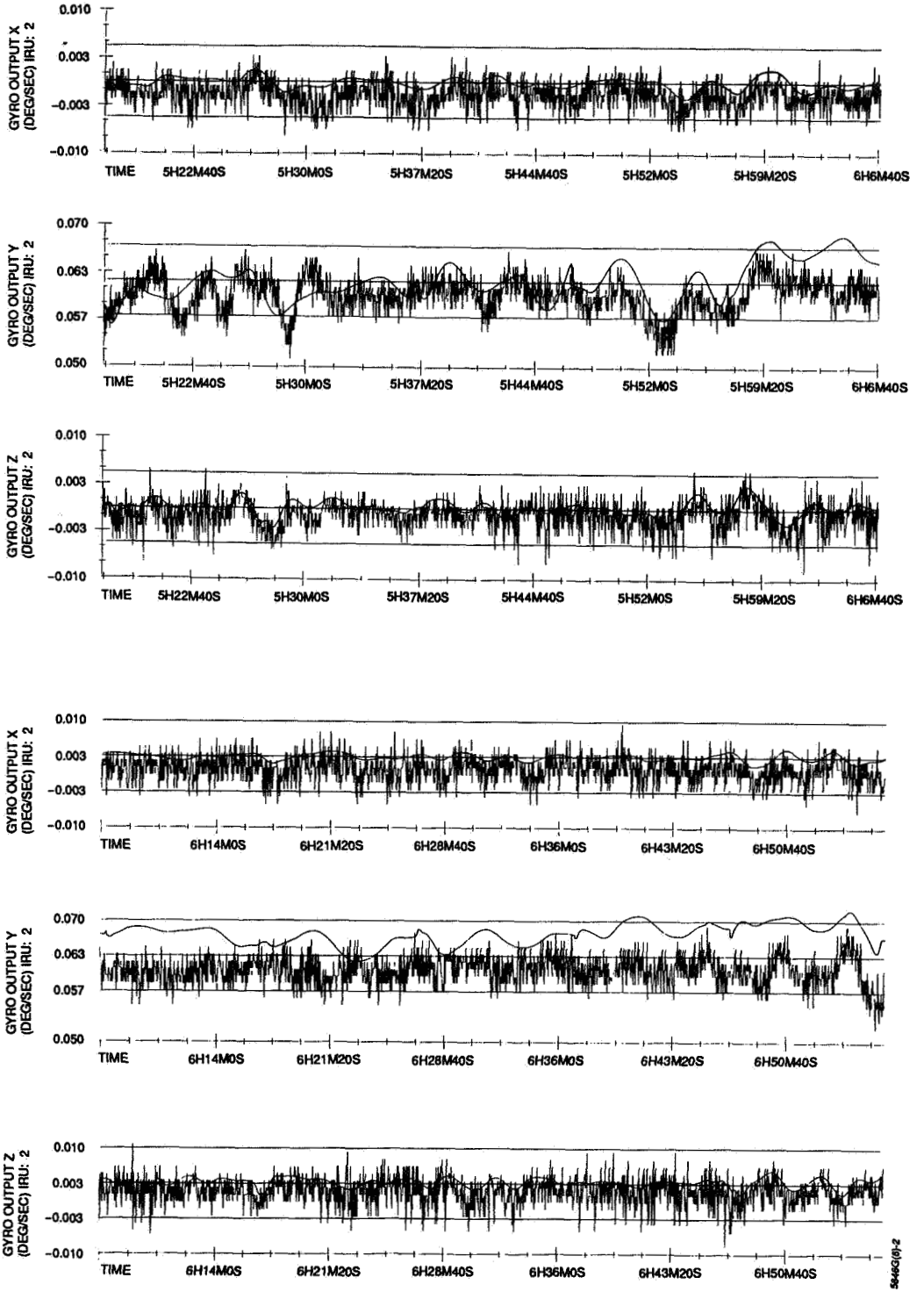


Figure 14. Comparison of Modeled Rates (Smooth Line) Excluding Environmental Torque Models to Measured IRU Rates

coarseness of the control telemetry, their effect on the actual spacecraft dynamics is diminished due to gyroscopic stiffness from the strong angular momentum bias of the ERBS.

ATTITUDE ANOMALIES

Beginning November 27,1989, the ERBS was configured under dual scanner MCS control during a period when the field of view trace of one Scanwheel viewed the Sun during sunrise and sunset events. The Sun pulse resulted in control anomalies of about 2 deg on all three axes. During this period, the pitch axis gyro was failed, and the roll axis gyro was severely degraded, precluding use of the FADS for reference attitudes to evaluate the dynamic model. The CADS was used for reference attitudes but gave accurate solutions only when Sun data were available following sunrise. Figures 15 through 20 show CADS angles computed with Sun, magnetometer, and uncorrupted IR Scanwheel data with corresponding data from modeled rates through a Sun interference occurrence. Agreement between the CADS attitudes and modeled FADS attitudes is within 0.2 deg even for pitch. The large change in momentum wheel speed reduces the effect of telemetry digitization and allows the pitch dynamics to be more accurately determined. Figure 21 shows the raw momentum wheel data from this period. The Sun interference occurrence demonstrates the accuracy of the dynamic model when large trends in control system activity make control torques more observable.

CONCLUSIONS

It has been demonstrated for the ERBS that dynamic modeling for attitude determination is feasible when control system data of sufficient frequency and resolution are available. Insufficient knowledge of control torques limits the accuracy of the modeled attitude dynamics. Poor resolution in the ERBS momentum wheel downlink telemetry results in poor pitch rate determination. It was found that improved granularity of the momentum wheel data was more important than increased time resolution, since the telemetry frequency provides sufficient input to the model relative to the timescale of the onboard control processes. When a significant perturbation to the nominal state occurs, the effect of telemetry resolution is reduced. The dynamic processes on all three axes are then observable in downlink telemetry, and the model performs well.

The dependence of the modeled dynamics on accurate control system information dominates the effect of environmental torques on the system. Better telemetry resolution is required to assess the accuracy of the disturbance torque models. The high momentum bias of the ERBS may diminish the effect of environmental torques on the actual dynamics, however. More detailed environmental torque models may be necessary for zero angular momentum spacecraft using this method.

One recommendation for further improvement in the model is to increase resolution of the momentum wheel control torques. Efforts to derive actual wheel behavior by preprocessing momentum wheel telemetry data resulted in limited accuracy. An alternative method would extract the commanded momentum wheel speed from a model of the onboard MCS pitch control law based on IR Scanwheel pitch angle input. Resolution in the Scanwheel fine pitch angle is 0.025 deg, which corresponds to a step change in pitch axis angular momentum of 0.25 N-m/sec, or a momentum wheel step of 27 rpm over the 8-sec interval. Since the Scanwheel pitch angle results from the sum total torques acting on the spacecraft pitch axis, this method would attribute all pitch motion to momentum wheel control. This would preclude estimation of the y-axis disturbance torques but could cause problems with magnetic dipole coupling in pitch axis control.

Another method for improving momentum wheel speed resolution would derive the analog tachometer input to the onboard analog to digital converter from the digitized telemetry data. A least-squares fit would be applied to the raw data, subject to known constraints on the control system. The success of this method would depend on the uniqueness of the analog input function.

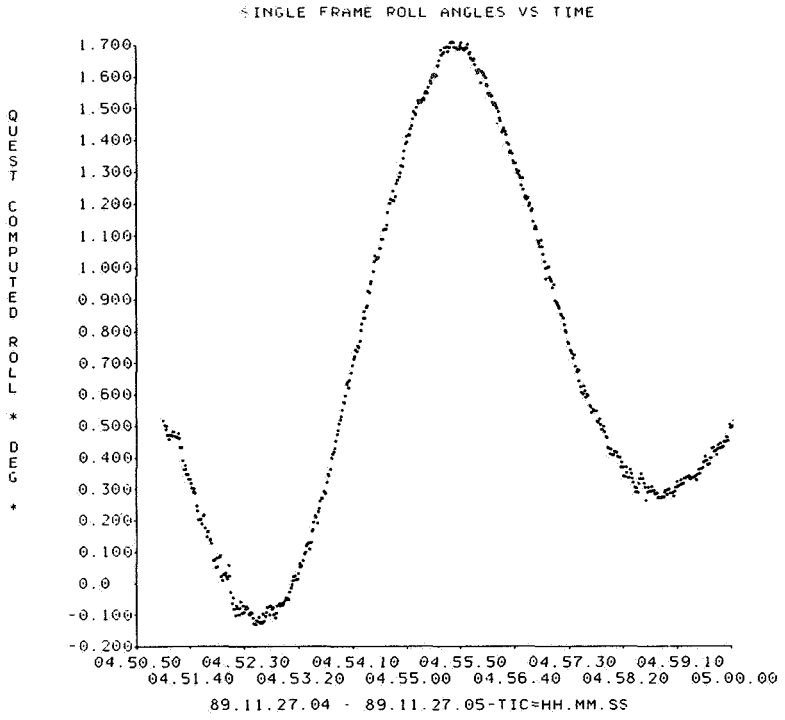


Figure 15. CADS Roll Angle During Sun Interference Anomaly

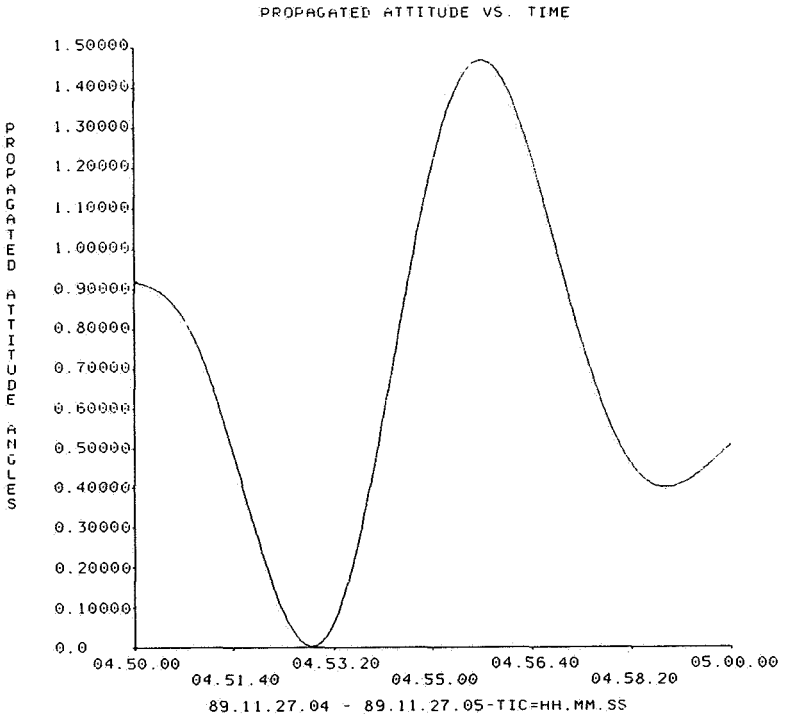


Figure 16. FADS Roll Angle During Sun Interference Anomaly Using Modeled Rates

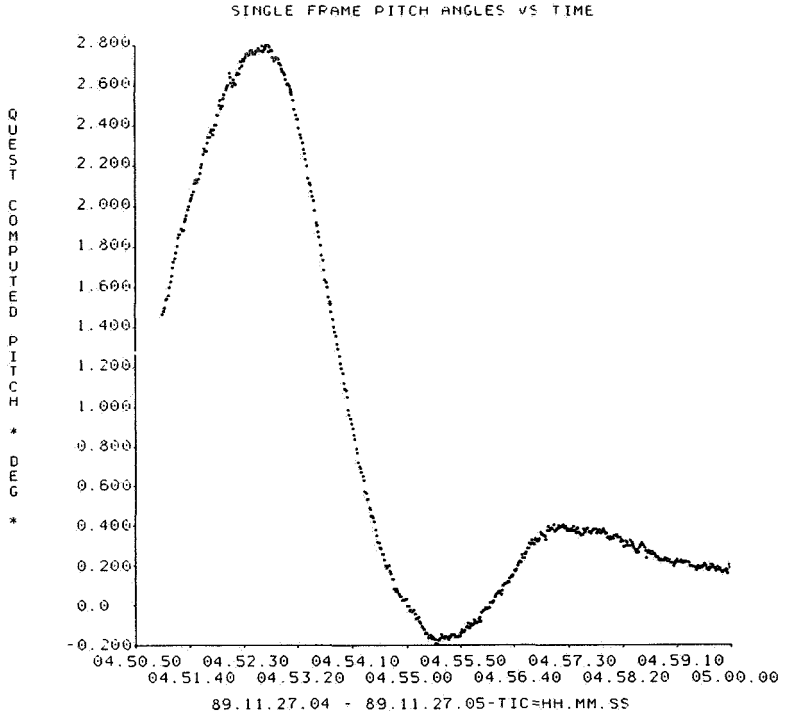


Figure 17. CADS Pitch Angle During Sun Interference Anomaly

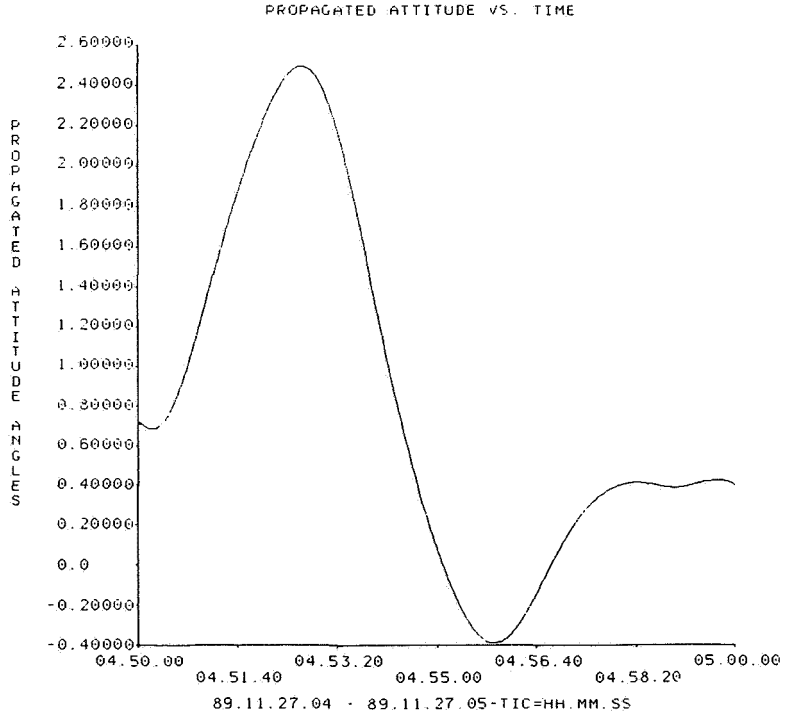


Figure 18. FADS Pitch Angle During Sun Interference Anomaly Using Modeled Rates

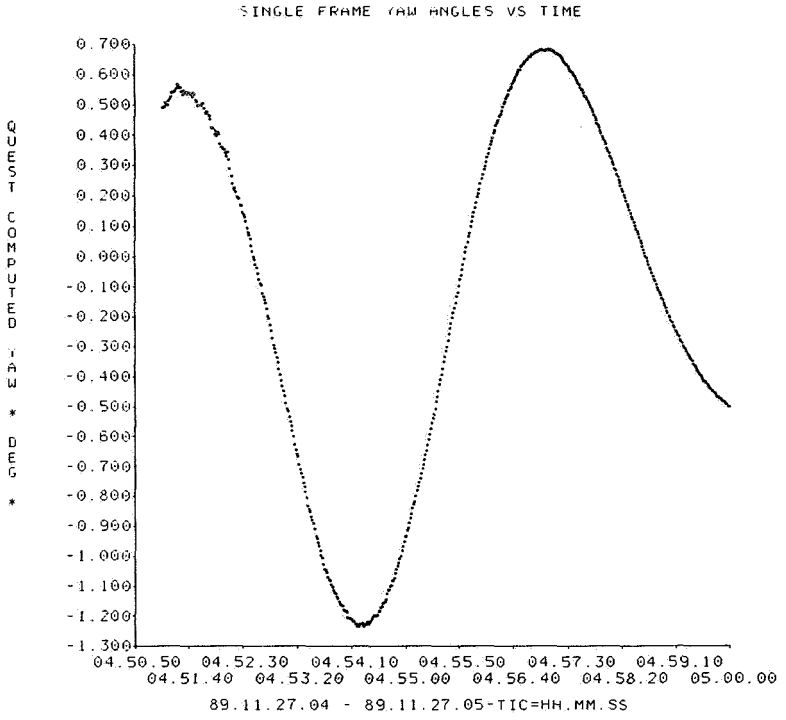


Figure 19. CADS Yaw Angle During Sun Interference Anomaly

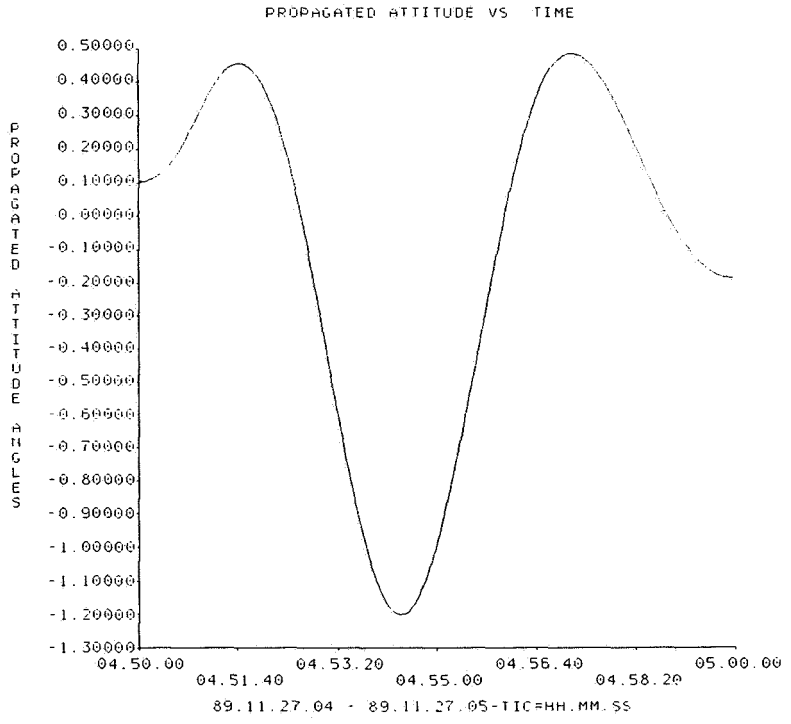


Figure 20. FADS Yaw Angles During Sun Interference Anomaly Using Modeled Rates

Integration of the dynamic model in the ERBS Attitude Determination System is very efficient. This approach exploits many capabilities of the existing ground support software, such as telemetry processing, data smoothing, graphic displays, and the fine attitude determination algorithm. For future work, incorporating this simple dynamic model into a state estimation algorithm using attitude sensor data would allow additional parameters, such as unmodeled torque biases, to be estimated. Improvement in the accuracy of attitudes propagated in this way could be expected.

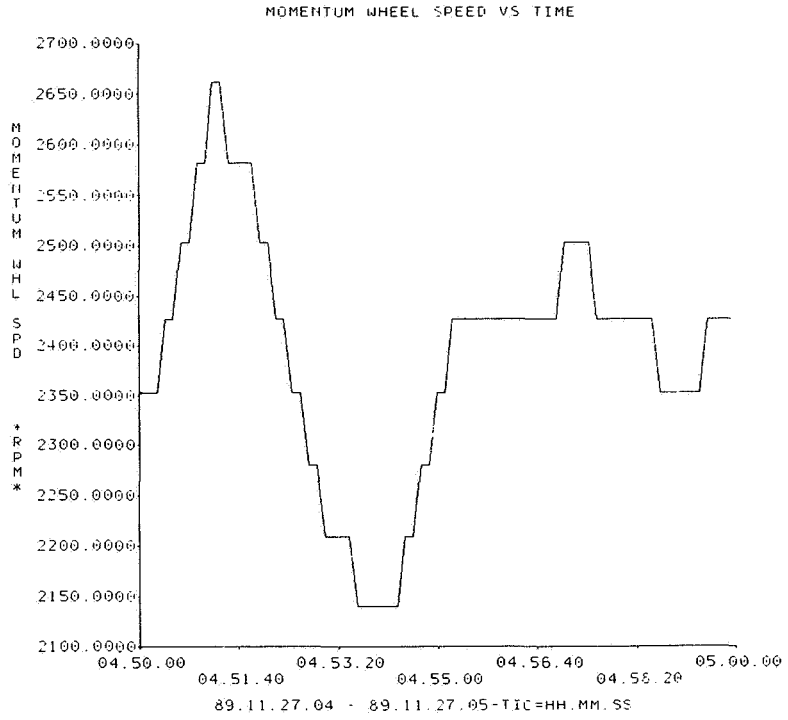


Figure 21. Raw Momentum Wheel Speed Data During Sun Interference Anomaly

REFERENCES

1. E. J. Lefferts and F. L. Markley, "Dynamic Modeling for Attitude Determination," paper presented at AIAA Guidance and Control Conference, San Diego, California, August 16-18, 1976

2. Computer Sciences Corporation, CSC/TM-76/6235, *Nimbus-G Attitude Determination Feasibility Study Utilizing the Attitude Dynamics Generator (ADGEN)*, F. L. Markley and J. W. Wood, December 1976

3. --, CSC/TM-77/6115, *Evaluation of Wheel Rate Data for LANDSAT Attitude Modeling*, J. Fein, April 1977

4. --, CSC/SD-82/6013, *Earth Radiation Budget Satellite (ERBS) Attitude Ground Support System Functional Specifications and Requirements*, G. Nair et al., September 1982

5. --, CSC/SD-84/6004UD1, *Earth Radiation Budget Satellite (ERBS) Attitude Determination System (ADS) System Description and User's Guide, Volume 1*, R. Strang, December 1985

6. --, CSC/TM-84/6071, *Earth Radiation Budget Satellite (ERBS) Attitude Analysis and Support Plan*, M. Phenneger and J. Rowe, August 1984

COMPARISON AND TESTING OF EXTENDED KALMAN FILTERS FOR ATTITUDE ESTIMATION OF THE EARTH RADIATION BUDGET SATELLITE

by

Julie Deutschmann
Flight Dynamics Analysis Branch
NASA Goddard Space Flight Center
Greenbelt, MD 20771

Itzhack Y. Bar-Itzhack
Aerospace Department
Technion-Israel Institute of Technology
Haifa 32000 Israel

Mohammad Rokni
Computer Sciences Corporation
Lanham-Seabrook, MD 20706

ABSTRACT

This paper presents the testing and comparison of two Extended Kalman Filters (EKFs) developed for the Earth Radiation Budget Satellite (ERBS). One EKF updates the attitude quaternion using a four component additive error quaternion . This technique is compared to that of a second EKF, which uses a multiplicative error quaternion. A brief development of the multiplicative algorithm is included. The mathematical development of the additive EKF was presented in the 1989 Flight Mechanics/Estimation Theory Symposium along with some preliminary testing results using real spacecraft data. A summary of the additive EKF algorithm is included. The convergence properties, singularity problems, and normalization techniques of the two filters are addressed. Both filters are tested with simulated ERBS sensor data in addition to real ERBS sensor data. The results of the two filters are also compared to those from the ERBS operational ground support software, which uses a batch differential correction algorithm to estimate attitude and gyro biases. Sensitivity studies are performed on the estimation of sensor calibration states. The potential application of the EKF for real time and non-real time ground attitude determination and sensor calibration for future missions such as the Gamma Ray Observatory (GRO) and the Small Explorer Mission (SMEX) is also presented.

I. INTRODUCTION

The purpose of this study was to test and compare two EKFs developed for ERBS. ERBS is equipped with the following sensors that are used for attitude determination: 2 redundant Inertial Reference Units (IRUs) each containing 3 single-axis gyroscopes, 2 digital fine Sun sensors (FSSs), 2 infrared (IR) horizon scanners, and 1 three-axis magnetometer. The state estimated by both EKFs consists of the attitude parameters, sensor misalignments for the Sun sensor, magnetometer and gyros, biases for the Sun sensor, horizon scanner, magnetometer and gyros, and scale factors for the Sun sensor, horizon scanner, magnetometer and gyros. A real time EKF was also developed which estimates only the attitude parameters and the gyro bias.

The development and initial testing of the additive EKF was presented in Reference 1. This filter was tested with only real data over short timespans. A multiplicative EKF was designed and tested and is presented in Reference 2. This work presents further testing of the additive EKF and comparison of the additive EKF to a multiplicative EKF adapted from that presented in Reference 2. The two filters are also compared, when possible, to the current ERBS batch algorithm which is used for fine attitude and gyro bias estimation. Both simulated data and real data are used for the testing and comparison.

In the additive EKF, the estimated quaternion is not necessarily normal unless it converges to the correct quaternion. Reference 3 shows that normalization speeds convergence of the filter and eliminates the need for filter tuning. The normalization techniques used in References 1 and 3 were external to the EKF algorithm. The covariance computation was not affected by the normalization, but the state estimation algorithm had to be modified to incorporate the part of the state estimate that was lost in the normalization. The realization of the normalization process as an update using a pseudo-measurement blends naturally into the EKF algorithm and does not require any modification of the EKF itself. This technique is tested on the additive EKF and compared to the original normalization process. The need for normalization in the multiplicative EKF is also presented.

II. THE EXTENDED KALMAN FILTER ALGORITHM

The EKF algorithm is based on the following assumed models

System model

$$\dot{\underline{X}} = \underline{f}(\underline{X}(t),t) + \underline{w}(t) \qquad (2.1)$$

Measurement model

$$\underline{z}_k = \underline{h}_k(\underline{X}(t_k)) + \underline{v}_k \qquad (2.2)$$

where: $\underline{X}(t)$ = state vector
$\underline{w}(t)$ = zero mean white process
$\underline{v}_k$ = zero mean white sequence

The measurement update and the propagation of the state estimate and of the error covariance are performed as

Update:
$$\hat{\underline{X}}_k(+) = \hat{\underline{X}}_k(-) + K_k[\underline{z}_k - \underline{h}_k(\hat{\underline{X}}_k(-))] \qquad (2.3)$$

where the gain matrix, K, is evaluated as

$$P_k(+) = [I - K_kH_k]P_k(-)[I - K_kH_k]^T + K_kR_kK_k^T \qquad (2.4)$$

$$K_k = P_k(-)H_k^T(\hat{\underline{X}}_k(-))[H_k(\hat{\underline{X}}_k(-))P_k(-)H_k^T(\hat{\underline{X}}_k(-)) + R_k]^{-1} \qquad (2.5)$$

Propagation:

$$\dot{\hat{\underline{X}}}(t) = \underline{f}(\hat{\underline{X}}(t),t) \qquad (2.6)$$

$$\dot{P}(t) = F(\hat{\underline{X}}(t),t)P(t) + P(t)F^T(\hat{\underline{X}}(t),t) + Q(t) \qquad (2.7)$$

where:

$$F(\underline{X}(t),t) = \left.\frac{\underline{f}(\underline{X}(t),t)}{\underline{X}(t)}\right|_{\underline{X}(t)=\hat{\underline{X}}(t)} \qquad (2.8a)$$

$$H(\underline{X}(-)) = \left.\frac{\underline{h}(\underline{X}(t))}{\underline{X}(t)}\right|_{\underline{X}(t)=\hat{\underline{X}}} \qquad (2.8b)$$

P_k = estimation error covariance matrix
R_k = covariance matrix of the white sequence $\underline{v}_k$
Q_k = spectral density matrix of $\underline{w}(t)$

The state vector was selected to be

$$\underline{X} = \begin{bmatrix} \underline{q} \\ \underline{S}_g \\ \underline{\Theta}_g \\ \underline{b}_g \\ \underline{\Theta}_s \\ \underline{S}_s \\ \underline{b}_s \\ \underline{b}_h \\ \underline{S}_m \\ \underline{\Theta}_m \\ \underline{b}_m \end{bmatrix} \qquad (2.9)$$

- $\underline{q}$ — 4 quaternion components
- $\underline{S}_g$ — 3 gyro scale factor errors
- $\underline{\Theta}_g$ — 6 gyro misalignment angles
- $\underline{b}_g$ — 3 gyro biases
- $\underline{\Theta}_s$ — 3 FSS misalignment angles
- $\underline{S}_s$ — 2 FSS scale factor errors
- $\underline{b}_s$ — 2 FSS biases
- $\underline{b}_h$ — 2 IR horizon scanner biases
- $\underline{S}_m$ — 3 magnetometer scale factor errors
- $\underline{\Theta}_m$ — 6 magnetometer misalignments
- $\underline{b}_m$ — 3 magnetometer biases

The effective measurements used to update the filter are defined as

$$\underline{y} = M_{AT}\underline{W}_{T',meas} - A(\hat{\underline{q}})\underline{V}_I \qquad (2.10)$$

where: $\underline{y}$ = effective measurements (or residuals)
M_{AT} = transformation matrix from nominal (nonmisaligned) sensor to body coordinates
$\underline{W}_{T',meas}$ = unit vector as measured by the sensor in the sensor misaligned coordinates

$A(\hat{\underline{q}})$ = transformation matrix from inertial to body coordinates as a function of the estimated quaternion

$\underline{V}_I$ = measured unit vector as known in inertial coordinates

While the traditional EKF algorithm updates the state estimate according to (2.3), we use $\underline{y}$ (as computed in (2.10)) to update the state estimate as

$$\hat{\underline{X}}_k(+) = \hat{\underline{X}}_k(-) + \underline{x}(t_k) \qquad (2.11) \qquad \text{where:} \quad \underline{x}(t_k) = K_k \underline{y}_k \qquad (2.12)$$

It can then be shown (Reference 1) that $\underline{z}_k - \underline{h}_k(\hat{\underline{X}}_k(-))$ is linearly related to $\underline{x}(t_k)$. The EKF estimates $\underline{x}(t_k)$ and then adds the estimate $\underline{x}(t_k)$ to $\hat{\underline{X}}_k(-)$, the best estimate of $\underline{X}(t_k)$. The linear relationship between $\underline{x}(t_k)$ and $\underline{y}_k$ will be shown in Section IV.

In the additive EKF, the first four components of $\underline{x}(t_k)$ are the corrections to the quaternion estimated by the EKF, denoted as $\delta\underline{q}$. These corrections are added to $\underline{q}(-)$, the best estimate of $\underline{q}$. The remaining elements in $\underline{x}(t_k)$ are the corrections to the calibration states which are also then added to the current best estimate of those states.

In the multiplicative EKF, the quaternion elements of $\underline{x}(t_k)$ are treated differently. The definition of $\underline{x}(t_k)$ is given as

$$\underline{x}(t_k) = \begin{bmatrix} \underline{\alpha} \\ \\ \underline{\delta\tau} \end{bmatrix} \qquad (2.13)$$

where:

$\underline{\alpha} = \begin{bmatrix} \phi \\ \Omega \\ \mu \end{bmatrix}$ = three small angles based on the assumption that the error quaternion is composed of three small angles (as the vector portion) and 1 (scalar portion)

$\underline{\delta\tau}$ = corrections to the calibration parameters given in (2.9)

The correction to the quaternion, given as $\delta\hat{\underline{q}}$, is then constructed according to

$$\delta\hat{\underline{q}}^T = \begin{bmatrix} \tfrac{1}{2}\phi & \tfrac{1}{2}\Omega & \tfrac{1}{2}\mu & 1 \end{bmatrix} \qquad (2.14)$$

The quaternion is updated as

$$\hat{\underline{q}}_{k+1}(+) = \hat{\underline{q}}_{k+1}(-)\delta\hat{\underline{q}}^{-1} \qquad (2.15)$$

The calibration components are updated according to (2.11). The updated values of the calibration components and $\hat{\underline{q}}_{k+1}(+)$ are augmented into (2.9) and are propagated in time using (2.6). The dynamics of the two EKFs will be presented in the next section. For further discussion of the algorithms see References 1 and 2.

III. THE DYNAMICS MODEL

The states that vary in time are the attitude parameters and "bias" states that are modeled as Markov rather than true bias states. The scale factors and misalignments are assumed to be constant in time. The elements of $\underline{x}(t_k)$ for the additive filter are the same as those shown in (2.9) with the quaternion error, $\delta\underline{q}$, replacing $\underline{q}$. The differential equation which governs the propagation of $\underline{x}$ is obtained by combining the linear differential equations of the components of the attitude augmented state vector. This yields an equation of the form

$$\dot{\underline{x}} = F(\hat{\underline{X}})\underline{x} + \underline{n} \qquad (3.1)$$

This equation is presented below with F and $\underline{n}$ given. The matrices $\tilde{\Phi}$, $\hat{B}$, $\tilde{U}$, $\tilde{W}$, and T which form F are derived and defined in Reference 1. F given below and in (3.1) is defined in (2.8a). It is used to propagate the covariance matrix in (2.7).

$$\frac{d}{dt}\begin{bmatrix} \delta\underline{x} \\ \underline{s}_g \\ \underline{\theta}_g \\ \underline{b}_g \\ \underline{\theta}_s \\ \underline{s}_s \\ \underline{b}_s \\ \underline{b}_h \\ \underline{s}_m \\ \underline{\theta}_m \\ \underline{b}_m \end{bmatrix} = \begin{bmatrix} \tilde{\Phi} & \hat{\tilde{B}}\tilde{U} & \hat{\tilde{B}}\tilde{W} & \hat{B} & & & & & & & \\ & & & T_g & & & & & & & \\ & & & & T_s & & & & & & \\ & & & & & T_h & & & & & \\ & & & & & & & & & T_m & \end{bmatrix} \begin{bmatrix} \delta\underline{x} \\ \underline{s}_g \\ \underline{\theta}_g \\ \underline{b}_g \\ \underline{\theta}_s \\ \underline{s}_s \\ \underline{b}_s \\ \underline{b}_h \\ \underline{s}_m \\ \underline{\theta}_m \\ \underline{b}_m \end{bmatrix} + \begin{bmatrix} \hat{B}\underline{n}_{g1} \\ \\ \underline{n}_{g1} \\ \\ \\ \underline{n}_s \\ \underline{n}_h \\ \\ \\ \underline{n}_m \end{bmatrix} \qquad (3.2)$$

In the multiplicative EKF, $\underline{x}$ is replaced with (2.13) and the first row of $F(\underline{X})$ is replaced with

$$\begin{bmatrix} \tilde{\Omega} & -\tilde{U} & \tilde{W} & I & & & & & & & \end{bmatrix} \qquad (3.3)$$

where:

$$\tilde{\Omega} = \begin{bmatrix} 0 & \tilde{w}_z & -\tilde{w}_y \\ -\tilde{w}_z & 0 & \tilde{w}_x \\ \tilde{w}_y & -\tilde{w}_x & 0 \end{bmatrix} \qquad (3.4)$$

I = 3x3 identity matrix

The first row of $\underline{n}$ in (3.1), defined in (3.2), is replaced by $\underline{n}_{g1}$. The spectral density of the elements of the white noise driving the Markov states in $\underline{x}$ in both EKFs is related to the individual states they drive according to the well known relation $Q_i = 2\sigma^2_{i,o}/T_i$ (Reference 4) where Q_i is the spectral density of the white noise driving state i, T_i is the time constant of this Markov state and $\sigma_{i,o}$ is the initial standard deviation of the state.

The estimation problem dealt with in this paper is characterized by a linear dynamics equation. The system dynamics can be augmented to obtain an equation of the form (Reference 1)

$$\dot{\underline{X}} = f(t)\underline{X} + \underline{n} \qquad (3.5)$$

where $\underline{X}$ is given in (2.9) (for both EKFs) and the first row of F(t) in (3.2) is replaced with the following to define f(t)

$$\begin{bmatrix} \tilde{\Phi} & & & & & & & & & & \end{bmatrix} \qquad (3.6)$$

The white noise vector $\underline{n}$ in (3.5) is of no consequence in the estimation process since according to (2.6) the propagation of $\underline{X}$ requires only the evaluation of f(t).

IV. THE MEASUREMENT MODEL

As mentioned in Section II the effective measurements used to update the filter are defined as

$$\underline{y} = M_{AT}\underline{W}_{T',meas} - A(\hat{\underline{g}})\underline{V}_I \qquad (4.1)$$

Incorporating the sensor misalignments, scale factor errors, and biases into (4.1) the linear relationship between $\underline{y}$ and $\underline{x}$ can be derived. This yields the following

$$\underline{y} = H_q d\underline{g} + M_{AT}[\underline{W}_{T',meas}\times]\underline{\theta} + M_{AT}d\underline{W}_{T'} \qquad (4.2)$$

The terms introduced in (4.2) are derived and defined in Reference 1. The term $[\underline{W}_{T'},_{meas}x]$ represents the replacement of the cross product with the multiplication of an anti-symmetric matrix, composed of $\underline{W}_{T'},_{meas}$. Equation (4.1) shows how to generate the effective measurement $\underline{y}$ which updates the estimate and (4.2) indicates the linear relationship between $\underline{y}$, the attitude errors, the misalignment errors of the sensor being used and $d\underline{W}_{T'}$, the total error generated by the sensor. Equation (4.2) is the first stage in finding the measurement matrix, H, (defined in (2.8b)) for each sensor used onboard ERBS.

The effective measurement (4.1) for the multiplicative EKF is computed in the same manner with A_{CI} replacing $A(\underline{q})$. A_{CI} is the transformation from the inertial to the computed (estimated) body orientation. The first term of (4.1) is manipulated the same as in Reference 1, but the second term is expanded in a different manner. This process is outlined below. The matrix A_{CI} can be written as

$$A_{CI} = A_{CA}A_{AI} \qquad (4.3)$$

where: A_{CA} = transformation matrix from the true body to the estimated body system
A_{AI} = transformation matrix from the inertial to the true body system

The matrix A_{CA} can be written as

$$A_{CA} = I - [\underline{\alpha}x] \qquad (4.4)$$

where $\underline{\alpha}$ are defined in (2.13). Using (4.4) in (4.3) gives

$$A_{CI} = (I - [\underline{\alpha}x])A_{AI} \qquad (4.5)$$

Substituting (4.5) into (4.1) with the expansion of the first term of (4.1) from Reference 1 gives

$$\underline{y} = [\underline{\alpha}x]A_{AI}\underline{V}_I - M_{AT}[\underline{W}_{T'},_{meas}x]\underline{\Theta} + M_{AT}d\underline{W}_{T'} \qquad (4.6)$$

Note that $\underline{V}_A = A_{AI}\underline{V}_I$ but since we don't know A_{AI} we use the estimated matrix, i.e. $\underline{V}_c = A_{CI}\underline{V}_I$, therefore we can write (4.6) as

$$\underline{y} = [-\underline{V}_c x]\underline{\alpha} - M_{AT}[\underline{W}_{T'},_{meas}x]\underline{\Theta} + M_{AT}d\underline{W}_{T'} \qquad (4.7)$$

The models for each of the sensors will now be given. From these models, the H matrix for each sensor is obtained. The derivation is shown in Reference 1.

FSS:

$$\underline{y} = \begin{bmatrix} & 0 \ldots 0 & & & & 0 \ldots 0 & \\ H_q & 0 \ldots 0 & M_{AT}[\underline{W}_{T'},_{meas}x] & M_{AT}W_s & \tan A \quad 0 & M_{AT}W_s & \underline{x} + M_{AT}W_s \begin{bmatrix} n_A \\ n_B \end{bmatrix} \\ & 0 \ldots 0 & & & 0 \quad \tan B & 0 \ldots 0 & \end{bmatrix} \qquad (4.8)$$

IR:

$$\underline{y} = \begin{bmatrix} & 0 \ldots 0 & & 0 \ldots 0 & \\ H_q & 0 \ldots 0 & W_h & 0 \ldots 0 & \underline{x} + W_h \begin{bmatrix} n_{hr} \\ n_{hp} \end{bmatrix} \\ & 0 \ldots 0 & & 0 \ldots 0 & \end{bmatrix} \qquad (4.9)$$

Magnetometer:

$$\underline{y} = \begin{bmatrix} & 0 \ldots\ldots\ldots 0 & & & \begin{bmatrix} n_{xm} \\ n_{ym} \\ n_{zm} \end{bmatrix} \\ H_q & 0 \ldots\ldots\ldots 0 & M_{AT}B' & \underline{x} + W_h & \\ & 0 \ldots\ldots\ldots 0 & & & \end{bmatrix} \qquad (4.10)$$

In 4.8, 4.9, and 4.10 $[-\underline{V}_c x]$ replaces H_q in the multiplicative filter.

V. QUATERNION NORMALIZATION

The quaternion that represents attitude is normal. Reference 3 shows that forcing normalization on the

estimated quaternion is advantageous since it speeds up convergence and eliminates the need for filter tuning. Reference 1 describes the forced normalization of the additive filter. That normalization technique is equivalent to removing a portion of the estimate. That method of normalization is external to the EKF algorithm. A new normalization method was derived that blends naturally into the Kalman filter algorithm. The idea is to use the normalized quaternion as a "pseudo-measurement". Performing a measurement update with very small noise (ideally zero) on the "measurement" then pushes the quaternion portion of the state to be the normalized quaternion. The algorithm is

$$\underline{z} = \frac{\hat{\underline{q}}_k(+)}{|\hat{\underline{q}}_k(+)|} \qquad (5.1)$$

where: $\hat{\underline{q}}_k(+)$ = the upper four elements of $\hat{\underline{X}}_k$ in (2.9)

The "measurement" covariance matrix can then be defined as:

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \delta \qquad (5.2)$$

where: δ = a very small dimensionless number

The measurement matrix, H, is defined as

$$H = \left[\begin{array}{cccc|ccc} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \end{array}\right] \qquad (5.3)$$

The gain is computed according to (2.4). The state is updated as

$$\hat{\underline{X}}_k^*(+) = \hat{\underline{X}}_k(+) + K[\underline{z} - H\hat{\underline{X}}(+)] \qquad (5.4)$$

The covariance is updated with (2.5). This measurement update process is according to a linear Kalman filter and not the extended Kalman filter since it handles the state $\underline{X}$ directly in the update and not the error state, $\underline{x}$.

The quaternion is also normalized in the multiplicative EKF for the same reasons as the additive EKF. The normalization process used was a forced normalization. The quaternion was simply normalized with no compensation performed.

VI. COMPENSATION

When propagating the state estimate and the covariance, we use the measured angular velocity. We know, however, that the propagated values are not accurate because the gyro output contains errors. We can better estimate those errors if we correct the gyro output for estimated errors. This operation is known as calibration.

We also want to compensate the measurements obtained from the FSS, the IR horizon scanner, and the magnetometers which are all orientation measuring devices whose output are used to update the filter. The reason we want to compensate the output from these sensors is different in nature than the reason for compensating the gyro output. In (2.11) we estimate the difference between the true value of $\underline{X}$ and its latest estimate, and add the estimate of the difference to the latest estimate of $\underline{X}$ to form its updated estimate. Let us consider an error term in one of the sensor measurements (say a bias). This bias is a part of $\underline{W}_{T',meas}$ and, thus, as indicated in (4.1) bears its signature on $\underline{y}$. Consequently, if certain observability conditions are met, it is estimated and added to the state estimate as indicated in (2.11). If no compensation takes place, the next time the measurements of this sensor are processed the bias is again estimated and added to the previous estimate of this bias, thus creating an estimate that is too large and incorrect. The proper way to handle this case is to eliminate the estimate of the bias from $\underline{W}_{T',meas}$ so that only the residual bias, which has not been estimated yet, is present in $\underline{y}$ as shown in (4.1). Only the

estimate of this residual is added to the existing estimate of the bias, which is a part of $\hat{X}$, yielding a correction to the previous estimate. This logic holds for the other error states, too. Reference 1 outlines this method. The compensation was applied to the multiplicative EKF in the same manner.

VII. SINGULARITIES

It was found that the models for both the additive and multiplicative EKFs presented in Section IV contain singularities. The IR and FSS measure only two quantities (direction). Artificial generation of a vector measurement based on the IR and FSS measurements constitutes a projection of two-dimensional information on a three dimensional space. Such projection yields a singular noise covariance matrix. Another singularity exists in the additive EKF. This is because only three parameters are necessary to specify attitude. Adding a fourth parameter results in the computation of HPH^T being singular.

To compensate for the first singularity, the noise covariance matrices for the FSS and IR were reduced to 2x2 matrices. This forced the elimination of the third row of H. Note that the magnetometer measures three independent components, and no alterations of its noise covariance and H matrices are necessary. In the additive EKF, the singularity of HPH^T is not removed by this operation. The nonsingular R keeps the computation of $(HPH^T + R)$ in equation (2.4) from becoming singular.

Several tests were run to verify that the singularities were eliminated. It was found that as the uncertainties in the measurement noises were reduced on the FSS and IR, the singular HPH^T in the additive EKF cannot be compensated for by the noise covariance matrices. The uncertainties in the FSS and IR measurement noises were kept at 0.01 degrees or higher to avoid this. It was found in the multiplicative EKF that near singular conditions can exist initially with a large initial attitude error. The filter overcomes the singularity as the state is updated. This singularity exists because the assumption given by (2.14), that the first three elements of $\delta\underline{q}$ are small, is violated.

VII. RESULTS

Several scenarios were run with both filters to study the characteristics of the two filters. Simulated data were used for most of the tests. The simulated data had an attitude solution of 0 degrees yaw, roll, and pitch. The x, y, and z gyro bias were approximately -1.7, 1.2, and 1.5 deg/hour during the first portion of the orbit. During the latter portion of the orbit when the FSS had coverage, the y bias was approximately -2.6 deg/hour for a short time period (around 200 seconds) and then it changed back to 1.2 deg/hour. (The y bias flipped in the simulated data due to the orbit eccentricity.) Other than the gyro bias the data were clean with no noise. Other sensor calibrations were studied by applying errors to the different sensors and seeing how well the two filters could estimate these errors. Real data were also used for some of the tests and the results will be presented after the simulated data test results. When possible, results from the batch estimator are also provided. The quaternion given in (2.9) defines the attitude in inertial coordinates. It was converted to geodetic pitch, roll, and yaw solutions for display and comparison to the batch solutions.

Simulated Data

The first study performed was a sensitivity study of the gyro bias to determine the best value for the a priori covariance. The a priori covariance was varied to determine what value gave the lowest error in the attitude and gyro bias solutions. As mentioned in Section III the gyro bias was modeled as a Markov process. It was found in this scenario that, since the bias does not change, both filters behaved better when the time constant on the Markov was set very high. This essentially models the bias as a constant. The time constant was set to $1x10^5$ seconds. The results of the sensitivity study showed that the best a priori value for the

Table 1. A Priori Values for Both Filters

Attitude - specified for run	P_o other states - 0 unless specified	FSS measurement σ- 0.01 degrees
Other states - 0	Q attitude - $2x10^{-9}$ rad²/sec²	IR measurement σ- 0.01 degrees
P_o attitude (quaternion) - 0.0625	Q Markov - see Section III	Magnetometer measurement
P_o gyro bias - $1x10^{-7}$ rad²/sec ²	time constants - $1x10^5$ sec	σ- 3 milliGauss

gyro bias covariance was $\sigma^2=1x10^{-7}$ rad^2/sec^2. All other a priori values are given in Table 1. These values were used in all runs except where noted.

The first case examined was a baseline case, solving only for attitude. The a priori attitude error was 50 degrees. Figures 1a and 1b show the yaw solutions from the additive EKF and batch algorithm. The multiplicative EKF yaw solution has the same shape as the additive EKF with the same final value. All three sensors were used in the estimation. Both filters show similar behavior, converging quickly. The batch solution converged onto the wrong solution. The a priori error of 50 degrees was too high. The filter pitch and roll solutions converge very quickly, reaching values of 0.008 deg for roll and -0.001 degrees for pitch (both filters). The batch epoch solution for roll was -2.8 degrees and 4.5 degrees for pitch. The sensor residuals for the two filter solutions were very small, but not for the batch due to the large error in the solution.

The next case involved adding the estimation of gyro bias to the baseline run. The attitude could not be given such a large error in order for the solutions to converge. The a priori error was set to 10 degrees (the limit was around 20 degrees). The yaw solutions for the two filters were similar to those shown in Figure 1. The batch algorithm converged for all angles with final values similar to the filters. The pitch and roll solutions also converge very quickly for the filters with nominal final results. Figures 2a and 2b show the estimation of the gyro bias for the two filters (in the span of data used the FSS had coverage so the y gyro bias was approximately -2.6 deg/hr). The batch epoch gyro bias solution is listed on the figure as well. The batch algorithm gives a slightly better estimate of the gyro bias. The filter, in general, needs more than 100 seconds (typically 200 or more) to converge to the gyro bias solution. After 200 seconds the additive and multiplicative filters give

Additive (deg/hr)	Multiplicative (deg/hr)	Batch (deg/hr)
x = -1.22	x = -1.54	x = -1.7
y = -1.71	y = -2.16	y = -2.6
z = 2.25	z = 1.40	z = 1.5

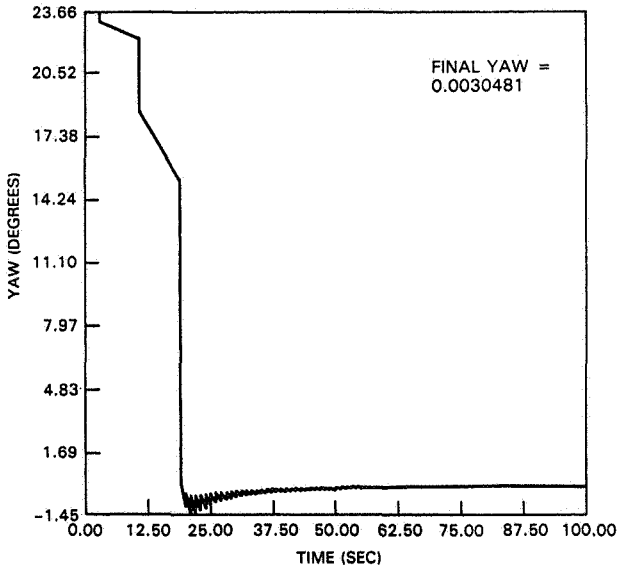
The multiplicative shows a little quicker convergence to the gyro bias solution. Both converge beyond 200 seconds and remain stable.

The next iteration in the baseline study examined the effects of normalization. The pseudo-measurement normalization technique outlined in Section V was implemented into the additive filter. It was found that it did not give good results when the noise covariance matrix was quite small. In this case the estimated quaternion was almost completely replaced by the normalized quaternion and the covariance matrix converged to a very small value. The filter then goes to the normalized quaternion with huge confidence but the normalized quaternion is not necessarily the correct quaternion. Table 2 shows the results at 100 seconds for various values of δ, the diagonal element of the noise covariance matrix (see Equation 5.2). Table 2 also gives the result from the baseline case above, starting with an a priori error of 10 degrees and estimating gyro bias. The pseudo-measurement normalization has the most effect on the pitch solution. As δ is decreased, the pitch solution eventually diverges beyond what is given in Table 2.

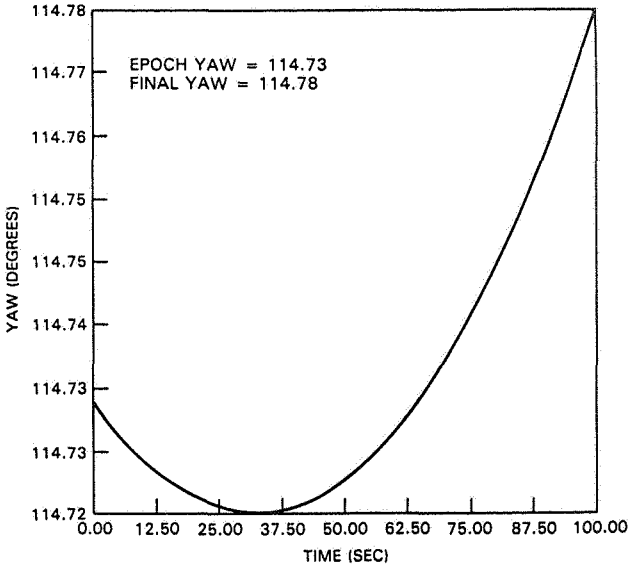
Table 2. Pseudo-Measurement Normalization Results for Different Noise Levels (A priori attitude = 10 degrees, gyro bias estimated)

δ	Yaw (deg)	Roll (deg)	Pitch (deg)
$1x10^{-5}$	0.0031	0.0077	-0.0073
$1x10^{-7}$	0.0030	0.0075	-0.0073
$1x10^{-9}$	0.00037	-0.0078	0.084
baseline	0.0039	0.0078	-0.00058

Another characteristic of the pseudo-measurement technique is the increased use of computer processing time. It took approximately 15 percent more computer time with the pseudo-measurement technique than the original technique to process 1000 seconds. (This is due to the inversion of a 4x4 matrix.) Normalization is not very critical for attitude determination in this scenario. The attitude solutions without normalization and with the pseudo-measurement normalization technique are similar to those in the baseline cases discussed above. Figures 3a, 3b, and 3c show the gyro bias estimates from the additive and multiplicative EKFs with no

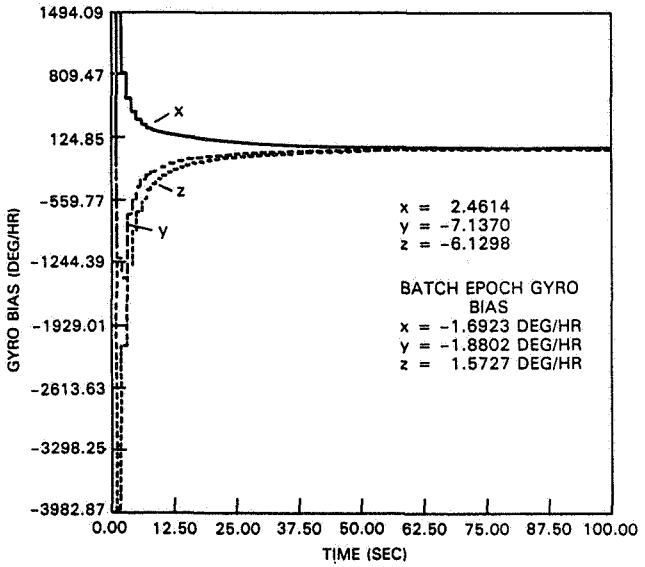


1a - ADDITIVE

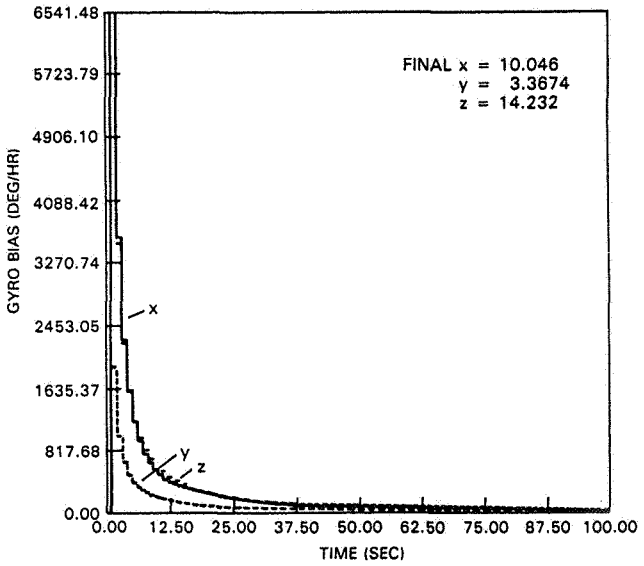


1b - BATCH

FIGURE 1. BASELINE YAW SOLUTIONS (APRIORI ATTITUDE = 50 DEG.)

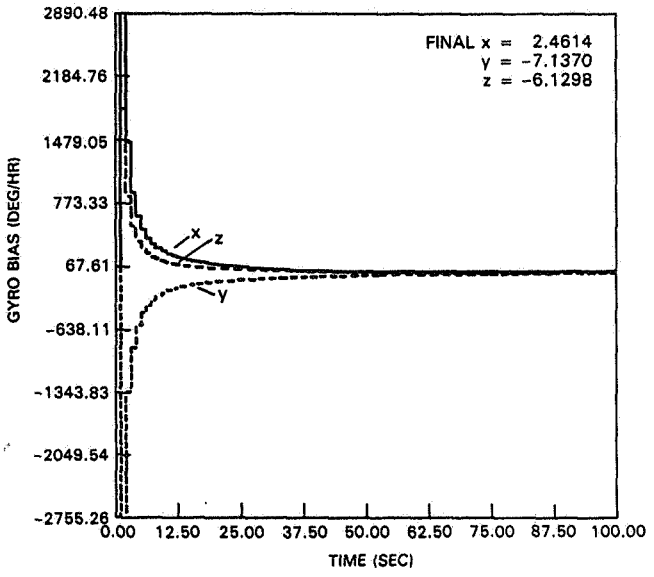


2a - ADDITIVE

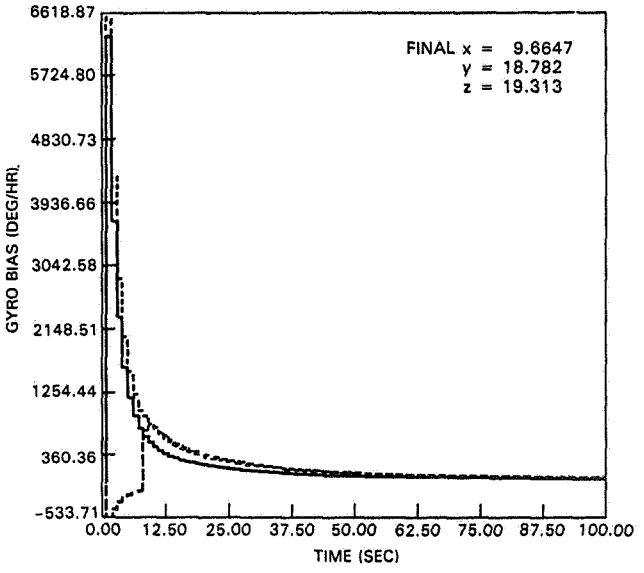


2b - MULTIPLICATIVE

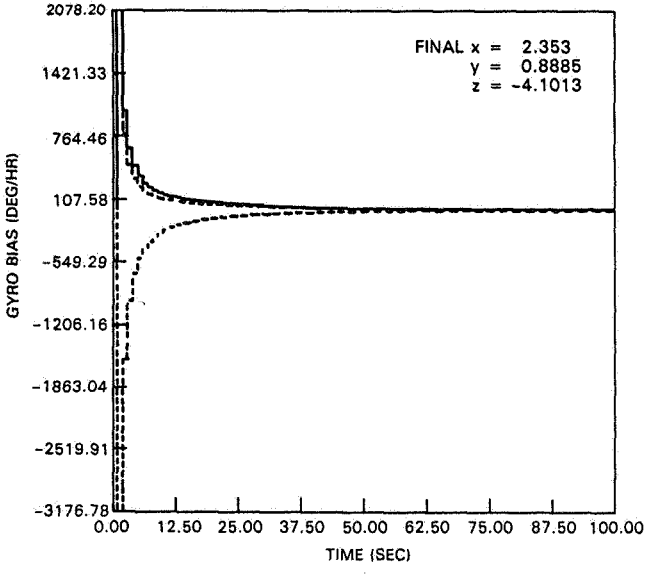
FIGURE 2. GYRO BIAS ESTIMATE (APRIORI ATTITUDE = 10 DEG.)



3a - ADDITIVE, NO NORMALIZATION



3b - MULTIPLICATIVE, NO NORMALIZATION



3c - ADDITIVE, PSUEDO-MEASUREMENT NORMALIZATION

FIGURE 3. GYRO BIAS ESTIMATES (APRIORI ATTITUDE = 10 DEG)

normalization and from the additive EKF with the pseudo-measurement normalization. The gyro bias is affected somewhat initially, but eventually the solutions converge. It was mentioned previously that Reference 3 shows normalization eliminates the need for filter tuning. Even though it is not crucial here it would be expected to have an effect similar to that found previously in other scenarios, perhaps cases that are not as nominal.

The next area examined was a yaw maneuver. The yaw was changed from 0 to 180 degrees at a constant rate. The gyro biases are different in this case. The values for x, y, and z start out with the original biases but after the maneuver they change to approximately 2.1, 1.2, and 1.6 deg/hr, respectively (due to geometry). Again attitude and gyro bias were estimated. This time an upper limit of 20 degrees initial attitude error was used. Figure 4a shows the yaw solution from the multiplicative EKF. The additive EKF and batch yaw estimates have the same shape with a final yaw values of 179.91 and 180.77 degrees. Both filter estimates and the batch estimate follow the maneuver very closely. The pitch and roll solutions exhibit nominal behavior and the residuals are extremely small. Figures 4b and 4c show the gyro bias estimates for both filters. After 100 seconds they have not quite converged due to the larger initial attitude error. At the end of the run (approximately 3000 seconds) the gyro bias has become quite stable with the values given below. The batch epoch gyro bias is also given below (using approximately 3000 seconds of data as well).

Additive (deg/hr)	Multiplicative (deg/hr)	Batch (deg/hr)
x = 1.712	x = 1.719	x = -0.442
y = 1.988	y = 1.467	y = 0.442
z = 1.446	z = 1.475	z = -0.050

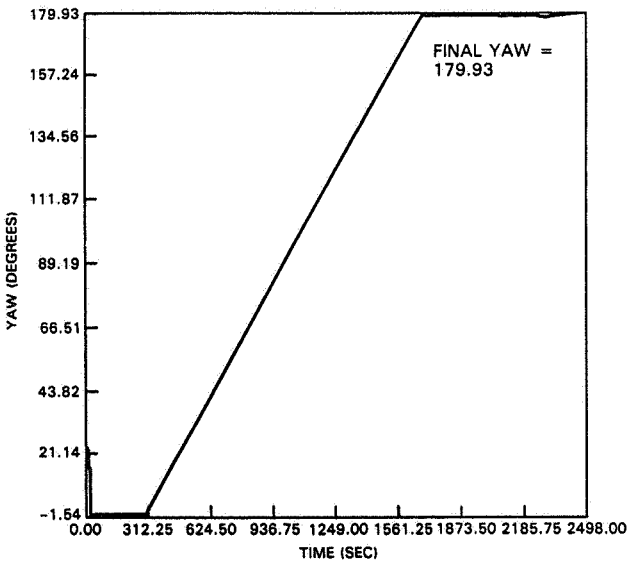
The batch cannot follow a change in the gyro bias since it gives only one solution at the epoch. The solved for biases are influenced by the initial bias and the bias after the maneuver and thus are somewhat between the two.

In the following cases both filters were used to study the characteristics of using different sensor combinations, as opposed to using data from all three sensors concurrently. The combinations were IR/MAG, FSS/MAG, IR only, and MAG only. In all cases both filters showed poor estimation of gyro bias, particularly in those cases with magnetometer data. The magnetometer, which suffers from a digitization of 6 milliGauss, is too coarse to estimate gyro bias. This digitization results in the magnetometer having only coarse attitude estimation ability. Without the availability of a fine attitude solution the gyro bias is not observable.

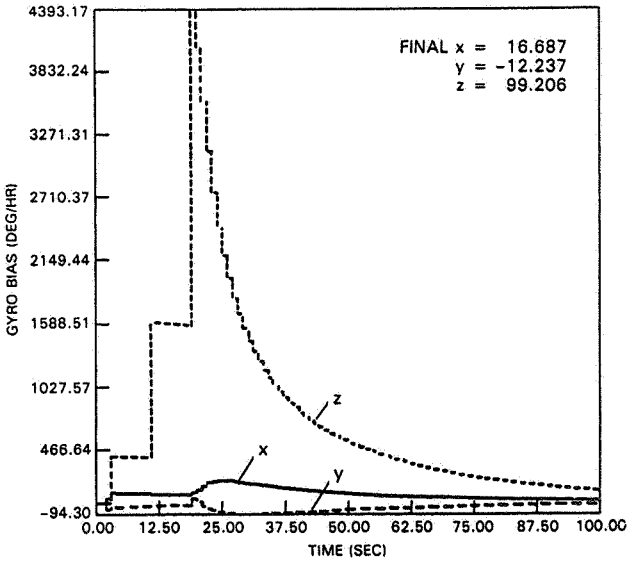
Figures 5a and 5b show the estimation of yaw by the additive EKF and the batch algorithm for an IR/MAG combination. The multiplicative EKF yaw solution again had the same shape as the additive EKF with a final yaw of -0.664 degrees. The a priori attitude was set to 0 degrees and the gyro bias was estimated. All show similar values for attitude. It was found when using two sensors, one being the magnetometer, that better results were achieved when the FSS or IR measurement uncertainties were increased to 0.1 degrees. At 0.01 degrees there was too much disparity between the uncertainties and the filter exhibited more fluctuations. Figure 5c shows the gyro bias estimation from the additive EKF with the batch epoch solution listed. The multiplicative gyro bias estimates look like the additive with final x, y, and z values of -4.413, 0.852, and -2.090 deg/hr, respectively. The filter solutions show quite a bit of fluctuation. Between the three algorithms the final results are similar for x and y, but differ for z; the batch algorithm giving a much better estimate. The x value for all three algorithms and the z for the filters is not estimated very well due to the magnetometer inaccuracy. The batch algorithm also weights the sensors differently which could account for the differences in the final solution. The coarse estimation of yaw by the magnetometer (yaw is not observable in the IR sensor) is shown in Figure 5a. These errors corrupt the gyro bias solution. The pitch and roll behave nominally since they are estimated mainly by the IR sensor. The IR residuals are also very small. A sample of the magnetometer residual is also shown in Figure 5d. The digitization is apparent in the residual curve. Also shown are the expected values (plus and minus) of the residual (dashed lines). The magnetometer residual settles near these values.

An attempt was made to improve the gyro bias variation by increasing the uncertainty on the magnetometer from 3 to 50 milliGauss. This did not improve the results at all.

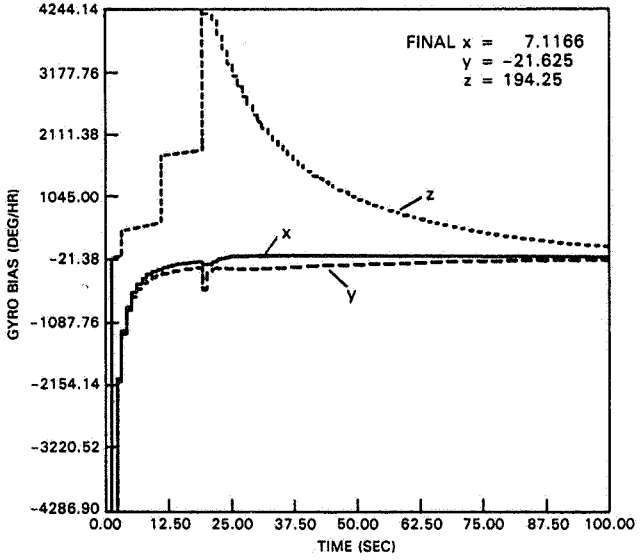
Without the gyro bias estimation, the IR/MAG solutions will converge from a much larger initial attitude error. Figures 5e and 5f show the yaw solutions for the multiplicative EKF and the batch algorithm with a 50 degree initial attitude error. The additive EKF yaw solution has the same shape as the multiplicative EKF with a final yaw of -0.123 degrees. All three algorithms have final solutions that are quite similar. The



4a - YAW, MULTIPLICATIVE

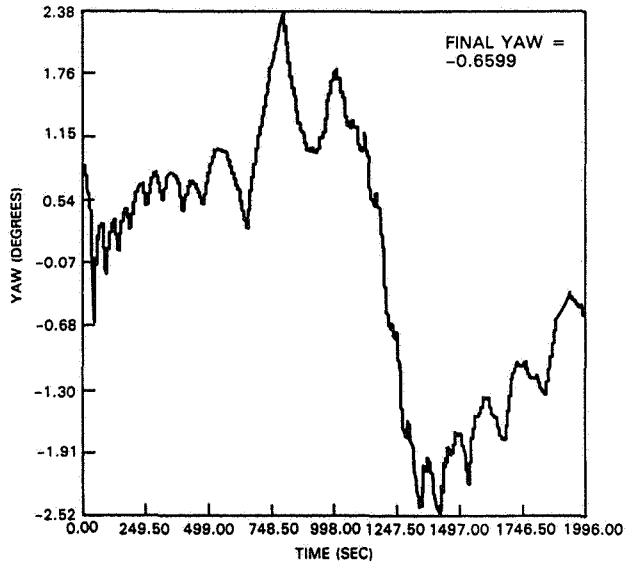


4b - GYRO BIAS, ADDITIVE

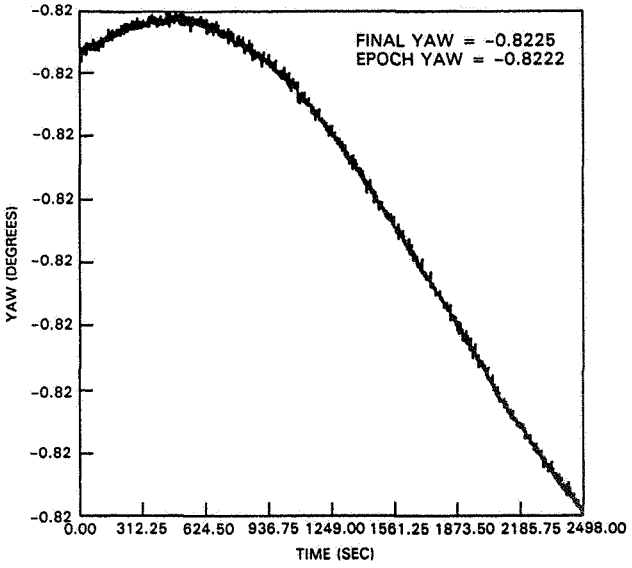


4c - GYRO BIAS, MULTIPLICATIVE

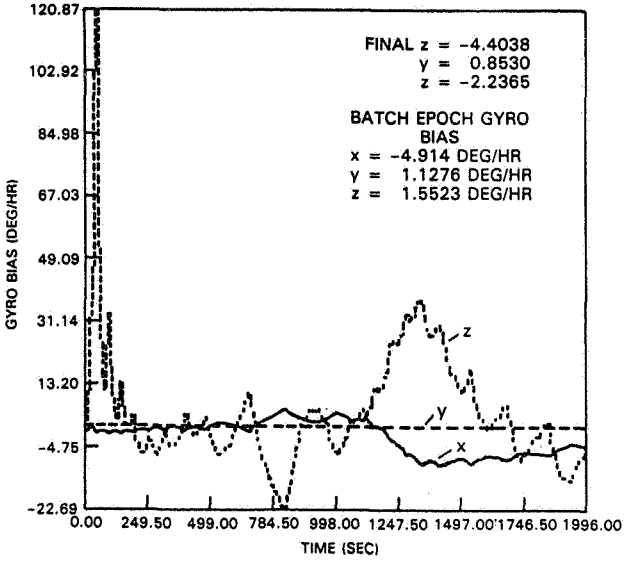
FIGURE 4. YAW MANEUVER (APRIORI ATTITUDE = 10 DEG)



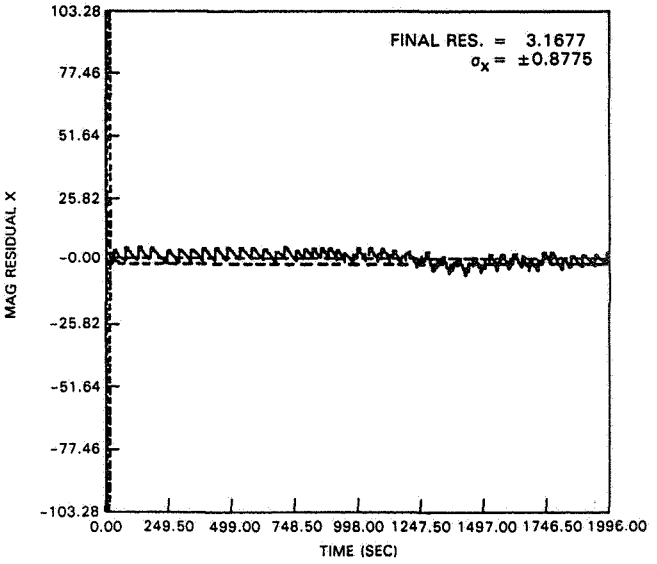
5a - YAW, ADDITIVE



5b - YAW, BATCH

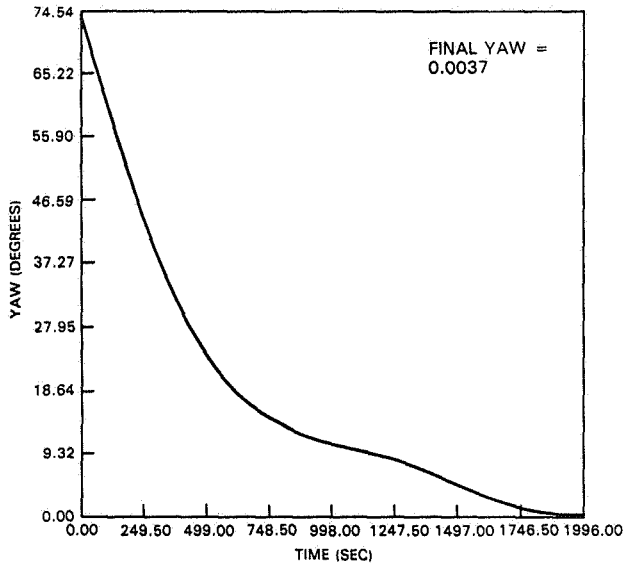


5c - GYRO BIAS, ADDITIVE

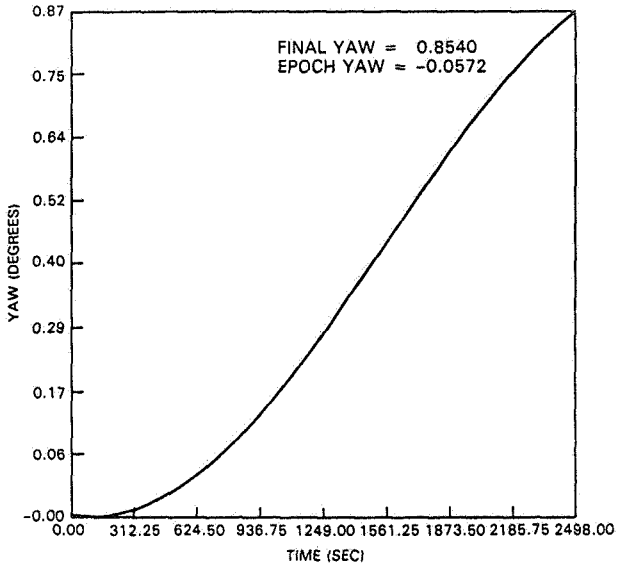


5d - MAGNETOMETER X RESIDUAL

FIGURE 5. IR/MAG SOLUTIONS (APRIORI ATTITUDE = 0 DEG)

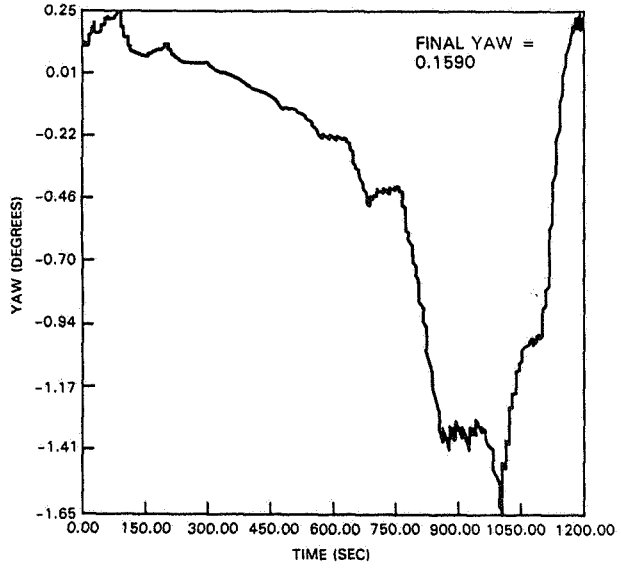


5e - YAW, MULTIPLICATIVE (APRIORI ATTITUDE = 50 DEG., NO GYRO BIAS)

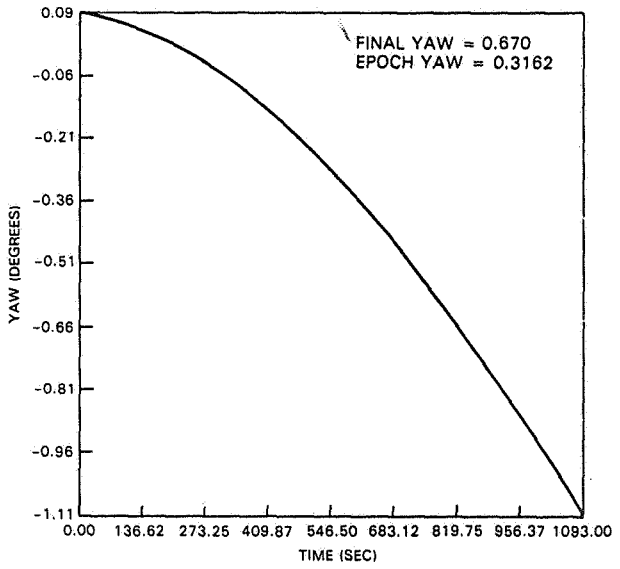


5f - YAW, BATCH (APRIORI ATTITUDE = 50 DEG., NO GYRO BIAS)

FIGURE 5. IR/MAG SOLUTIONS



6a - YAW, ADDITIVE



6b - YAW, BATCH

FIGURE 6. FSS/MAG SOLUTIONS (APRIORI ATTITUDE = 0 DEG)

pitch and roll solutions are nominal. The IR residuals are very small and the magnetometer residuals show similar behavior to that shown in Figure 5.d.

As in the IR/MAG case, solutions using an FSS/MAG combination were generated. Again the gyro bias was poorly estimated for the same reasons stated above. The algorithms were again started with a 0 degree initial attitude. Figures 6a and 6b show the additive EKF and the batch yaw solutions. The multiplicative EKF yaw solution again has the same shape as the additive but with a final yaw of 0.162 degrees. Both filter solutions exhibit quite a lot of fluctuation. The FSS loses coverage at 980 seconds. This causes more divergence in the solution. The roll solution exhibits similar behavior, with considerable fluctuation (the roll also shows this behavior because the FSS does not directly measure the roll as the IR sensor does). The final roll values are 1.64, 1.63, 0.7 degrees for the additive, multiplicative, and batch algorithms, respectively. The pitch shows a slightly more nominal behavior. Figure 6c shows the additive gyro bias estimates. The multiplicative gyro bias estimates look the same with final x, y, and z values of -5.786, -0.582, and -26.219 deg/hr. The batch again gives a much better estimate of gyro bias for z than the filters do. The gyro bias z solutions for the filters are better before the loss of the FSS. Near 750 seconds, the multiplicative gives a gyro bias solution of -1.3, 1.7, 1.3 deg/hour. The additive filter is similar. Once the FSS is lost, though, the filter is relying solely on the magnetometer. The residuals in this case are similar to previous results.

As in the case of the IR/MAG solutions, an attempt was made to stabilize the gyro bias estimation by increasing the measurement noise on the magnetometer from 3 to 50 milliGauss. Figure 6.d shows this result for the additive filter (the multiplicative filter is similar). In this case, the gyro bias is improved by increasing the magnetometer measurement noise. The final results are much closer to the actual values. The x gyro bias estimate is now better than the batch estimate. This technique may work better in the FSS/MAG combination since the FSS gives some estimate of the yaw solution, and thus the problem of relying only on the magnetometer for yaw estimation is avoided.

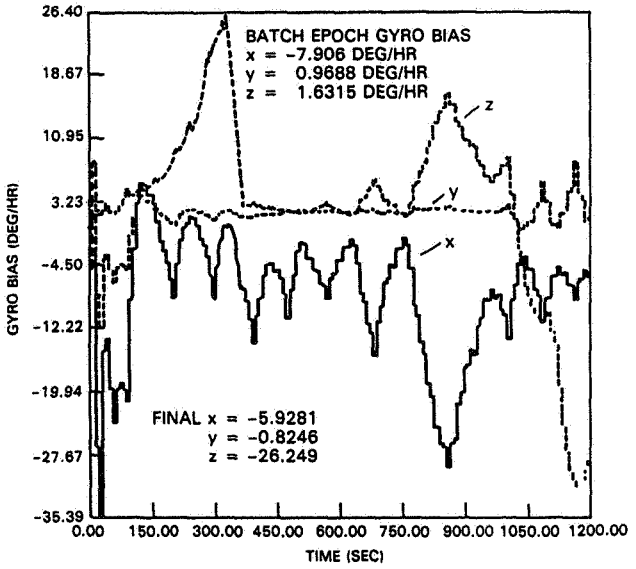
As in the case of the IR/MAG solutions, without the estimation of gyro bias the filters converge from much larger initial attitude errors. The yaw solutions are shown in Figures 6e and 6f for the additive and the batch. The multiplicative is similar with a final yaw of -0.647 degrees. The roll solutions converge to final values of 1.94, 1.19, and 0.46 degrees for the additive, multiplicative, and batch algorithms, respectively. The pitch solutions show a nominal behavior. For the additive filter a 50 degree initial attitude error was used, for the multiplicative and the batch algorithms the maximum initial attitude error was 30 degrees. The solutions converge quickly and the three algorithms give similar solutions.

The final two cases investigate the use of only one sensor. Figures 7a and 7b show the multiplicative EKF and the batch yaw solutions using only magnetometer data. The additive EKF yaw solution is similar to the multiplicative with a final value of -1.254 degrees. The a priori attitude error was set at 50 degrees and gyro bias was not estimated. All three converge to less than 1 degree (the additive eventually converges this far as well) error using approximately a half orbit of data. The pitch and roll solutions show similar results with the pitch solution converging within approximately 300 seconds. Figure 8 shows the yaw solution from the multiplicative filter using only IR data. The additive filter and batch algorithm did not converge using only IR data with an initial attitude error of 50 degrees. Figure 8 shows that the yaw solution eventually converges with only IR data due to quarter orbit coupling (related to the geometry of the ERBS attitude). The pitch and roll solutions converge within 10 seconds to values less than 0.05 degrees.

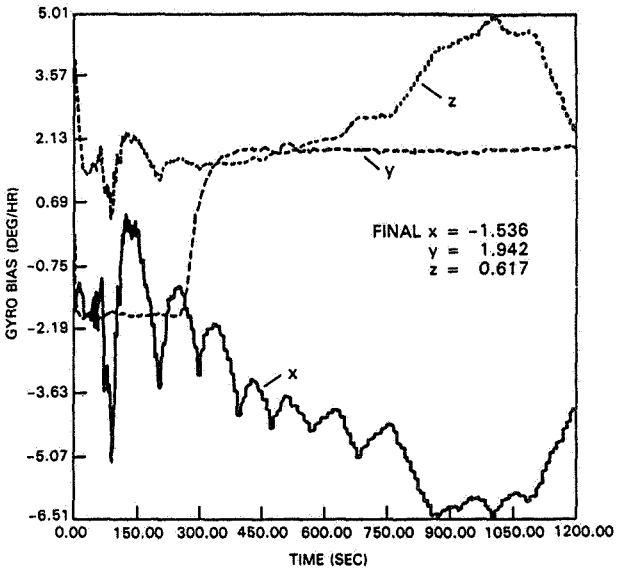
Sensor Calibration

A further test of gyro bias estimation was performed to determine if the filter could follow a change in the bias. In this case the gyro bias was started at 3.6 deg/hour on all axes, and after 300 seconds it was switched to -3.6 deg/hour. Figure 9 shows the estimation of this gyro bias by the additive filter. The filter follows the change and converges again within 200 seconds. The multiplicative filter exhibits similar behavior. The batch algorithm cannot estimate a gyro bias change since it estimates only a single solution at the epoch (and then the epoch attitude is propagated using the gyro data corrected for the epoch gyro bias).

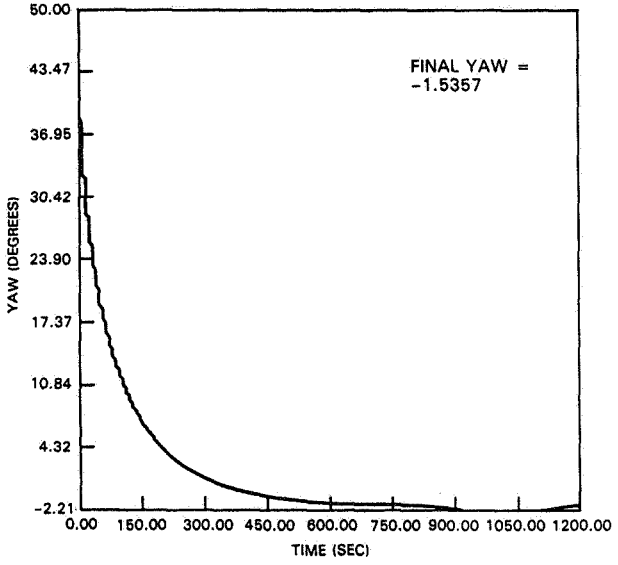
The remaining calibration study focuses on the FSS, IR, and magnetometer. The same clean simulated data were used and biases were applied to the sensors being studied. A priori covariances were selected to be $1x10^{-7}$ (units for particular bias) for biases based on the sensitivity study performed on the gyro bias. In the future, sensitivity studies should be performed on each sensor calibration to determine the best a priori covariance. The first studies involved applying a single calibration error to only one sensor at a time. Several calibration errors were then applied to study the capability of the filters to estimate several



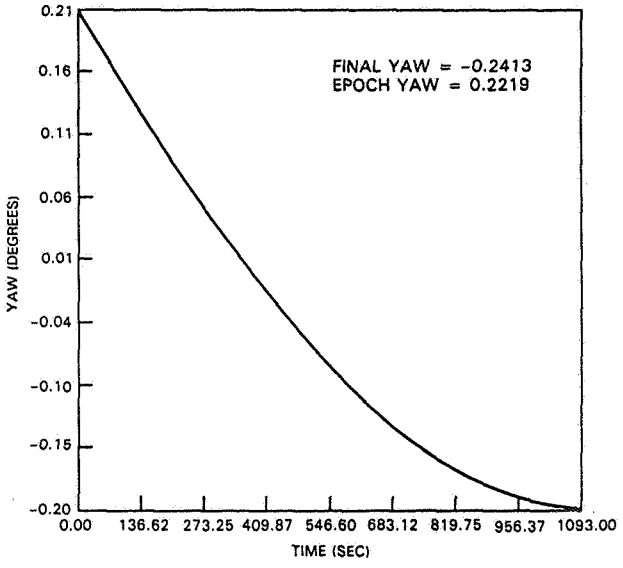
6c - GYRO BIAS, ADDITIVE



6d - GYRO BIAS, ADDITIVE (σ_{mag} = 50 mG)

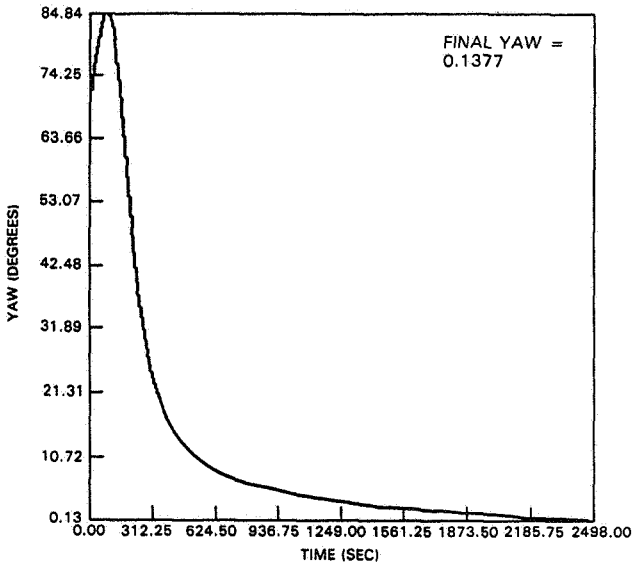


6e - YAW, ADDITIVE APRIORI ATTITUDE = 50 DEG)
(NO GYRO BIAS)

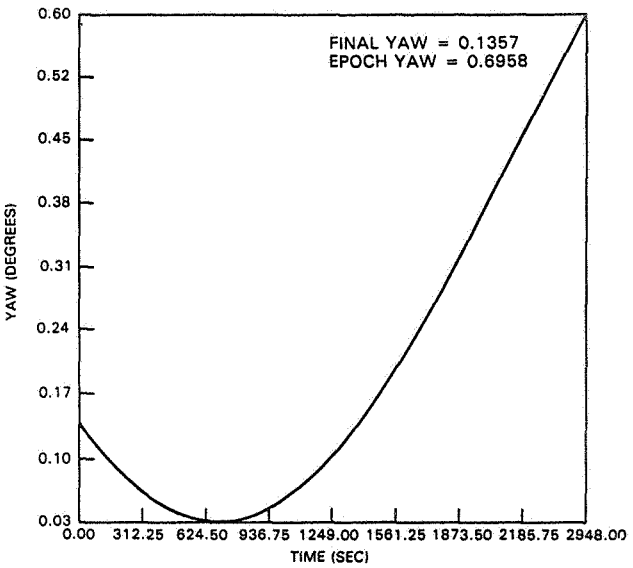


6f - YAW, BATCH (APRIORI ATTITUDE = 30 DEG)
(NO GYRO BIAS)

FIGURE 6. FSS/MAG SOLUTIONS



7a - YAW, MULTIPLICATIVE



7b - YAW, BATCH

FIGURE 7. MAGNETOMETER ONLY (APRIORI ATTITUDE = 50 DEG)

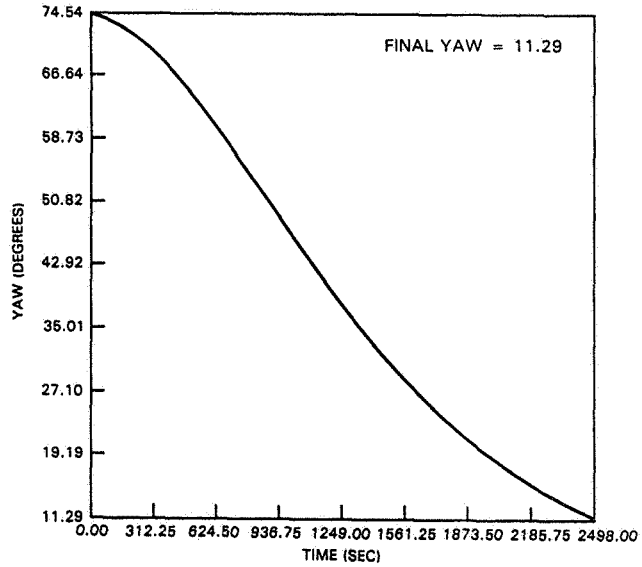


FIGURE 8. IR ONLY, MULTIPLICATIVE
(APRIORI ATTITUDE = 50 DEG)
(NO GYRO BIAS)

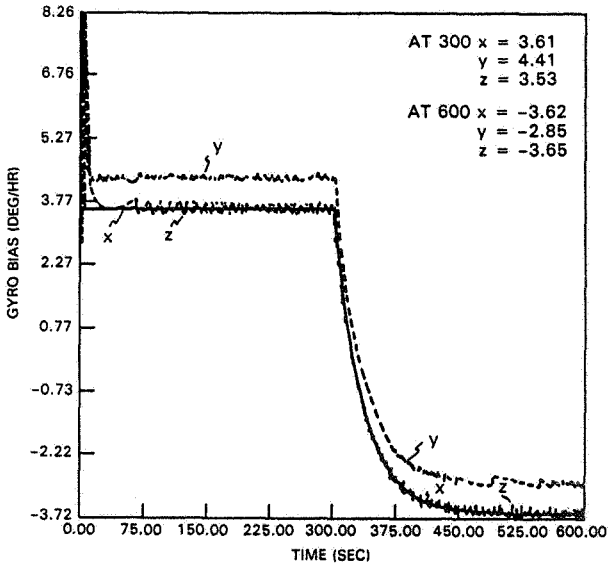


FIGURE 9. CHANGING GYRO BIAS ADDITIVE
(APRIORI ATTITUDE = 0 DEG)

calibration parameters at once.

The first calibration parameter to be studied was the IR bias. Table 3 shows the IR bias for the multiplicative and additive filters. The final entry in the table is the highest correlation coefficient, indicating the least observable component. A 0.1 degree bias was applied to both the measured pitch and roll. Table 3 shows that the multiplicative filter estimates the pitch bias quite well. The roll bias in the multiplicative filter is not observable. It is highly correlated with the first angle in the error state vector due to the geometry in this case. The FSS does not supply enough roll information. With

Table 3. Estimation of IR Horizon Scanner Pitch and Roll Bias (A priori attitude = 0 degrees)

		Bias (deg)		Uncertainty (deg)		Highest
Filter	Length	Roll	Pitch	Roll	Pitch	Correlation
Mult.	200 sec	-0.177	0.103	0.014	0.0006	$(\Theta_1, r)=0.989$
Add.	200 sec	-50.01	3.528	0.003	0.0105	all high

different geometry (the sun in a different location in the FOV or sun in the other FSS) the roll bias should be estimated as well as the pitch bias in the present example. The additive filter does not estimate either the pitch or roll bias well.

The next bias studied was an FSS bias. Table 4 shows the estimated values when a 0.1 degree bias was applied.

Table 4. Estimation of FSS Alpha and Beta Bias (A priori attitude = 0 degrees)

		Bias (deg)		Uncertainty (deg)		Highest
Filter	Length	alpha	beta	alpha	beta	Correlation
Mult.	200 sec	0.208	0.110	0.012	0.0011	$(\Theta_1, \alpha)=0.986$, $(\Theta_3, \alpha)=0.982$
Add.	200 sec	53.38	124.3	0.0005	0.0001	$(q_1, \alpha)=0.873$

Again the multiplicative filter estimated one of the biases (beta) quite well. With a longer run, the beta bias converges to approximately 0.1 degrees. The alpha bias is not observable again due to geometry. The additive filter shows poor estimation without significant observability problems as reflected by the correlation matrix.

The next bias estimated was magnetometer bias. A 10 milliGauss bias was applied on all axes. Table 5 shows that both filters estimate the bias quite well.

Table 5. Estimation of Magnetometer X, Y, and Z Biases (A priori attitude = 10 degrees, gyro bias estimated)

		Bias (mG)			Uncertainty (mG)			Highest
Filter	Length	X	Y	Z	X	Y	Z	Correlation
Mult.	200 sec	13.46	12.87	13.68	0.173	0.173	0.173	small
Add.	200 sec	13.48	12.90	13.72	0.173	0.173	0.173	small

The correlations were all very small. The estimated biases are all within 6 milliGauss (the resolution of the magnetometer) of the applied biases.

FSS misalignment was the final calibration error studied with the multiplicative filter. (The additive filter was not studied because of its poor performance estimating the previous calibration errors.) A 0.1 degree misalignment was applied to the FSS x and y axes and a 0.05 degree misalignment was applied to the FSS z axis. The misalignment uncertainty was 0.18 degrees. After 200 seconds the estimated misalignments were -0.070, -0.22, -0.51 degrees for the FSS x, y, and z axes, respectively, with uncertainties of 0.032, 0.043, and 0.093 degrees. The x and z misalignments are highly correlated to one another (the correlation coefficient was 0.998) but the three misalignments are not highly correlated to the attitude. Next, the filter was run with the same misalignments applied, but alpha and beta biases were solved for instead of

misalignments. This revealed that the filter could not distinguish the misalignments from biases. The beta bias, which corresponds to a y misalignment, was estimated well but the alpha bias was not since it is not observable at this attitude (this means that the x and z misalignments are also not observable). The estimated biases after 200 and 400 seconds were

α = 0.12	β = 0.09	σ = 0.005 degrees	(200 seconds)
α =-0.01	β = 0.09	σ = 0.0006 degrees	(400 seconds)

Data from different attitudes would be necessary to estimate the alpha bias and also to try and estimate the misalignments and distinguish them from biases.

The multiplicative filter was set up to solve for several calibration errors in one run. The estimated components consisted of attitude, gyro bias, FSS y misalignment, FSS beta bias, IR pitch bias, and magnetometer bias. The sensor errors applied were the same as those used previously. After 400 seconds the filter gave estimates of gyro and magnetometer bias like those given above (in the nominal cases). The filter did not give good estimates of the other calibration errors. With all the errors combined the FSS y misalignment and beta bias and the IR pitch bias were not observable, even though the beta and pitch biases were observable to the filter when applied alone. The geometry in this case does not give enough information to solve for all the parameters together. Again, data from several attitudes should be used and perhaps the state should be kept smaller when performing sensor calibration. It would be necessary to iterate to solve for all the sensor errors.

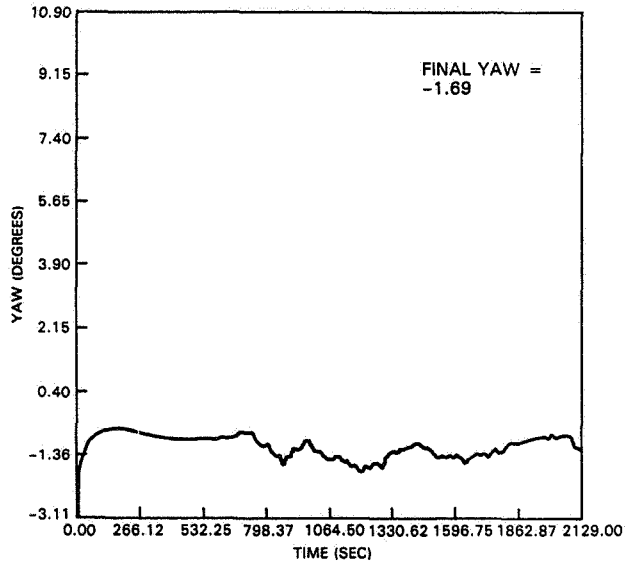
The gyro scale factor and misalignments were not estimated because a single attitude would not give sufficient observability. Attitude maneuvers would be necessary. This is also true for the FSS scale factor. The magnetometer misalignment or scale factors were not estimated in this initial study as they are likely not to be observable with the coarse ERBS magnetometer data.

Real Data

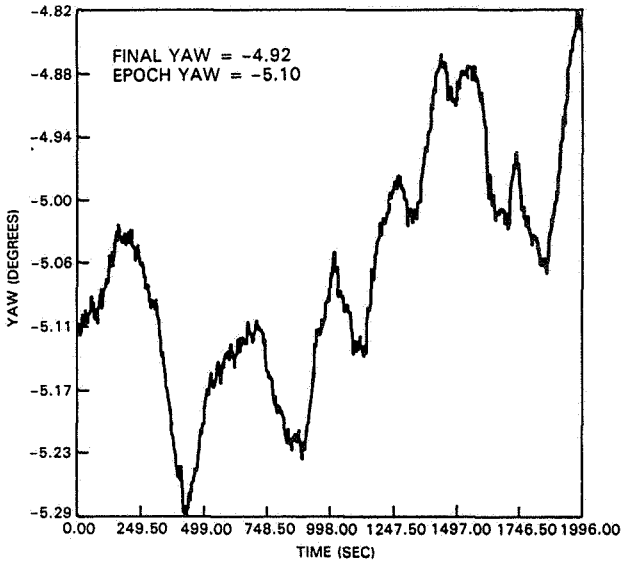
The last test studies the behavior of the filters when using real ERBS data. The orbit of data selected contained approximately 10 minutes of FSS data at the beginning and end of the orbit. Figures 10a and 10b show the multiplicative EKF and the batch algorithm estimates of yaw. The additive EKF yaw solution looks like the multiplicative with the same final value. Figure 10c shows the multiplicative gyro bias estimates. The additive has final x, y, and z gyro bias estimates of -7.393, -4.795, and 1.627 deg/hr, respectively. The roll and pitch solutions look similar to the yaw solution with final values of 0.07, 0.07, 0.10 degrees roll and 0.30, 0.30, 0.30 degrees pitch for the additive, multiplicative, and batch algorithms, respectively. The filters have slightly smaller estimates of yaw, and the three algorithms have similar values for pitch and roll. The gyro bias estimates are similar in y, but x and z are somewhat smaller for the two filters than for the batch (comparing the final filter results with the batch epoch). The two filters exhibit very similar behavior. The residuals for the FSS were considerably smaller for the two filters than for the batch algorithm. The IR residuals were similar and the magnetometer residuals exhibited behavior similar to that shown in Figure 5.d. In the case of the real data, the true reference is not known. The smaller residuals in the filters tend to give those solutions more credibility.

VIII. POTENTIAL APPLICATION

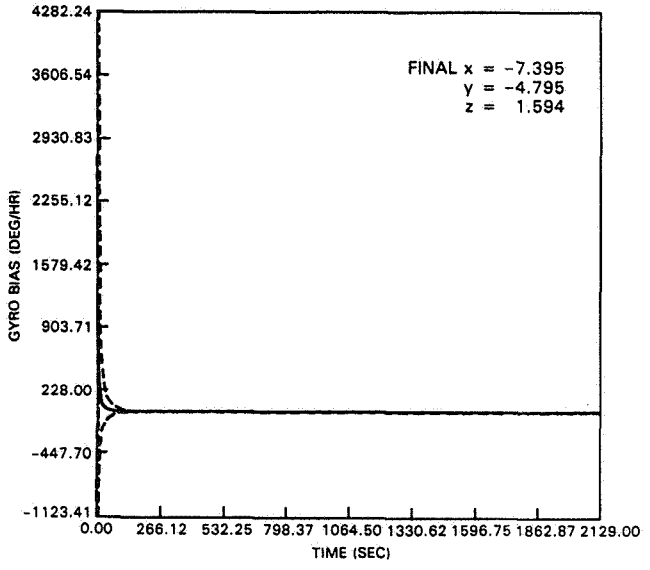
A version of the additive filter was created in which only attitude and gyro bias are solved for (the dimension of Equation 2.9 was reduced to 7). This filter was then tested with simulated ERBS data in a real-time attitude determination system (Reference 5). The filter was able to process the data and generate the attitude and gyro bias solutions shown in Figures 1 and 2 (baseline) in real time (the nominal ERBS data rate is 1 set of measurements every second). Currently, real-time attitude support in the Flight Dynamics Division is performed with single-frame estimators that solve only for attitude. The real-time filter has the capability of giving much more accurate solutions in a relatively short period of time (see Figures 1 and 2 for the length of time to converge). Gyro bias estimates could also be generated when sufficient data coverage is available. This real-time filter gives results comparable to the ERBS batch algorithm, which requires considerably more processing time and memory. This real-time filter is currently being adapted as a prototype system to be used by the GRO. Future missions, such as the SMEX series, are also planning to



10a - YAW, MULTIPLICATIVE



10b - YAW, BATCH



10c - GYRO BIAS, MULTIPLICATIVE

FIGURE 10. REAL DATA
(APRIORI ATTITUDE = 10 DEG)

BATCH EPOCH GYRO BIAS
x = -25.903 DEG/HR
y = -4.408 DEG/HR
z = -3.774 DEG/HR

implement a real-time filter, based on that developed for ERBS.

The full-state filter has potential application for non-real-time support. The initial test results given in Section VII demonstrate that it can give solutions comparable to the batch algorithm (in certain scenarios, others need further investigation as to why the filter solutions differ from the batch). It is also advantageous to combine the attitude estimation with the sensor calibration since the correlation matrix provides valuable information on observability conditions. The combined attitude determination and sensor calibration can be implemented into a batch algorithm. The batch algorithm, though, requires much more memory and has no means of compensating for dynamic noise which can affect an epoch attitude solution propagated over a long period of time.

IX. CONCLUSIONS

In the scenarios presented (using nominal simulated data) both filters, the additive and multiplicative, are very robust in attitude estimation. The filters can be started with large initial errors and still have quick convergence. The batch algorithm is more sensitive to large initial attitude errors in some cases. The filters also exhibit good estimation of gyro bias (when the optimal sensor data are available), although the batch converges with less data. The filters follow changes in attitude and gyro bias closely; the batch algorithm also follows an attitude maneuver closely but it cannot follow changes in gyro bias. When the optimal sensor combinations are not available, the filters must rely on the magnetometer data. The attitude solutions are still estimated well, but the gyro bias is not estimated as well. For ERBS this is a result of the coarseness of the magnetometer data. The batch algorithm does a better job estimating gyro bias in these cases. Further investigation into the weighting is necessary to find ways of improving the filter results (the results would also be improved with a more accurate magnetometer). Both filters also gave reasonable results when using real ERBS data. The filters had slightly smaller residuals than the batch algorithm. Since the true solution is not known, the residuals are the only real figure of merit.

The pseudo-measurement normalization technique is an acceptable normalization method when computer processing time is not critical. When the noise level is selected properly, the pseudo-measurement technique gives results comparable to the original normalization technique. When the noise level is not selected properly, the attitude solutions converge to the wrong value. In the scenario presented this had a significant impact on the pitch solution.

The multiplicative filter has better sensor calibration characteristics than the additive filter. The reasons why are not clear as of this writing. Perhaps the attitude singularity in the additive filter has an affect on the ability of that filter to distinguish and estimate sensor errors. With a single set of data, many calibration errors are not observable. Further study of the sensor calibration with different data spans, giving different sensor coverage, is necessary to fully determine the capabilities of the multiplicative filter. Additional studies should also be performed with sensor corruption, such as noise and data dropout.

REFERENCES

1. Deutschmann, J., and Bar-Itzhack, I.Y., "Extended Kalman Filter for Attitude Estimation of the Earth Radiation Budget Satellite", Flight Mechanics/Estimation Theory Symposium, NASA Conference Publication 3050, May 23-24, 1989

2. Bar-Itzhack, I.Y., and Harman, R.R., "True Covariance Simulation of the EUVE Update FIlter", Flight Mechanics/Estimation Theory Symposium, NASA Conference Publication 3050, May 23-24, 1989

3. Bar-Itzhack, I.Y., and Oshman, Y., "Recursive Attitude Determination from Vector Observations: Quaternion Estimation", IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-21, Jan. 1985, pp.128-136

4. Gelb, A. (ed.), Applied Optimal Estimation, M.I.T. Press, Cambridge, MA 1988

5. Rokni, M., and Deutschmann, J., "Attitude Determination A Prototype Real Time Sequential Filter (RTSF) Evaluation Report", CSC/TM-90/6050, February 1990

Attitude Determination and Calibration using a Recursive Maximum Likelihood-Based Adaptive Kalman Filter

D. A. Kelly, A. Fermelia*
G. K. F Lee†

23 May 1990

Abstract

Kalman filter technniques are widely used in the areas of attitude and orbit determination, prediction and calibration. These techniques work well if the system dynamics are well-defined. Problems arise, however, when the system parameters are unknown ahead of time or changing over time.

This paper discusses an adaptive Kalman filter design that utilizes recursive maximum likelihood parameter identification. At the center of this design is the Kalman filter itself, which has the responsibility for attitude determination. At the same time, however, the identification algorithm is continually identifying the system parameters. The approach is applicable to nonlinear as well as linear systems. This adaptive Kalman filter design has much potential for real-time implementation, especially considering the fast clock speeds, cache memory and internal RAM available today.

The recursive maximum likelihood (RML) identification algorithm is discussed in detail, with special attention directed towards its unique matrix formulation. Next, the procedure for using the algorithm is described along with comments on how this algorithm interacts with the Kalman filter.

Finally, a spacecraft attitude determination/calibration example is provided. In the development of the dynamics for this example, the angular velocity of one of the axis is assumed to be constant. In the simulation, however, this velocity varies slowly. The RML identifier is used to continually identify this changing velocity. This AKF-RML method may be used to identify multiple parameters such as sensor biases or external torques.

*Hughes Aircraft Company, Aurora, Colorado
†Mechanical and Aerospace Department, North Carolina State University

1 Introduction

A problem that has received recent attention is the estimation of systems whose dynamics are unknown or changing. In these situations, it is often neccessary to utilize an estimation algorithm that is able to adjust the system dynamics model automatically. Self-adjusting estimation schemes are known as *adaptive filters*. There are many examples of systems with unknown or changing dynamics, several of which have just recently become more of a concern.

Spacecraft attitude and orbit estimation are two such areas. Many current operational spacecraft do most of the attitude estimation, prediction and calibration on the ground using large mainframe computers. Recent literature, [19] and [23] for example, has suggested on-board attitude prediction and calibration using adaptive filtering techniques. Adaptive filtering methods have already been used for some time in spacecraft orbit estimation [27].

Adaptive filter techniques have also been applied in the steering of ships [28] and has been suggested for use in calibrating ocean navigational gyros [21]. Industrial processes research has led to the development of adaptive filter algorithms for monitoring chemical processes. Chen, Wadhwani and Roberts [3], for example, offer an adaptive filter technique for monitoring changes in raw material composition.

Adaptive control methods often utilize adaptive filters. An extensive amount of literature has surfaced within the last fifteen years in the field of adaptive control of robotic manipulators alone. An example of a robotic manipulator adaptive control scheme is discussed in detail in an article by Lee, Kelly and Karim [15].

A relative newcomer to the field of adaptive filtering is large space structures (LSS). Larger spacecraft offer several advantages over smaller spacecraft including longer on-orbit lifespans (thus, fewer launch vehicles are required), on-orbit refueling (for lower orbit spacecraft) and more or larger payloads. As these spacecraft increase in size and complexity, some problems arise. A larger, more complex spacecraft will have lower bending frequencies, more fuel slosh, larger disturbances [11] and the possibility of greater interaction between multiple payloads [10]. Adaptive filtering techniques have been proposed for use in identifying parameters such as vibrational frequencies, damping coefficients, and attitude estimation.

Accurately estimating the states of a system whose dynamics are time-invariant and known is often easily accomplished. There are numerous estimation methods for accomplishing this task, depending on the particular application. New problems are created, however, when the system has unknown or time-varying dynamics. If estimation is attempted for system dynamics that are incorrectly modelled, large errors in the estimated states are likely. Even more problems may arise if control is to be applied based upon the estimated states. It is clear that a state feedback regulator applying control to a system based upon these incorrect states may have problems determining the correct amount of control to apply, possibly causing the system to go unstable.

This paper discusses an adaptive filter technique that utilizes a Kalman filter. Adaptive filters based upon Kalman filters are known as *adaptive Kalman filters* (AKF). This AKF design is developed in the second and third sections, then applied to a satellite attitude determination problem in Section 4.

2 A RML Identification Technique

2.1 Motivation for Development

Using a Kalman filter-based adaptive filter offers several advantages over other adaptive filter techniques. Even though it may be computationally intense depending upon the application, the adaptive Kalman filter design provides the capability to easily provide the system states where and when they may be needed. A large space structure system could be designed so that the states are the structure's attitude (pitch, roll, yaw and their rates). For this example, the AKF could serve as not only the state estimator, but the provider of these estimated states to other payloads as well.

To be adaptive, however, the AKF must be able to accurately identify unknown or time-varying system parameters. The burden of this task falls upon the identifier. Since it may be desireable to use a Kalman filter-based adaptive filter, it seems reasonable that the identifier algorithm chosen should be one that works well with a Kalman filter. This chapter will cover the development of such an identifier.

Maximum likelihood techniques have been used for parameter identification for many years. Lee [18], in his book published in 1964, claims that Fisher [8] first developed identification using maximum likelihood techniques. The concept of recursive maximum likelihood identification using a Newton-Raphson type optimization technique and Kalman filter equations was originally conceived by Stepner and Mehre [25] in 1973. Their algorithm, which was designed to handle nonlinear as well as linear dynamics, was very computationally intensive. In 1986, Fermelia [7] further expanded the concepts, and Sjodin and Fermelia [24] developed working code for a first order dynamics model in 1987. Sjodin and Fermelia's algorithm was much simpler in concept, thus giving it a much better chance for real-time implementation. Kelly [14] and Fermelia extended this work to multiple second order dynamic models in 1989.

2.2 System Model

This dissertation discusses an adaptive Kalman filter design where the system model is defined in state-space form. Consider a system described by

$$\dot{\underline{x}} = F\underline{x} + G\underline{u} + \underline{w} \tag{1}$$

$$\underline{z} = H\underline{x} + \underline{v}. \tag{2}$$

The noise vectors, $\underline{w}$ and $\underline{v}$, are assumed to be normally distributed with a mean of zero.

In order to obtain a discrete-time model of the above system model, it is assumed that the set of discrete-time points $\{0, 1, \ldots, k, k+1\}$ are sufficiently close together that piecewise constant approximations may be made. The solution of $\dot{\underline{x}} = F\underline{x} + G\underline{u}$ for such an interval may be expressed as [2]

$$\underline{x}_{k+1} = \Phi_{k+1,k}\underline{x}_k + \int_k^{k+1} \Phi_{k+1,\tau} G_\tau d\tau \underline{u}_k. \tag{3}$$

2.3 Defining the Performance Index

In this section, a performance index is defined for RML identification. Since the identification technique is a recursive maximum likelihood method, the log likelihood function must first be defined. To identify the system parameters, the log likelihood function must be maximized. This function is maximized by minimizing its negative term. This negative term is chosen as the performance index.

2.3.1 The Log Likelihood Function

The identification method that is developed in this section is a recursive maximum likelihood method. Variables that are vectors rather than scalars are underlined. Variables that are estimated or identified have a hat, such as $\hat{\underline{x}}$. First, define a log likelihood function, $\mathcal{L}(\underline{\theta})$, as

$$\mathcal{L}(\underline{\theta}) = \ln[p(\underline{Z}|\underline{\theta})] \tag{4}$$

where $\underline{\theta}$ is a vector of unknown parameters and $p(\underline{Z}|\underline{\theta})$ is the conditional probability density function of the observations, $\underline{Z}$, given $\underline{\theta}$. The maximum likelihood estimate, $\underline{\theta}$, is the parameter vector that most likely caused the observations [24].

Repeated use of Bayes' Law is used to derive an expression for $\mathcal{L}(\underline{\theta})$. Let $\underline{Z}_{k+1}$ be the set of all observations at time $k+1$, or

$$\underline{Z}_{k+1} = [\underline{z}_1\, \underline{z}_2 \cdots \underline{z}_{k+1}]. \tag{5}$$

Now, $p(\underline{Z}_{k+1}|\underline{\theta})$ may be expressed as

$$p(\underline{Z}_{k+1}|\underline{\theta}) = p(\underline{z}_1\, \underline{z}_2 \cdots \underline{z}_{k+1}|\underline{\theta}). \tag{6}$$

$\mathcal{L}(\underline{\theta})$ may now be written as

$$\mathcal{L}(\underline{\theta}) = \ln \prod_{i=1}^{k+1} p(\underline{z}_i|\underline{Z}_{i-1}, \underline{\theta}). \tag{7}$$

If the noise values, $\underline{w}_k$ *and* $\underline{v}_{k+1}$, are normally distributed, $p(\underline{z}_i|\underline{Z}_{i-1}, \underline{\theta})$ is normally distributed. From [24],

$$p(\underline{z}_i|\underline{Z}_{i-1}, \underline{\theta}) = [(2\pi)^m det\, B]^{-1/2} exp(-\frac{1}{2}\underline{\nu}_i^T B^{-1} \underline{\nu}_i) \tag{8}$$

where

$$\begin{aligned} m &\stackrel{\text{def}}{=} dim\,[\underline{\nu}] \\ B &\stackrel{\text{def}}{=} E\,(\underline{\nu}\,\underline{\nu}^T) \\ \underline{\nu}_i &\stackrel{\text{def}}{=} \underline{z}_i - \hat{\underline{z}}_i. \end{aligned}$$

Substituting this into the Equation 8 gives

$$\begin{aligned} \mathcal{L}(\underline{\theta}) &= \ln(\prod_{i=1}^{k+1}[(2\pi)^m\, det\, B]^{-1/2} exp(-\frac{1}{2}\,\underline{\nu}_i^T\, B^{-1}\, \underline{\nu}_i)) \\ &= -\frac{1}{2}\, m(k+1)\, \ln(2\pi) \\ &\quad -\sum_{i=1}^{k+1} \frac{1}{2}\,(\ln(det\, B) + \underline{\nu}_i^T\, B^{-1}\, \underline{\nu}_i). \end{aligned} \tag{9}$$

2.3.2 The Performance Index

Now that the log likelihood function has been defined, an expression for the performance index must be found. In Equation 9, $\mathcal{L}(\underline{\theta})$ is maximized if the summation expression is minimized. Consider the following performance index,

$$J_{k+1}(\underline{\theta}) = \sum_{i=1}^{k+1} \frac{1}{2}\,(\ln(det\, B) + \underline{\nu}_i^T\, B^{-1}\, \underline{\nu}_i) \tag{10}$$

to be minimized.

If $J_{k+1}(\underline{\theta})$ is expanded using a Taylor Series expansion, one obtains

$$\begin{aligned} J_{k+1}(\underline{\theta}) &= J_{k+1}(\underline{\theta}^*) + \frac{\partial J_{k+1}(\underline{\theta})}{\partial \underline{\theta}}|_{\underline{\theta}=\underline{\theta}^*}\delta\underline{\theta} \\ &\quad + \frac{1}{2}\frac{\partial^2 J_{k+1}(\underline{\theta})}{\partial \underline{\theta}^2}|_{\underline{\theta}=\underline{\theta}^*}\delta\underline{\theta}^2 + \cdots \end{aligned} \tag{11}$$

where $\underline{\theta}^*$ is the current value of the parameter, $\underline{\theta}$.

If third and higher order partial terms are assumed to be negligible, the above expression can be solved for $\delta\underline{\theta}$, the parameter error. At steady state, it is desired that $J_{k+1}(\underline{\theta}) - J_{k+1}(\underline{\theta}^*) = 0$. What remains of the Taylor Series expansion is

$$0 = \frac{\partial J_{k+1}(\underline{\theta})}{\partial \underline{\theta}}|_{\underline{\theta}=\underline{\theta}^*}\delta\underline{\theta} + \frac{1}{2}\frac{\partial^2 J_{k+1}(\underline{\theta})}{\partial \underline{\theta}^2}|_{\underline{\theta}=\underline{\theta}^*}\delta\underline{\theta}^2. \tag{12}$$

Solving for the parameter update, the following expression is obtained:

$$\delta\underline{\theta} \;=\; -\,2\left[\frac{\partial^2 J_{k+1}(\underline{\theta})}{\partial\underline{\theta}^2}|_{\underline{\theta}=\underline{\theta}^*}\right]^{-1}\;\frac{\partial J_{k+1}(\underline{\theta})}{\partial\underline{\theta}}|_{\underline{\theta}=\underline{\theta}^*}. \tag{13}$$

The parameter update, $\underline{\theta}_{new}$ is found by adding the error in the parameter, $\delta\underline{\theta}$, to the current value of the parameter, $\underline{\theta}_{old}$.

$$\underline{\theta}_{new} \;=\; \underline{\theta}_{old} \;+\; \delta\underline{\theta} \tag{14}$$

Equation 13 is very similar in form to the recursive Newton-Raphson algorithm.

The first and second partials are found by differentiating $J_{k+1}(\underline{\theta})$ from Equation 10. The error covariance matrix, B, is assumed to be constant once steady state has been reached. Thereby, the first term of Equation 10 may be ignored when the partial is taken, giving [1] [24] [25]

$$\frac{\partial J_{k+1}(\underline{\theta})}{\partial\underline{\theta}} \;=\; \sum_{i=1}^{k+1}\frac{\partial\underline{\nu}_i^T}{\partial\underline{\theta}}B^{-1}\underline{\nu}_i \tag{15}$$

$$\frac{\partial^2 J_{k+1}(\underline{\theta})}{\partial\underline{\theta}^2} \;=\; \sum_{i=1}^{k+1}\frac{\partial\underline{\nu}_i^T}{\partial\underline{\theta}}B^{-1}\frac{\partial\underline{\nu}_i}{\partial\underline{\theta}^T}. \tag{16}$$

The error covariance matrix B is defined as [20] [25]

$$B \;=\; \frac{1}{k+1}\sum_{i=1}^{k+1}\underline{\nu}_i\underline{\nu}_i^T. \tag{17}$$

The error covariance matrix B contains a vector, $\underline{\nu}$, multiplied by its transpose, $\underline{\nu}$. Because of this, B will tend to be singular for the first and second iterations. This singularity will cause difficulties in solving for B^{-1}. Depending upon the dynamics of the system, the identifier generally works well if B^{-1} is reinitialized on or about the third iteration. Hence, one solves for B on the first two iterations, but not for B^{-1}, $\frac{\partial J}{\partial\underline{\theta}}$ or $\frac{\partial^2 J}{\partial\underline{\theta}^2}$. Then on the third iteration, begin solving for all identifier values.

2.4 RML Identification Algorithm

The next two sections of this chapter discuss the matrix (multiple identified parameters) form of RML identification. The first section covers the details of the derivation of the matrix form of RML identification. This is followed by a section describing the procedure for matrix RML identification.

2.5 RML Identification Algorithm

The recursive identification parameter update algorithm is described by Equations 13 through 17. Equation 13 is comprised of Equations 15 and 16. The $\frac{\partial \underline{\nu}_{k+1}^T}{\partial \underline{\theta}}$ term of these last two equations is now described. Since the following derivations rely heavily upon the Kalman filter equations, Table 1 [9] is presented as an quick reference of the set of Kalman filter equations.

Table 1: Summary of Kalman Filter Equations

	Kalman Filter Equations
Process Model	$\underline{x}_{k+1} = \Phi_{k+1,k}\underline{x}_k + \underline{w}_k \qquad \underline{w}_k \sim N(\underline{0}, Q_k)$
Measurement Model	$\underline{z}_{k+1} = H_{k+1}\underline{x}_{k+1} + \underline{v}_{k+1} \qquad \underline{v}_{k+1} \sim N(\underline{0}, R_k)$
State Estimate Extrapolation	$\hat{\underline{x}}_{k+1/k} = \Phi_{k+1,k}\hat{\underline{x}}_{k/k}$
Error Covariance Extrapolation	$\overline{P}_{k+1/k} = \Phi_{k+1,k}P_{k/k}\Phi_{k+1,k}^T + Q_k$
Error Covariance Update	$P_{k+1/k+1} = [I - K_{k+1}H_{k+1}]\overline{P}_{k+1/k}$
Kalman Gain	$K_{k+1} = \overline{P}_{k+1/k}H_{k+1}^T\left[H_{k+1}\overline{P}_{k+1/k}H_{k+1}^T + R_{k+1}\right]^{-1}$
State Estimate Update	$\hat{\underline{x}}_{k+1/k+1} = \hat{\underline{x}}_{k+1/k} + K_{k+1}\left[\underline{z}_{k+1} - H_{k+1}\hat{\underline{x}}_{k+1/k}\right]$

First, define a variation in $\underline{\nu}_{k+1}$ as the error between the desired innovations, $\underline{\nu}_{k+1}(D)$, and the incorrect innovations, $\underline{\nu}_{k+1}(I)$.

$$\delta\underline{\nu}_{k+1} = \underline{\nu}_{k+1}(D) - \underline{\nu}_{k+1}(I). \tag{18}$$

The desired innovations can be viewed as the innovations expected for a Kalman filter whose dynamics match those of the system being modeled. The incorrect innovations can be viewed as the innovations expected from a Kalman filter whose model dynamics do not match those of the system being modeled.

In order to obtain an expression for $\delta\underline{\nu}_{k+1}$, begin with the Kalman filter expression for the innovations.

$$\underline{\nu}_{k+1} = \underline{z}_{k+1} - \hat{\underline{z}}_{k+1/k}. \tag{19}$$

Then, observing the variation of both sides,

$$\delta\underline{\nu}_{k+1} = \delta(\underline{z}_{k+1} - \hat{\underline{z}}_{k+1/k}). \tag{20}$$

Since there is no error in $\underline{z}_{k+1}$ due to errors in the system dynamics or modeling, $\delta\underline{z}_{k+1} = 0$. Thus,

$$\delta\underline{\nu}_{k+1} = -\delta(\hat{\underline{z}}_{k+1/k}) \tag{21}$$

Assume that the unknown or changing parameters exist only in $\hat{\Phi}_{k+1,k}$, the Kalman filter state transition matrix. Using the Kalman filter equation for $\hat{\underline{z}}_{k+1/k}$ from Table 1, expand to get

$$\delta\underline{\nu}_{k+1} = -\delta\left[H_{k+1}\hat{\Phi}_{k+1,k}\hat{\underline{x}}_{k/k}\right]. \tag{22}$$

The parameters to be identified occur only in $\hat{\Phi}_{k+1,k}$, the state transition matrix. Therefore,

$$\begin{aligned}\delta\underline{\nu}_{k+1} &= -\left[H_{k+1}\delta\hat{\Phi}_{k+1,k}\hat{\underline{x}}_{k/k} + H_{k+1}\hat{\Phi}_{k+1,k}\delta\hat{\underline{x}}_{k/k}\right] \\ &= -\left[H_{k+1}\frac{\partial\hat{\Phi}_{k+1,k}}{\partial\underline{\theta}^T}\hat{\underline{x}}_{k/k} + H_{k+1}\hat{\Phi}_{k+1,k}\frac{\partial\hat{\underline{x}}_{k/k}}{\partial\underline{\theta}^T}\right]\delta\underline{\theta}.\end{aligned} \tag{23}$$

The term $\partial\hat{\Phi}$ may be expressed as

$$\partial\hat{\Phi} = \begin{bmatrix} \partial\phi_{11} & \partial\phi_{12} & \cdots & \partial\phi_{1n} \\ \partial\phi_{21} & \partial\phi_{22} & \cdots & \partial\phi_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \partial\phi_{n1} & \partial\phi_{n2} & \cdots & \partial\phi_{nn} \end{bmatrix} \tag{24}$$

where

$$\begin{aligned}\partial\phi_{ij} &= \frac{\partial\phi_{ij}}{\partial\theta_1}\delta\theta_1 + \frac{\partial\phi_{ij}}{\partial\theta_2}\delta\theta_2 + \cdots + \frac{\partial\phi_{ij}}{\partial\theta_n}\delta\theta_n \\ &= \frac{\partial\phi_{ij}}{\partial\underline{\theta}^T}\delta\underline{\theta}.\end{aligned} \tag{25}$$

Let the terms contained by the brackets in Equation 23 be represented by a matrix A_{k+1}.

$$A_{k+1} = H_{k+1}\frac{\partial\hat{\Phi}_{k+1,k}}{\partial\underline{\theta}^T}\hat{\underline{x}}_{k/k} + H_{k+1}\Phi_{k+1,k}\frac{\partial\hat{\underline{x}}_{k/k}}{\partial\underline{\theta}^T}. \tag{26}$$

Then,

$$\frac{\partial\nu_{k+1}^T}{\partial\underline{\theta}} = -A_{k+1}. \tag{27}$$

The first term of A_{k+1} may be realized by the following algorithm [5]:

$$\begin{aligned}H_{k+1}\delta\hat{\Phi}_{k+1,k}\hat{\underline{x}}_{k/k} &= H_{k+1}\begin{bmatrix} \hat{\underline{x}}_{k/k}^T & \frac{\partial\phi_1}{\partial\underline{\theta}^T} \\ \hat{\underline{x}}_{k/k}^T & \frac{\partial\phi_2}{\partial\underline{\theta}^T} \\ & \vdots \\ \hat{\underline{x}}_{k/k}^T & \frac{\partial\phi_n}{\partial\underline{\theta}^T} \end{bmatrix}\delta\underline{\theta} \\ &= H_{k+1}\,M_{k+1}\,\delta\underline{\theta}\end{aligned} \tag{28}$$

or

$$H_{k+1}\frac{\partial\hat{\Phi}_{k+1,k}}{\partial\underline{\theta}^T}\hat{\underline{x}}_{k/k} = H_{k+1}M_{k+1} \tag{29}$$

where ϕ_{ij} are elements of $\hat{\Phi}_{k+1,k}$ and

$$\begin{aligned}
\phi_1 &= [\,\phi_{11}\ \ \phi_{12}\ \ \cdots\ \ \phi_{1n}\,]^T \\
\phi_2 &= [\,\phi_{21}\ \ \phi_{22}\ \ \cdots\ \ \phi_{2n}\,]^T \\
&\vdots \\
\phi_n &= [\,\phi_{n1}\ \ \phi_{n2}\ \ \cdots\ \ \phi_{nn}\,]^T .
\end{aligned}$$

An expression must now be found for the second term of A_{k+1}. If an expression can be found for $\frac{\partial\hat{\underline{x}}_{k+1/k+1}}{\partial\underline{\theta}^T}$, then $\frac{\partial\hat{\underline{x}}_{k/k}}{\partial\underline{\theta}^T}$ is simply the value of $\frac{\partial\hat{\underline{x}}_{k+1/k+1}}{\partial\underline{\theta}^T}$ from the last iteration. Let a matrix B_{k+1} be defined as

$$B_{k+1} = \frac{\partial\hat{\underline{x}}_{k+1/k+1}}{\partial\underline{\theta}^T}. \tag{30}$$

Then, B_{k+1} from the last iteration is defined as

$$B_k = \frac{\partial\hat{\underline{x}}_{k/k}}{\partial\underline{\theta}^T}. \tag{31}$$

Again, using the Kalman filter equation for $\hat{\underline{x}}_{k+1/k+1}$ from Table 1, the following expression is derived

$$\begin{aligned}
B_{k+1}\delta\underline{\theta} &= \delta\hat{\underline{x}}_{k+1/k+1} \\
&= \delta\left[\hat{\Phi}_{k+1,k}\hat{\underline{x}}_{k/k} + G_{k+1}\underline{\nu}_{k+1}\right] \\
&= \left[\frac{\partial\hat{\Phi}_{k+1,k}}{\partial\underline{\theta}^T}\hat{\underline{x}}_{k/k} + \hat{\Phi}_{k+1,k}\frac{\partial\hat{\underline{x}}_{k/k}}{\partial\underline{\theta}^T}\right. \\
&\qquad \left. + \frac{\partial G_{k+1}}{\partial\underline{\theta}^T}\underline{\nu}_{k+1} + G_{k+1}\frac{\partial\underline{\nu}_{k+1}}{\partial\underline{\theta}^T}\right]\delta\underline{\theta}.
\end{aligned} \tag{32}$$

Thus, B_{k+1} may be written in terms of M_{k+1}, B_k and A_{k+1} as

$$\begin{aligned}
B_{k+1} &= M_{k+1} + \hat{\Phi}_{k+1,k}B_k \\
&\qquad + \frac{\partial G_{k+1}}{\partial\underline{\theta}^T}\underline{\nu}_{k+1} - G_{k+1}A_{k+1}.
\end{aligned} \tag{33}$$

Let a matrix C_{k+1} be defined such that

$$C_{k+1} = \frac{\partial G_{k+1}}{\partial\underline{\theta}^T}\underline{\nu}_{k+1} \tag{34}$$

or

$$C_{k+1}\delta\underline{\theta} = \delta G_{k+1}\underline{\nu}_{k+1}. \tag{35}$$

An expression for $\frac{\partial G_{k+1}}{\partial \underline{\theta}^T}$ is found next. Again, expand using the Kalman filter equation for G_{k+1} from Table 1.

$$\delta G_{k+1} = \left[\frac{\partial P_{k+1/k+1}}{\partial \underline{\theta}^T} H_{k+1}^T R_{k+1}^{-1}\right] \delta\underline{\theta}. \tag{36}$$

Expand $\frac{\partial P_{k+1/k+1}}{\partial \underline{\theta}^T}$ the same way.

$$\delta P_{k+1/k+1} = [I - G_{k+1}H_{k+1}]\,\delta P_{k+1/k} - \delta G_{k+1}\, H_{k+1}P_{k+1/k}. \tag{37}$$

Letting

$$\Gamma_{k+1} = I - G_{k+1}H_{k+1} \tag{38}$$

Equation 37 becomes

$$\delta P_{k+1/k+1} = \left[\Gamma_{k+1}\frac{\partial P_{k+1/k}}{\partial \underline{\theta}^T} - \frac{\partial G_{k+1}}{\partial \underline{\theta}^T} H_{k+1}P_{k+1/k}\right] \delta\underline{\theta}. \tag{39}$$

Thus

$$\frac{\partial P_{k+1/k+1}}{\partial \underline{\theta}^T} = \Gamma_{k+1}\frac{\partial P_{k+1/k}}{\partial \underline{\theta}^T} - \frac{\partial G_{k+1}}{\partial \underline{\theta}^T} H_{k+1}P_{k+1/k}. \tag{40}$$

Now, $\frac{\partial P_{k+1/k+1}}{\partial \underline{\theta}^T}$ may be described as

$$\begin{aligned}\frac{\partial P_{k+1/k+1}}{\partial \underline{\theta}^T} &= \frac{\partial P_{k+1/k}}{\partial \underline{\theta}^T} - \frac{\partial G_{k+1}}{\partial \underline{\theta}^T} H_{k+1}P_{k+1/k} \\ &\quad - G_{k+1}H_{k+1}\frac{\partial P_{k+1/k}}{\partial \underline{\theta}^T}\end{aligned} \tag{41}$$

and substituting $\delta P_{k+1/k+1}$ into Equation 36 yields

$$\begin{aligned}\delta G_{k+1} &= \left[\Gamma_{k+1}\frac{\partial P_{k+1/k}}{\partial \underline{\theta}^T} H_{k+1}^T R_{k+1}^{-1}\right] \\ &\quad \left[I + H_{k+1}P_{k+1/k}H_{k+1}^T R_{k+1}^{-1}\right]^{-1} \delta\underline{\theta}.\end{aligned} \tag{42}$$

Introduce the matrix V_{k+1} where

$$V_{k+1} = H_{k+1}^T R_{k+1}^{-1}\left[I + H_{k+1}P_{k+1/k}H_{k+1}^T R_{k+1}^{-1}\right]^{-1} \underline{\nu}_{k+1}. \tag{43}$$

Then, post multiplying Equation 42 by $\underline{\nu}_{k+1}$, an expression for C_{k+1} is obtained as

$$\begin{aligned}C_{k+1} &= \frac{\partial G_{k+1}}{\partial \underline{\theta}^T}\underline{\nu}_{k+1} \\ &= \Gamma_{k+1}\frac{\partial P_{k+1/k}}{\partial \underline{\theta}^T} V_{k+1}\end{aligned} \tag{44}$$

Letting

$$D_{k+1} = \frac{\partial P_{k+1/k}}{\partial \underline{\theta}^T} V_{k+1} \tag{45}$$

allows C_{k+1} to be expressed in terms of a matrix D_{k+1}.

$$C_{k+1} = \Gamma_{k+1} D_{k+1} \tag{46}$$

Next, an expression for D_{k+1} is found as

$$\begin{aligned} D_{k+1}\delta\underline{\theta} &= \delta P_{k+1/k} V_{k+1} \\ &= \delta\hat{\Phi}_{k+1,k} W_{k+1} + \hat{\Phi}_{k+1,k}\, \delta P_{k/k}\, U_{k+1} \\ &\quad + \hat{\Phi}_{k+1,k} P_{k/k}\, \delta\hat{\Phi}^T_{k+1,k} V_{k+1} \end{aligned} \tag{47}$$

where

$$W_{k+1} = P_{k/k}\hat{\Phi}^T_{k+1,k} V_{k+1} \tag{48}$$

$$U_{k+1} = \hat{\Phi}^T_{k+1,k} V_{k+1}. \tag{49}$$

Using the same technique as in Equation 28, the first term of the D_{k+1} expression may be written as

$$\begin{aligned} \delta\hat{\Phi}_{k+1,k} W_{k+1} &= \begin{bmatrix} W^T_{k+1} \frac{\partial \phi_1}{\partial \underline{\theta}^T} \\ W^T_{k+1} \frac{\partial \phi_2}{\partial \underline{\theta}^T} \\ \vdots \\ W^T_{k+1} \frac{\partial \phi_n}{\partial \underline{\theta}^T} \end{bmatrix} \delta\underline{\theta} \\ &= S_{k+1}\delta\underline{\theta}. \end{aligned} \tag{50}$$

The third term can be written as

$$\begin{aligned} \hat{\Phi}_{k+1,k} P_{k/k}\, \delta\hat{\Phi}^T_{k+1,k} V_{k+1} &= \Phi_{k+1,k} P_{k/k} \begin{bmatrix} V^T_{k+1} \frac{\partial \phi^*_1}{\partial \underline{\theta}^T} \\ V^T_{k+1} \frac{\partial \phi^*_2}{\partial \underline{\theta}^T} \\ \vdots \\ V^T_{k+1} \frac{\partial \phi^*_n}{\partial \underline{\theta}^T} \end{bmatrix} \delta\underline{\theta} \\ &= \hat{\Phi}_{k+1,k} P_{k/k} N_{k+1} \delta\underline{\theta} \end{aligned} \tag{51}$$

where

$$\begin{aligned} \phi^*_1 &= [\, \phi_{11} \;\; \phi_{21} \;\; \cdots \;\; \phi_{n1} \,]^T \\ \phi^*_2 &= [\, \phi_{12} \;\; \phi_{22} \;\; \cdots \;\; \phi_{n2} \,]^T \\ &\vdots \\ \phi^*_n &= [\, \phi_{1n} \;\; \phi_{2n} \;\; \cdots \;\; \phi_{nn} \,]^T. \end{aligned}$$

Let the second term of the D_{k+1} expression be written in terms of a matrix E_{k+1},

$$\hat{\Phi}_{k+1,k}\delta P_{k/k}U_{k+1} = \hat{\Phi}_{k+1,k}E_{k+1}\delta\underline{\theta}. \tag{52}$$

Combining all three terms for D_{k+1} gives

$$D_{k+1} = \hat{\Phi}_{k+1,k}E_{k+1} + S_{k+1} + \hat{\Phi}_{k+1,k}P_{k/k}N_{k+1}. \tag{53}$$

An expression for E_{k+1} must now be found. Equation 52 defines E_{k+1} as

$$E_{k+1}\delta\underline{\theta} = \delta P_{k/k}U_{k+1}. \tag{54}$$

Using the Kalman filter equation for $P_{k/k}$ yields

$$\delta P_{k/k} = \delta P_{k/k-1} - \delta P_{k/k}H_k^T R_k^{-1} H_k P_{k/k-1} - P_{k/k}H_k^T R_k^{-1} H_k \delta P_{k/k-1}. \tag{55}$$

Solving for $\delta P_{k/k}$ and postmultiplying by U_{k+1} yields

$$\delta P_{k/k}U_{k+1} = \Gamma_k \delta P_{k/k-1}Y_{k+1} \tag{56}$$

or

$$E_{k+1}\delta\underline{\theta} = \Gamma_k \delta P_{k/k-1}Y_{k+1} \tag{57}$$

where

$$Y_{k+1} = T_k U_{k+1} \tag{58}$$

$$T_k = I + H_k^T R_k^{-1} P_{k/k-1}. \tag{59}$$

Let a matrix F_{k+1} be defined as

$$F_{k+1}\delta\underline{\theta} = \delta P_{k/k-1}Y_{k+1} \tag{60}$$

such that E_{k+1} is expressed in terms of F_{k+1}.

$$E_{k+1} = \Gamma_k F_{k+1} \tag{61}$$

The matrix F_{k+1} is expressed in terms of D_k, the previous value of the D_{k+1} matrix. Since $\delta P_{k/k-1}$ is symmetric,

$$\begin{aligned} F_{k+1}\delta\underline{\theta} &= \delta P_{k/k-1}Y_{k+1} \\ &= V_k^* V_k^T \delta P_{k/k-1}Y_{k+1} \\ &= V_k^* \left(D_k\delta\underline{\theta}\right)^T Y_{k+1} \end{aligned} \tag{62}$$

where V_k^* is defined as the pseudo-inverse of V_k^T [26], or

$$V_k^* = V_k \left[V_k^T V_k\right]^{-1}. \tag{63}$$

This allows F_{k+1} to be expressed as

$$F_{k+1} = \left[\begin{array}{cccc} V_k^* Y_{k+1}^T d_1 & V_k^* Y_{k+1}^T d_2 & \cdots & V_k^* Y_{k+1}^T d_j \end{array}\right] \tag{64}$$

where d_j is defined as the j^{th} column of D_k.

3 Matrix RML Identification Procedure

The identification algorithm has now been derived for the matrix (multiple parameter) case where only the state transition matrix, $\hat{\Phi}_{k+1,k}$, is a function of the parameters. The procedure for using the matrix form of RML identification is now explained. There are six major steps, with steps five and six having several substeps.

1. Set B_k equal to B_{k+1} of the last identifier iteration. If this is the first iteration, set it equal to some small reasonable number. The user may have to try several different starting values in addition to letting the identifier iterate several times in order to determine a reasonable initial value.
2. Calculate $\frac{\partial \underline{\nu}_{k+1}^T}{\partial \underline{\theta}}$ using Equation 27.
3. Calculate the error covariance, B , using Equation 17.
4. Calculate the first and second partials of J_{k+1} using Equations 15 and 16.
5. If ready to update,
   - Calculate the parameter error vector, $\delta\underline{\theta}$, using Equation 13.
   - Update the parameter vector, $\underline{\theta}$, using Equation 14.
6. If not ready to update, solve for B_{k+1} to be used as B_k in the next iteration. To do this,
   - Set D_k, T_k, V_k and Γ_k equal to D_{k+1}, T_{k+1}, V_{k+1} and Γ_{k+1}, respectively, from the previous iteration. From the last iteration of the Kalman filter, get the values for $P_{k/k}$, $\underline{\hat{x}}_{k/k}$, $P_{k+1/k}$, G_{k+1}, H_{k+1} and R_{k+1}.
   - Calculate:
     (a) V_k^* from Equation 63
     (b) V_{k+1} from Equation 43
     (c) U_{k+1} from Equation 49
     (d) Y_{k+1} from Equation 58
     (e) F_{k+1} from Equation 64
     (f) E_{k+1} from Equation 61
     (g) N_{k+1} from Equation 51
     (h) W_{k+1} from Equation 48
     (i) S_{k+1} from Equation 50
     (j) D_{k+1} from Equation 53
     (k) Γ_{k+1} from Equation 38
     (l) C_{k+1} from Equation 46
     (m) B_{k+1} from Equation 33

4 Attitude Determination Example

This section discusses an application of AKF-RML to a spacecraft attitude scenario. First, the dynamic equations are developed. Next, the state space model for the dynamic equations is explained. Finally, results are shown using simulated data.

4.1 Dynamic Equations

Assume the attitude of a spacecraft may be described using Euler's equations [12],

$$I_1 \frac{\delta \omega_1}{\delta t} = N_1 + (I_2 - I_3)\omega_2\omega_3 \tag{65}$$

$$I_2 \frac{\delta \omega_2}{\delta t} = N_2 + (I_3 - I_1)\omega_3\omega_1 \tag{66}$$

$$I_3 \frac{\delta \omega_3}{\delta t} = N_3 + (I_1 - I_2)\omega_1\omega_2 \tag{67}$$

where I_i represents the inertia about the ith axis, ω_i is the angular velocity about the ith axis and N_i is the applied torque about the ith axis.

If torque-free motion ($N_i = 0$), symmetry of two of the inertias ($I_1 = I_2 = I_T$) and constant velocity in the third axis ($\omega_3 = n$) is assumed, then the above equations become

$$I_T \frac{\delta \omega_1}{\delta t} = -(I_3 - I_T)\omega_2\omega_3 \tag{68}$$

$$I_T \frac{\delta \omega_2}{\delta t} = (I_3 - I_T)\omega_3\omega_1 \tag{69}$$

$$I_T \frac{\delta \omega_3}{\delta t} = 0. \tag{70}$$

Differentiating Equation 68 yields

$$I_T \ddot{\omega}_1 = -(I_3 - I_T)\dot{\omega}_2\omega_3. \tag{71}$$

Multiplying by I_T and substituting in Equation 69 gives

$$I_T^2 \ddot{\omega}_1 = -(I_3 - I_T)\omega_3 I_T \dot{\omega}_2 \tag{72}$$

$$= -(I_3 - I_T)\omega_3 (I_3 - I_T)\omega_3\omega_1 \tag{73}$$

$$= -(I_3 - I_T)^2 \omega_3^2 \omega_1 \tag{74}$$

or

$$\ddot{\omega}_1 = -\left(\frac{I_3 - I_T}{I_T}\right)^2 \omega_1. \tag{75}$$

Integrating Equation 70 then yields an expression for ω_2 as

$$\omega_2 = \left(\frac{I_3 - I_T}{I_T}\right)\omega_3\omega_1 t \tag{76}$$

where t is the step size.

4.2 State Space Model

Equation 75 is in the form of a simple undamped harmonic oscillator. Therefore, it may be expressed in continuous state space form as

$$\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\omega_n^2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{77}$$

or in discrete state space form as

$$\begin{bmatrix} \dot{x_1} \\ \dot{x_2} \end{bmatrix}_{k+1} = \begin{bmatrix} \cos\omega_n t & \frac{1}{\omega_n}\sin\omega_n t \\ -\omega_n \sin\omega_n t & \cos\omega_n t \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_k \tag{78}$$

where

$$\omega_n = \left(\frac{I_3 - I_T}{I_T}\right)\omega_3. \tag{79}$$

4.3 Attitude Determination Results

The example described in Sections 4.1 and 4.2 were implemented on a XT-compatable PC. The true natural frequency, ω_n, was set to -6.0 radians. The natural frequency in the Kalman filter dynamics was assumed to be unknown, and the identifier set up to solve for the true value.

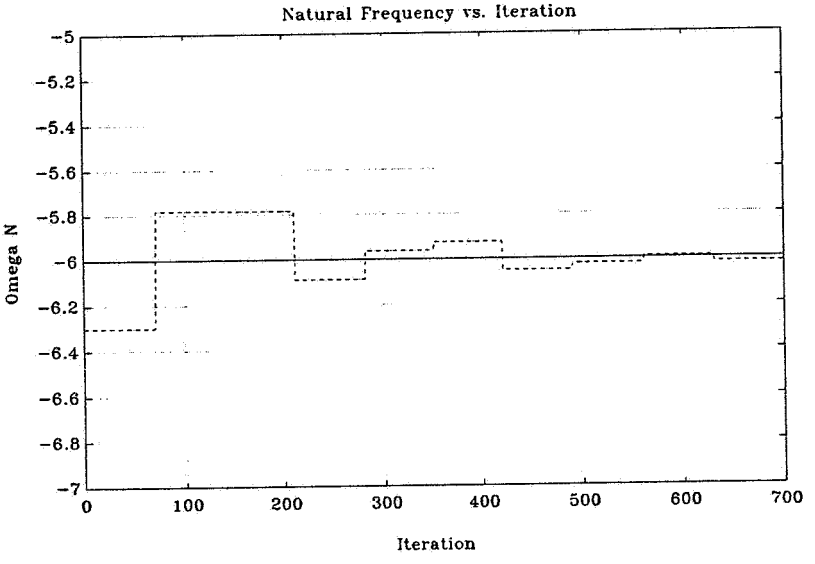
In the first test case, AKF-RML was set up to operate in a pseudo-batch mode. In this mode, the identifier is actually solving for the unknown parameter at each iteration, but the parameter is updated every k iterations. Figure 1(a) shows the results of this test case with k set to 70 iterations and the initial guess on ω_n equal to -6.3 radians.

While this pseudo-batch mode is useful for analysis and to gain insight into the operation of the AKF-RML algorithm, the goal is still to operate recursively. Figure 1(b) shows the results of recursive AKF-RML with the initial guess of ω_n at -6.6 radians.

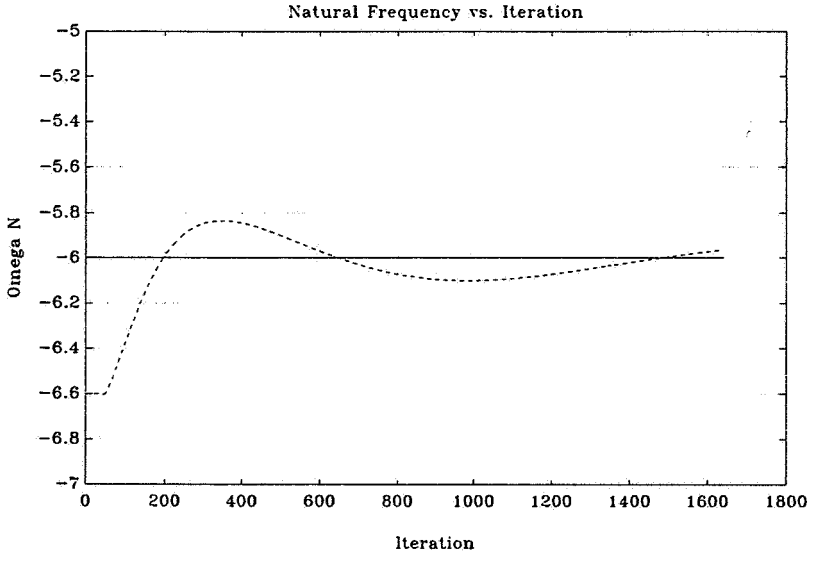
5 Summary

Adaptive filtering is rapidly gaining popularity as a method of estimating systems with unknown or changing dynamics. This paper offers a recursive identification algorithm that is designed to be used in conjunction with a Kalman filter.

In [14], the recursive maximum likelihood identification algorithm is developed and tested. Extensive testing is accomplished using simulated data, beginning with the simple first order, decaying exponential case. This is extended to second order spring-mass dynamics, with excellent results obtained for both of these cases. Next, damping is added to create the well-known spring-mass-damper dynamic case. Very



(a)



(b)

Figure 1: AKF-RML Results in (a) Batch Mode and (b) Recursive Mode

good results are obtained when identifying the natural frequency or damping coefficient individually. Problems arise, however, when both parameters are identified concurrently. This problem presents itself whenever the parameters to be identified are not of the same magnitude, and is solved by the addition of a scaling matrix to the parameter update equation.

The identification algorithm developed in this paper may be computationally intense in some applications. As with all Kalman filter algorithms, the computational needs rise quickly as the number of states and observations increase. The identification algorithm computational needs will rise not only with an increase in the states and observations, but with an increase in the number of identified parameters as well. Another possible disadvantage is the trial and error method that is currently used to find the parameter scale factors. The use of scale factors is actually not a disadvantage against this identification algorithm alone, as similar scaling or weighting matrices are used in many current identification algorithms when identifying parameters of different magnitudes. In addition, maximum likelihood techniques tend to converge poorly when the initial conditions are far away from the true conditions. In most practical applications the users should be able to arrive at reasonable initial conditions.

The advantages of using this adaptive filter design are many. Incorporation of a Kalman filter estimator in AKF-RML allows the designer to choose system states that relate to real-world entities such as position, velocity and acceleration. In the case of aircraft or satellite applications, for example, those states can then be passed on to various payloads for pointing requirements. RML identification is designed specifically to be used in conjunction with the Kalman filter. Therefore, this adaptive Kalman filter design may be implemented for systems whose dynamics are unknown or varying.

In addition, the RML identification algorithm described in this paper is able to handle relatively high levels of noise. The common problem of insufficient excitation prevents parameters from being identified at very low noise levels, but this problem disappears as noise levels are brought up enough to provide sufficient excitation.

Results are shown for a simple spacecraft attitude determination example. Plots are presented for two cases. The first case shows the results of AKF-RML in batch mode where the parameter is updated after a series of iterations. Next, results are provided of AKF-RML in recursive mode. More research needs to done, especially in the area of choosing initial conditions, in order to get AKF-RML to operate efficiently in recursive mode.

6 References

[1] M. Athans and F. C. Schweppe. Gradient Matrices and Matrix Calculations. MIT Tech. Note 1965-53, November 1965.

[2] W. L. Brogan. Modern Control Theory, Second Edition. Prentice-Hall, Inc., Englewood Cliffs, N. J., 1985.

[3] S. Chen, D. S. Wadhwani and P. D. Roberts, 'Heirarchical Integrated System Optimisation (Optimization) and Parameter Estimation using a Microcomputer Based System,' IEEE Proceedings of the 25th Conference on Decision and Control, Athens, Greece, December 1986.

[4] M. Ciletti and J. Tyler, 'Proceedings of the Symposium on Control Theory and Navy Applications,' U.S. Naval Postgraduate School, Monterey, California, July 15-17, 1975.

[5] A. Fermelia. Informal Notes. This realization was derived by A. Fermelia in 1985 and used by B. Sjodin in his Masters thesis (see Sjodin reference in this bibliography).

[6] A. Fermelia and D. A. Kelly, 'Least-Squares Formulation to a Newton-Raphson Parameter Update Algorithm,' manuscript written in December 1989.

[7] A. Fermelia and R. H. Larson, 'Closed Loop Methodology (CLM) Applied to Signal Processing; A Revisitation,' 1986 American Control Conference, 18-20 June, 1986, Volume 2 of 3, pp. 847-851.

[8] R. A. Fisher, 'On an Absolute Criterion for Fitting Frequency Curves,' Mess. Math., vol. 41, p. 155, 1912.

[9] A. Gelb, editor. Applied Optimal Estimation. The MIT Press, Cambridge, Massachusetts, 1986.

[10] G. Goodwin and K. Sin. Adaptive Filtering Prediction and Control. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1984.

[11] B. Govin and B. Claudinon, 'Adaptive Control of Flexible Space Structures,' Proceedings of the Ninth IFAC/ESA Symposium, The Netherlands, pp. 361-371, 5-9 July 1982.

[12] M. H. Kaplan. Modern Spacecraft Dynamics and Control. John Wiley and Sons, New York, New York, 1976.

[13] D. A. Kelly, G. K. F. Lee and A. Fermelia,'A Self-Tuning Regulator Approach to the Control of Large Space Structures,' SIAM Conference on Control in the 90's: Achievements, Opportunities, and Challenges, San Francisco, California, 17-19 May 1989.

[14] D. A. Kelly, 'An Adaptive Kalman Filter Design Using Recursive Maximum Likelihood Identification,' Dissertation, Department of Electrical Engineering, Colorado State University, Colorado, Spring 1990.

[15] G. K. F. Lee and D. A. Kelly, 'A Self-Tuning Regulator Design Using Recursive Maximum Likelihood Identification,' a memo sent to Dr. R. Montgomery, Spacecraft Control Branch, NASA Langley Research Center, Hampton, Virginia, May 1989.

[16] G. K. F. Lee, D. A. Kelly and A. Karim, 'A Discrete-Time Adaptive Controller for Nonlinear Systems Using a Recursive Maximum Likelihood Structure,' ISMM International Symposium on Computer Applications in Design, Simulation and Analysis, New Orleans, Louisiana, 5-7 March 1990.

[17] G. K. F. Lee and D. A. Kelly, 'A Self-Tuning Regulator Algorithm for Discrete-Time Decentralized Systems,' ISMM International Conference on Microcomputer Applications, Los Angeles, California, 14-16 December 1989.

[18] R. Lee. Optimal Estimation, Identification, and Control. Research Monograph No. 28, The M.I.T. Press, Cambridge, Massachusetts, 1964.

[19] E. J. Lefferts, F. L. Markley and M. D. Shuster, 'Kalman Filtering for Spacecraft Attitude Estimation,' AIAA J. Guidance, vol. 5, no. 5, September/October 1982.

[20] R. K. Mehra, 'On-Line Identification of Linear Dynamic Systems With Applications to Kalman Filtering,' IEEE Trans. Auto. Control, April 1970.

[21] C. E. Mueller and G. Stein,'Calibration of High Precision Navigators with Maximum Likelihood Estimation,' Proceedings of the Symposium on Control Theory and Navy Applications, U.S. Naval Postgraduate School, Monterey, California, 15-17 July 1975.

[22] A. Sage and J. Melsa. System Identification. Academic Press, New York, New York, 1971, pp. 66-81.

[23] M. D. Shuster, D. M. Chitre and D. P. Niebur, 'In-Flight Estimation of Spacecraft Attitude Sensor Accuracies and Alignments,' AIAA J. Guidance, vol. 5, no. 4, July/August 1982.

[24] R. Sjodin, 'CLM: A Closed Loop Methodology for the Modeling of Dynamic Systems,' Masters Thesis, Department of Electrical Engineering, University of Colorado, Boulder, Colorado, 1987.

[25] D. E. Stepner and R. K. Mehra, 'Maximum Likelihood Identification and Optimal Input Design for Identifying Aircraft Satbility and Control Derivatives,' NASA Contractor Report CR-2200, NASA, Washington, D. C., March 1973.

[26] Y. Takahashi, M. J. Rabins and D. M. Auslander. Control and Dynamic Systems. Addison-Wesley Publishing Company, Reading, Massachusetts, pp. 765-790, 1970.

[27] 'TRACE66: Trajectory Analysis and Orbit Determination Program,' SAMSO-TR-71-141, volumes 1-10, The Aerospace Corporation, El Segundo, California, 31 May 1974.

[28] 'Identification of Ship Steering Dynamics Using Inertial Sensors,' IFAC Identification and System Parameter Estimation 1985, pp. 175-182, York, UK, 1985.

An Experimental Study of Nonlinear Dynamic System Identification

Greselda I. Stry[1] and D. Joseph Mook[2]

Department of Mechanical and Aerospace Engineering
State University of New York at Buffalo
Buffalo, New York 14260
716–636–3058

Abstract

A technique for robust identification of nonlinear dynamic systems is developed and illustrated using both digital simulations and analog experiments. The technique is based on the Minimum Model Error optimal estimation approach. A detailed literature review is included in which fundamental differences between the current approach and previous work is described. The most significant feature of the current work is the ability to identify nonlinear dynamic systems without prior assumptions regarding the form of the nonlinearities, in contrast to existing nonlinear identification approaches which usually require detailed assumptions of the nonlinearities. The example illustrations indicate that the method is robust with respect to prior ignorance of the model, and with respect to measurement noise, measurement frequency, and measurement record length.

Introduction

The widespread existence of nonlinear behavior in many dynamic systems is well-documented, e.g, Thompson and Stewart [1]; Nayfeh and Mook [2]. In particular, virtually every problem associated with orbit estimation, flight trajectory estimation, spacecraft dynamics, etc., is known to exhibit nonlinear behavior. Many excellent methods for analyzing nonlinear system models have been developed. However, a key practical link is often overlooked, namely: How does one obtain an accurate mathematical model for the dynamics of a particular complicated nonlinear system?

Accurate dynamic models are necessary for analysis, filter design, and/or control system design. For example, most filter design assumes white process noise, yet many nonlinear effects are inherently non-zero mean; e.g., quadratic nonlinearities are always positive. In order to obtain a model with truly zero mean process noise for filter design

---

1 Graduate Resarch Assistant; NASA Graduate Researcher

2 Assistant Professor

purposes, all of the quadratic terms (and many other nonlinearities) must be well modeled. However, the complexity of many real systems precludes the possibility of accurately constructing a dynamic model purely from analysis using the laws of physics.

Identification is the process of developing an accurate mathematical model for a system, given a set of output measurements. Much work has been done on identification of linear systems, resulting in a number of efficient algorithms. The accuracy and ease of application of these algorithms has given linear identification an enormous popularity. It is, therefore, a common practice to use linearized models to describe nonlinear systems. However, linearization does not work in every application, and even when it does provide a reasonable approximation, the approximation is normally limited to a small region about the operating point of linearization. Consequently, there is a real need for nonlinear identification algorithms. If nonlinearities are a predominant part of a system's behavior, using a linear model to describe such a system leads to inconsistencies ranging from inaccurate numerical results to misrepresentation of the system's qualitative behavior. Since nonlinearities are seldomly easily characterized, identification techniques may prove beneficial in developing accurate mathematical representations of nonlinear systems.

Numerous methods for the identification of nonlinear systems have been developed in the past two decades (Natke, Juang and Gawronski [3]). Most methods fall into one of the following categories:

1. describing the nonlinear system using a linear model
2. the direct equation approach
3. representing the nonlinear system in a series expansion, and obtaining the respective coefficients either by using a regression estimation technique, by minimizing a cost functional, by using correlation techniques, or by some other approach
4. obtaining a graphical representation of the nonlinear term(s), then finding an analytical model for the nonlinearity

With such diversity of nonlinear identification techniques, the choice of an algorithm may be based on criteria such as: iterations required, robustness in the presence of measurement noise, number of measurements needed, robustness with respect to knowledge of the inital conditions, and robustness with respect to initial assumptions regarding the form of the nonlinearity, depending on the needs of the particular application.

Among the methods which linearize the nonlinear system are those presented by Jedner and Unbehauen [4] and Ibanez [5]. Jedner and Unbehauen [4] represent a nonlinear system, which may often function at a number of operating points, by an equivalent number of linear submodels. It is assumed that the system operates at only a few points. Although the model is good for controller design, the point at which the system is operating must be known and the linear models apply only within the operating regions. Ibanez [5] takes a slightly different approach by assuming the system response to be periodic at the forcing frequency. An approximate transfer function is constructed. The

tranfer function is dependent on the amplitude as well as on the exciting frequency and is valid only within the region of exciting frequencies.

The direct equation approach is used by Yasuda, Kawamura and Watanabe [6], [7]. The input and output measurements of a dynamic process are expressed in a Fourier Series using, for example, an FFT algorithm. The system nonlinearity is represented as a sum of polynomials with unknown coefficients. Applying the principle of harmonic balance, the polynomial coefficients as well as the other system parameters are obtained accurately. Knowledge of the nonlinearity is needed to construct the polynomial. Truncation in the Fourier Series expansion of the input or output may lead to error.

The regression estimation approach is used by Billings and Voon [8] and Greblick and Pawlak [9]. Billings and Voon [8] use the NARMAX model (Nonlinear Auto Regressive Moving Average model with eXogenous inputs) to represent the nonlinear system. A stepwise regression method determines the significant terms in the NARMAX model. Then a prediction-error algorithm provides optimal estimates of the final model parameters. Greblick and Pawlak [9] represent the linear dynamic submodel by an ARMA model and the nonlinearities by a Borel function. A non-parametric kernel regression estimation is employed to obtain the final analytical model.

Kortman and Unbehauen [10] and Distefano and Rath [11] use the minimization of an error cost function as a means of obtaining the coefficients of the functions used to represent the nonlinearities. The method presented by Kortman and Unbehauen [10] uses only system input and output information to estimate the polynomial representing the nonlinearities and the parameters of the linear components. It is robust in the presence of noise, although iteration is necessary. Distefano and Rath [11] present two techniques, a non-iterative direct identification and an iterative direct identification. In the first technique, measurement of all variables is required and the model parameters are obtained through the minimization of an error function. In the second technique, iteration is used to minimize a cost function yielding the system parameters in addition to the state trajectories. In Distefano and Rath [11], the nonlinear model form is also taken to be known.

In other techniques, as in statistical linearization, a nonlinear relation is replaced by a linear equivalent gain. Broersen [12] extends the technique of statistical linearization by representing the nonlinearity as a linear combination of a number of arbitrary functions. Correlation techniques are then used to determine the coefficients of these functions. The number and type of functions selected depends on the desired accuracy as well as some knowledge of the system nonlinearity. Reasonable accuracy is obtained in the presence of noise and no iterations are necessary. Although some of the basic properties of the true nonlinear output are preserved, it is limited to only random excitation, and knowledge of all states and forcing terms is required.

In the method of multiple scales (Hanagud, Mayyappa and Craig [13]), a perturbation solution to the nonlinear equation of motion is obtained. An objective function is built

employing an integral least squares approach. The minimization of the functional yields the unknown parameters. Data on only one field variable is necessary, and the method is effective in the presence of high noise. The method of multiple scales, however, is restricted to systems with small damping and slight nonlinearities and, as in most other methods, the form of the nonlinearity is assumed a priori. The method typically requires some algebraic manipulations which may be quite involved, and these manipulations are only valid for a particular assumed nonlinear form. If the assumed nonlinear form is changed, the algebra must be repeated.

Another popular approach is to describe the nonlinear system using the Volterra or Wiener kernels. The Volterra series consists of the summation of impulse responses of increasing dimensionality. The Wiener series is also a set of orthogonal functions in which the input is white gaussian noise. Marmarelis and Udwadia [14], for example, estimate the first and higher order kernels appearing in the Volterra series using correlation techniques. Chen, Ishii and Suzumura [15] use cross-correlation functions in addition to the Volterra and Wiener series to describe nonlinear models and to show the relation between the system inner structure and the series. Although weakly nonlinear systems can be described by the first few kernels, for strongly nonlinear systems these series give accurate numerical results only at the expense of an excessive number of coefficients. This renders the analytical model impractical for control applications.

Other popular series are orthogonal polynomials such as Legendre (Wang and Chan [16]), Chebyshev, and Jacobi (Horn and Chou [17]). Horn and Chou [17] expand the variables of the system into a shifted Jacobi series, reducing the nonlinear state equation into a linear algebraic matrix equation. The unknown parameters of the nonlinear system are then estimated using least squares. Even though the algorithm works well in the presence of noise, the nonlinear form must be known a priori.

Another technique used for the identification of nonlinear systems is the extended Kalman filter. The extended Kalman filter is the linear Kalman filter applied to nonlinear systems by linearizing the nonlinear model into a Taylor series expansion about the estimated state vector. Yun and Shinozuka [18] apply the extended Kalman filter for the parameter estimation of a quadratic term. The state vector is augmented by including the unknown parameters in addition to the state variables. Through a series of iterations, the response, as well as the unknown parameters, are estimated by the Kalman filter. Among its disavantages are high sensitivity to initial conditions, in particular if the initial conditions are barely known.

Hammond, Lo and Seager-Smith [19] use an optimal control technique based on optimal control methods employed for linear system deconvolution. The form of the linear model is assumed to be known as well as the input and the output. A cost functional consisting of the weighted sum of the square of the error (between the actual and estimated output) yields an optimal estimated input. The estimated input and the actual input are

used to obtain the nonlinearity as a function of the state variables. Although no previous assumption is made of the nonlinearities, there is no provision to deal with noise.

In previous papers, the Minimum Model Error algorithm (MME) was explained in detail (Mook and Junkins [20]), modified for nonlinear identification (Mook [21]), and shown to accurately identify exotic nonlinearities in higher order systems (Stry and Mook [22]). In this paper, it is shown how the MME algorithm successfully identifies nonlinearities using experimental data. An analytical model representing a harmonic oscillator with quadratic position feedback is studied. First, output data is obtained from a digital computer simulation of the nonlinear system and the quadratic term is identified to illustrate the accuracy of the technique on a known system. Second, an attempt is made to duplicate the nonlinear model using an analog computer. It is shown that despite the inability of the analog computer to produce a true quadratic term, the Minimum Model Error algorithm is capable of identifying a nonlinear model which accurately reproduces the analog output. The Minimum Model Error method produces a numerically stable identification regardless of the analog data initial conditions or record length.

MME Algorithm

In this section, we review the MME algorithm and how it is used to identify nonlinear dynamic systems. A more detailed explanation may be found in Mook and Junkins [20], Mook [21], and Stry and Mook [22].

The MME may be summarized as follows. Suppose there is a nonlinear system whose exact analytical representation is unknown, but for which output measurements are available. Using "normal" means (analysis, finite elements, etc.), a system model is constructed. As shown in [21]-[22], the MME will work well even if this system model is poor. The MME combines the assumed model with the measurements to determine the correct form of the nonlinear system. A correction term which represents the error in the model is added to the assumed model and a cost functional is formed. Minimization of the cost functional yields the model error. Subsequently, a least squares fit is performed on the error term to determine the correct form of the nonlinear system.

Consider a forced nonlinear dynamic system which may be modeled in state-space form by the equation

$$\dot{\underline{x}}(t) = A\underline{x}(t) + \underline{F}(t) + \underline{f}(\underline{x}(t), \dot{\underline{x}}(t)) \tag{1}$$

where $\underline{x}(t)$ is the $n \times 1$ state vector consisting of the system states, A is the $n \times n$ state matrix, $\underline{F}(t)$ is an $n \times 1$ vector of known external excitation, and $\underline{f}(\underline{x}(t), \dot{\underline{x}}(t))$ is an $n \times 1$ vector which includes all of the system nonlinearities. State-observable discrete time domain measurements are available for this system in the form

$$\tilde{\underline{y}}(t_k) = \underline{g}_k(\underline{x}(t_k), t_k) + \underline{v}_k, \qquad t_0 \leq t_k \leq t_f \tag{2}$$

where $\underline{\tilde{y}}(t_k)$ is an $m \times 1$ measurement vector at time t_k, $\underline{g}_k$ is the accurate model of the measurement process, and $\underline{v}_k$ represents measurement noise. $\underline{v}_k$ is assumed to be a zero-mean, gaussian distributed process of known covariance R_k. The measurement vector $\underline{\tilde{y}}(t_k)$ may contain one or more of the system states. To implement MME, assume that a model, which is generally not the true system model because of the difficulties inherent in obtaining the true system model, is constructed in state-vector form as

$$\dot{\underline{x}}(t) = A\underline{x}(t) + \underline{F}(t) \tag{3}$$

Here, we show a linear model because in practice, linearization is the most common approach to modeling nonlinear systems. MME uses the assumed linear model in (3) and the noisy measurements in (2) to find the model error.

The model error, which includes the unknown nonlinear term(s) of the system, is represented by the addition of a correction term to the assumed linear model as

$$\dot{\underline{x}}(t) = A\underline{x}(t) + \underline{F}(t) + \underline{d}(t) \tag{4}$$

where $\underline{d}(t)$ is the $n \times 1$ correction term (dynamic model error) to be estimated later.

A cost functional, J, that consists of the weighted integral square of the correction term plus the weighted sum square of the measurement-minus-estimated measurement residuals, is formed:

$$J = \sum_{k=1}^{M} \left\{ [\underline{\tilde{y}}(t_k) - \underline{g}_k(\hat{\underline{x}}(t_k), t_k)]^T R_k^{-1} [\underline{\tilde{y}}(t_k) - \underline{g}_k(\hat{\underline{x}}(t_k), t_k)] \right\}$$

$$+ \int_{t_o}^{t_f} \underline{d}(\tau)^T W \underline{d}(\tau) d\tau \tag{5}$$

where M is the number of measurement times, $\hat{\underline{x}}(t_k)$ is the estimated state vector and W is a weight matrix to be determined.

J is minimized with respect to the correction term, $\underline{d}(t)$. The necessary conditions for the minimization lead to the following two point boundary value problem (TPBVP), (see Geering [23]),

$$\dot{\underline{x}}(t) = A\underline{x}(t) + \underline{F}(t) + \underline{d}(t) \tag{5a}$$

$$\dot{\underline{\lambda}}(t) = -A^T \underline{\lambda}(t) \tag{5b}$$

$$\underline{d}(t) = -\frac{1}{2} W \underline{\lambda}(t) \tag{5c}$$

$$\underline{\lambda}(t_k^+) = \underline{\lambda}(t_k^-) + 2H_k R_k^{-1} [\underline{\tilde{y}}(t_k) - \underline{g}_k(\hat{\underline{x}}(t_k), t_k)] \tag{5d}$$

$$H_k = \frac{\delta g}{\delta \underline{x}} \Big|_{\hat{\underline{x}}(t_k), t_k}$$

$$\underline{x}(t_o) = \underline{x}_o \quad or \quad \underline{\lambda}(t_o) = 0 \tag{5e}$$

$$\underline{x}(t_f) = \underline{x}_f \quad or \quad \underline{\lambda}(t_f) = 0 \tag{5f}$$

where $\underline{\lambda}(t)$ is a vector of costates (Lagrange multipliers). Estimates of the states and of the dynamic model error are produced by the solution of this two-point boundary value problem. The estimates depend on the particular value of W. The solution is repeated until a value of W is obtained which produces state estimates which satisfy the "covariance constraint", explained next.

According to the covariance constraint, the measurement-minus-estimated measurement residual covariance matrix must match the measurement-minus-truth error covariance matrix. This may be written as

$$[\underline{\tilde{y}}(t_k) - \underline{g}_k(\underline{\hat{x}}(t_k), t_k)]^T [\underline{\tilde{y}}(t_k) - \underline{g}_k(\underline{\hat{x}}(t_k), t_k)] \approx R_k \tag{6}$$

During the minimization, the weight W is varied until the state estimates satisfy the covariance constraint, i.e., the left hand side of Eq. (6) is approximately equal to the right hand side. The correction term or model error is, therefore, the minimum adjustment to the model required for the estimated states to predict the measurements with approximately the same covariance as the measurement error.

The TPBVP represented by Eqs. (5a) to (5f) contains jumps in the costates and, consequently, in the correction term. As evident from Eq. (5d), the size of the jump is directly proportional to the measurement residual at each measurement time. The noisier the measurements, the larger the jump size. A multiple shooting algorithm, developed by Mook and Lew [24], converts this jump-discontinuous TPBVP into a set of linear algebraic equations which may be solved using any linear equation solver. Multiple shooting also facilitates the analysis of a large number of measurements, by processing the solution at the end of every set of jumps.

After W has been determined such that the state estimates satisfy the covariance constraint, the final step in the identification procedure is to use a least squares algorithm to fit the model error $d(t)$ to the unknown dynamic term(s). The error is expanded into some combination of linear and nonlinear terms, for example,

$$d(t) = \alpha x(t) + \beta x^2(t) + \gamma x^3(t) + \dots \tag{7}$$

where α, β, γ, ... are unknown coefficients to be determined by least squares. Presumably, the least squares fit of Eq. (7) will find zero coefficients for the terms in the expansion which are not part of the true model, and nonzero coefficients for the actual model terms. Eq. (7) may be sampled repeatedly to obtain

$$\begin{aligned}
d(t_1) &= \alpha x(t_1) + \beta x^2(t_1) + \gamma x^3(t_1) + \dots \\
d(t_2) &= \alpha x(t_2) + \beta x^2(t_2) + \gamma x^3(t_2) + \dots \\
\vdots &= \qquad\qquad \vdots \\
d(t_l) &= \alpha x(t_l) + \beta x^2(t_l) + \gamma x^3(t_l) + \dots
\end{aligned}$$

or, in matrix form,

$$\underline{D}_{l\times 1} = M_{l\times p}\underline{P}_{p\times 1} \tag{8}$$

where $\underline{P} = [\alpha \quad \beta \quad \gamma \quad \ldots]^T$ is the vector of coefficients for the terms in $d(t)$. Since estimates of $d(t)$ are available continuously throughout the time domain, the parameter l may be chosen quite large to improve the least squares fit. Generally, because of the jump discontinuities in the model error estimates at the measurement times, it is desirable to pick the least squares sampling times in Eq. (8) at points other than the measurement times. The least squares estimate is found by minimizing the following cost functional with respect to P:

$$\Phi = [\underline{D} - M\underline{P}]^T[\underline{D} - M\underline{P}] \tag{9}$$

The solution is given by

$$\underline{P} = (M^T M)^{-1} M^T \underline{D} \tag{10}$$

The multiple shooting algorithm presented by Mook and Lew [24] was used to obtain the MME solutions used in the tests presented in this paper. It was assumed in the examples that MME obtained the dynamic error term without knowledge of the boundary conditions on $\underline{x}$, so some distortion of the correction term at the initial and final times was expected due to the constraints of Eqs. (5e-5f), i.e., by assuming no state knowledge is available at t_0 or t_f, we constrain $\lambda(t_0) = 0$ and $\lambda(t_f) = 0$. Therefore, in all test cases, the initial and final ten percent of the correction term data was ignored in the least squares fit.

Application Examples

Two nonlinear equations of motion were studied, which represent the motion of an undamped harmonic oscillator with different amounts of quadratic position feedback (identical equations may arise in other physical systems as well). The equations in state space form are

$$\begin{pmatrix} \dot{x} \\ \ddot{x} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} + \begin{pmatrix} 0 \\ -0.526x^2 \end{pmatrix} \tag{11}$$

$$\begin{pmatrix} \dot{x} \\ \ddot{x} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ \dot{x} \end{pmatrix} + \begin{pmatrix} 0 \\ -1.137x^2 \end{pmatrix} \tag{12}$$

where x is position, and the dot indicates differentiation with respect to time. No forcing was applied.

In the following discussion, Eq. (11) is denoted Model A and Eq. (12) is denoted Model B. Different initial conditions were used for each system, for a total of five different tests. These are shown in Table 1.

Table 1. List of conditions used for each test

Test #	$x(t_0)$	$\dot{x}(t_0)$	x^2
A1	0.00	0.08	-0.526
A2	0.00	0.26	-0.526
A3	0.00	0.52	-0.526
B1	0.00	0.08	-1.137
B2	0.00	0.26	-1.137

To utilize MME, the linear part of Eqs. (11) and (12) was assumed to be known, rendering the model error equivalent to the nonlinear term, $c * x^2$. Data for the MME nonlinear identification was generated from two different sources. First, noiseless position measurements were gathered from a digital computer simulation for all five tests. Application of MME to the measurements yielded an accurate estimate of the nonlinear term in each case. Then, an attempt was made to duplicate each system on an analog computer. Even though the analog computer did not reliably reproduce the quadratic term, the position measurements for all five tests were recorded and nonlinear models identified. MME proved capable of identifying accurate nonlinear models for the analog output.

Digital computer simulation results

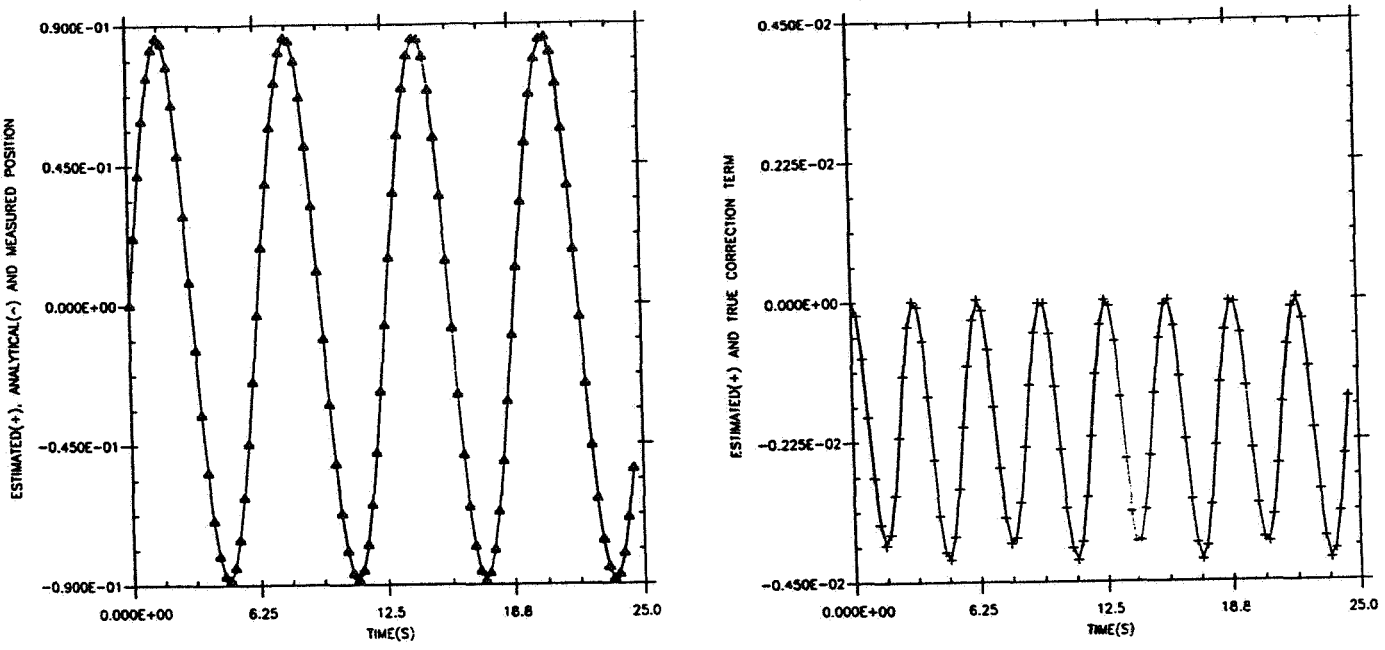
One hundred noiseless position measurements were generated on a VAX 780 for the five test cases shown in Table 1.. A sampling rate of 4 Hertz was used. Three terms were employed in the least squares fit: x, $\dot{x}$ and x^2. The resulting numerical values are shown in Table 2.

Table 2. Least Squares estimates of the dynamic model error employing analytically generated measurements in the MME algorithm.

Test #	MME x	MME $\dot{x}$	MME x^2	Analytic x	Analytic $\dot{x}$	Analytic x^2
A1	4.24×10^{-4}	4.90×10^{-4}	-0.526	0.0	0.0	-0.526
A2	1.21×10^{-3}	-3.00×10^{-4}	-0.527	0.0	0.0	-0.526
A3	2.49×10^{-3}	2.35×10^{-3}	-0.528	0.0	0.0	-0.526
B1	-2.86×10^{-4}	-6.39×10^{-3}	-1.138	0.0	0.0	-1.137
B2	-3.46×10^{-4}	-6.80×10^{-3}	-1.138	0.0	0.0	-1.137

MME identifies the quadratic term with great accuracy in all five tests. A plot of the estimated, analytical and measured position is shown in Figure (1a) for test case A1.

Figure (1b) presents the predicted model error and the dynamic model error estimated by MME for test case A1. MME estimates the position and correction term with great accuracy. Similar results are obtained for tests A2, A3, B1, and B2, but are not shown.



(a) (b)

Figure 1 (a) Analytical, measured (Δ), and MME estimated position for test A1 using digital computer simulated measurements. The analytical and MME estimates are essentially identical (solid line). (b) MME estimate (+) and actual model error.

Analog computer results

Three hundred fifty position measurements were generated on a Comdyna GP-6 analog computer for all five test cases. One hundred measurements with a sampling rate of 4 Hertz were used in the analysis. The identification procedure yielded the numerical values shown in Table 3.

Table 3. Least Square estimates of the nonlinear terms using measurements generated by the analog computer.

Test #	MME x	MME $\dot{x}$	MME x^2
A1	-0.20	1.04×10^{-2}	-6.17
A2	-5.06×10^{-2}	-1.89×10^{-3}	-1.322
A3	-6.44×10^{-3}	-4.81×10^{-3}	-0.689
B1	0.10	-3.78×10^{-3}	-3.47
B2	2.55×10^{-2}	7.42×10^{-3}	-1.265

The numerical results for the least squares fit of the error term did not match the analytically predicted coefficients. The reason for the numerical discrepancy was the analog's failure to produce a dependable quadratic term. Table 4 shows some position values squared by the analog. The analog consistently produced an error in the quadratic term. The recorded data, although containing errors due to quadratic term, is believed to be practically noiseless.

Table 4. Quadratic term produced by the analog computer.

x	x^2	Analog
2.00	4.00	4.30
2.50	6.25	7.00
3.00	9.00	9.50

Figures (2)-(6) show the analytical position, analog measurements and position predicted by the MME analysis for all analog tests. The MME identification produced good state estimates and a model which matched the measured data much better than the analytical models in Eqs. (11) and (12).

Note, these results were obtained without knowledge of the initial or final state vector value. As shown in Eqs. (5e) and (5f), by setting the initial and final costate values to zero, no knowledge of the initial or end conditions are necessary. Also, the same results presented in Table 3 were obtained when using all three hundred and fifty measurements instead of one hundred measurements in the MME algorithm.

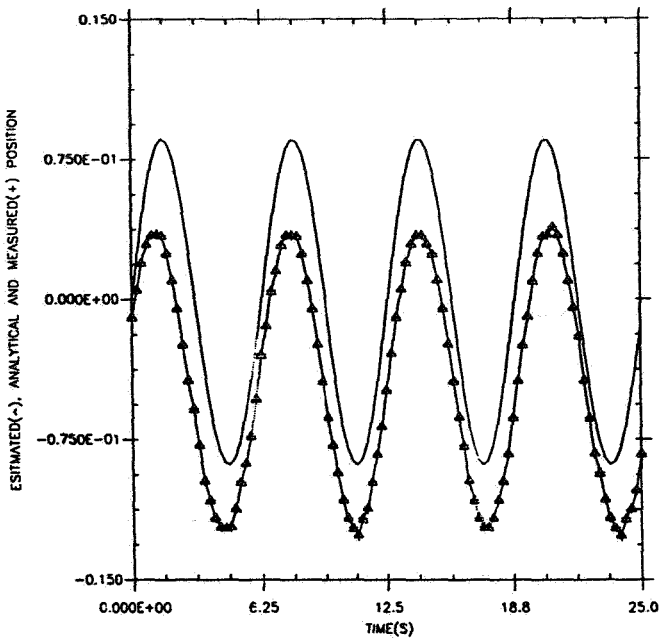


Figure 2 Analytical, measured (Δ), and MME estimated position for test A1 using analog computer measurements. The MME estimates are essentially identical to the measurements

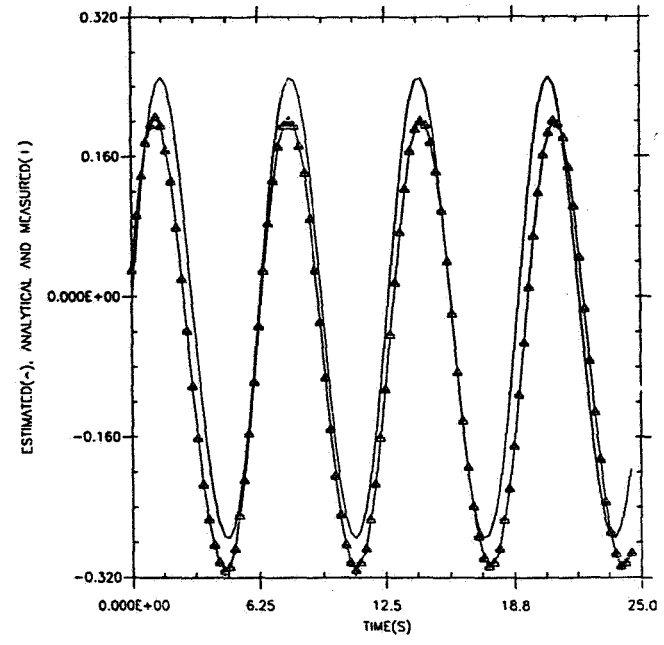


Figure 3 Analytical, measured (Δ), and MME estimated position for test A2 using analog computer measurements. The MME estimates are essentially identical to the measurements

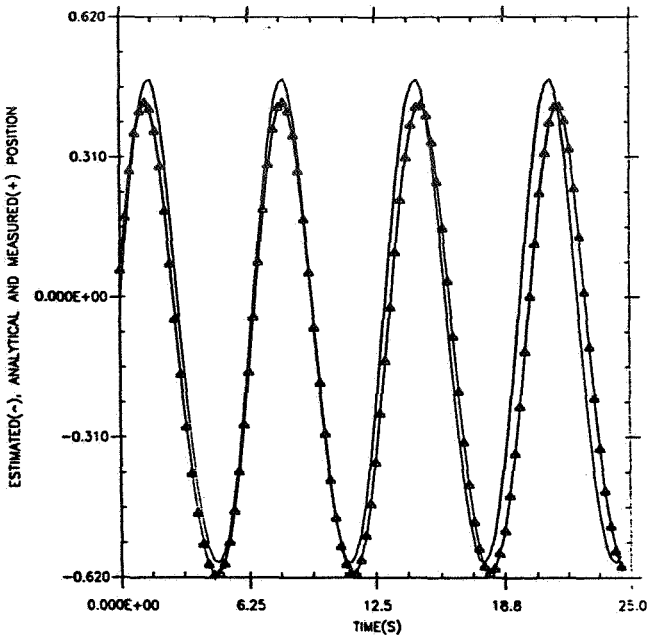


Figure 4 Analytical, measured (Δ), and MME estimated position for test A3 using analog computer measurements. The MME estimates are essentially identical to the measurements

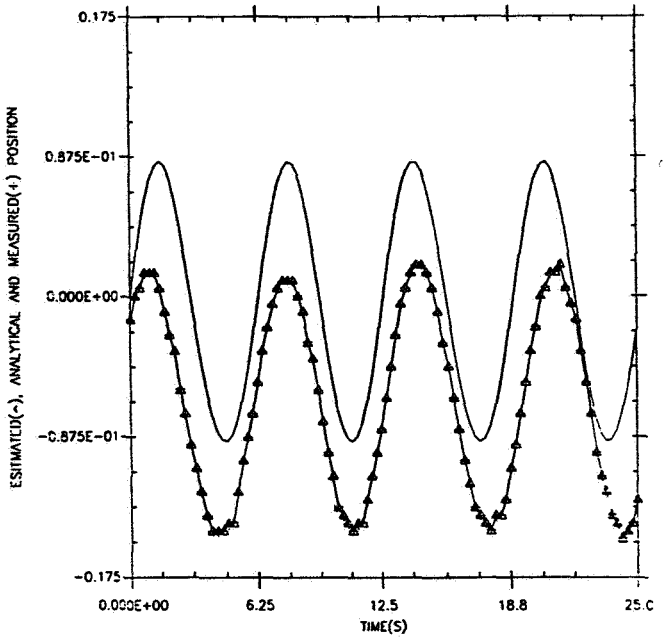


Figure 5 Analytical, measured (Δ), and MME estimated position for test B1 using analog computer measurements. The MME estimates are essentially identical to the measurements

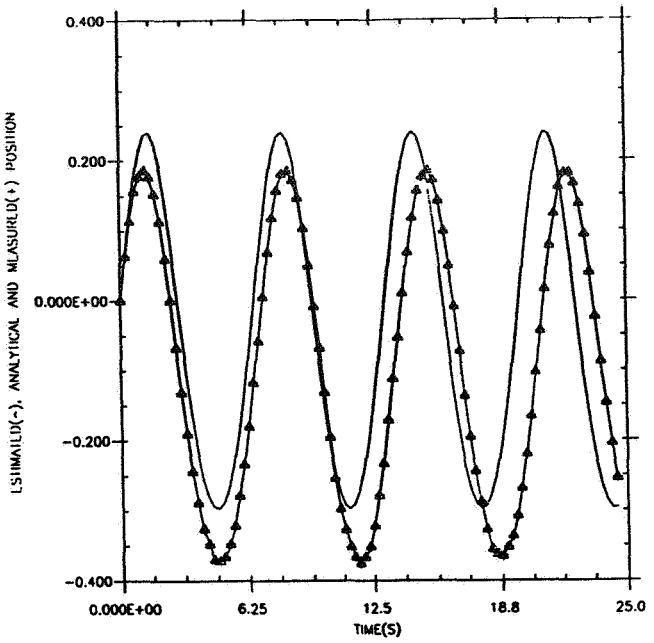


Figure 6 Analytical, measured (Δ), and MME estimated position for test B2 using analog computer measurements. The MME estimates are essentially identical to the measurements

Conclusion

In this paper, an MME—based algorithm was used to identify the quadratic term of a nonlinear harmonic oscillator. For demonstration purposes, data was obtained from two sources. Output data obtained from a digital computer simulation was used to verify the accuracy of the method. Then, data from an analog computer was used as a test of the method on "real" data. It was shown that despite the inability of the analog computer to reproduce a true quadratic term, the MME algorithm was capable of identifying a nonlinear model which accurately reproduced the analog output. This result indicates that the method is robust with respect to (lack of) a priori knowledge of the system dynamics. The identification was accurate regardless of initial conditions or data record length, indicating that the method is also robust with respect to those variables.

References

1. Thompson, J.M.T., and Stewart, H.B., **Nonlinear Dynamics and Chaos**, Wiley, New York, 1986.
2. Nayfeh, A. H., and Mook, D. T., **Nonlinear Oscillations**, Wiley, New York, 1979.

3. Natke, H.G., Juang, J.-N., and Gawronski, W., "Identification of Nonlinear Mechanical Systems: A Brief Review," NASA Langley Research Center, Hampton, VA, U.S.A. .
4. Jedner, U., and Unbehauen, H., "Identification of a Class of Nonlinear Systems by Parameter Estimation of a Linear Multi-Model," *IMACS, Modeling and Simulation for Control of Lumped and Distributed Parameter Systems*, June 3-6, 1986, pp. 11–15.
5. Ibanez, P., "Identification of Dynamic Parameters of Linear and Non-Linear Structural Models from Experimental Data," *Nuclear Engineering and Design*, Vol. 25, 1973, pp. 32–41.
6. Yasuda, K., Kawamura, S., and Watanabe, K., "Identification of Nonlinear Multi-Degree-of-Freedom Systems (Identification Under Noisy Measurements)," *JSME International Journal* , Vol. 3, No. 1, 1988, pp. 502–509.
7. Yasuda, K., Kawamura, S., and Watanabe, K., "Identification of Nonlinear Multi-Degree-of-Freedom Systems (Presentation of an Identification technique)," *JSME International Journal* , Vol. 3, No. 1, 1988, pp. 8–14.
8. Billings, S.A., and Voon, W.S.F., "A Prediction-Error and Stepwise Regression Estimation Algorithm for Nonlinear Systems," *Int. J. Control*, Vol. 44, No. 3, 1986, pp. 803–822.
9. Greblick, W., and Pawlak, M., "Hammerstein System Identification by Non-Parametric Regression Estimation," *Int. J. Control*, Vol. 45, No. 1, 1987, pp. 343–354.
10. Kortmann, M., and Unbehauen, H., "Application of a Recursive Prediction Error Method to the Identification of Nonlinear Systems Using the Weiner Model," *IMACS, Modeling and Simulation for Control of Lumped and Distributed Parameter Systems*, June 3-6, 1986, pp. 3–9.
11. Distefano, N., and Rath, A., "System Identification in Nonlinear Seismic Dynamics," *Computer Methods in Applied Mechanics and Engineering*, Vol. 5, 1975, pp. 353–372.
12. Broersen, P.M.T., "Estimation of Parameters of Non-Linear Dynamical Systems," *Int. J. Non-Linear Mechanics* , Vol. 9, 1974, pp. 355–361.
13. Hanagud, S.V., Meyyappa, M., and Craig, J.I., "Method of Multiple Scales and Identification of Nonlinear Dynamical Systems," *AAIA J.*, Vol. 23, No. 5, 1985, pp. 802–807.
14. Marmarelis, P.Z., and Udwadia, F.E., "The Identification of Building Structural Systems-PartII: The Nonlinear Case," *Bulletin of the Seismological Society of America*, Vol. 66, 1979, pp. 153–171.
15. Chen, H., Ishii, N., and Suzumura, N., "Structural Classification of Non-Linear Sytems by Input and Output," *Int. J. Systems Sci.*, Vol. 17, No. 5, 1986, pp. 741–774.
16. Wang, M.L., and Chang, R.Y., "Model Reduction and Control System Design by Shifted Legendre Polynomial Functions," *A.S.M.E. J. Dynam. Sys. Meas. Control*,

Vol. 105, 1983, pp. 52–55.
17. Horn, I-R., and Chou, J-H., "Analysis and Identification of Non-Linear Systems via Shifted Jacobi Series," *Int. J. Control*, Vol. 45, No. 1, 1987, pp. 279–290.
18. Yun, C-B., and Shinozuka, M., "Identification of Nonlinear Structural Dynamic Systems," *J. Struct. Mech.*, Vol. 8, No. 2, 1980, pp. 187–203.
19. Hammond, J.K., Lo, H.R., and Seager-Smith, E.J., "Identification of Nonlinearities in Vibrating Systems Using Optimal Control Techniques," *IMAC*, 1987, pp. 1467–1473.
20. Mook, D.J., and Junkins, "Minimum Model Error Estimation for Poorly Modeled Dynamic Systems", *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 11, No. 3, 1988, pp. 256–261.
21. Mook, D.J., "Estimation and Identification of Nonlinear Dynamic Systems," *AIAA J.*, Vol. 27, No. 7, 1989, pp. 968–974.
22. Stry, G.I., and Mook, D.J., "Identification of Higher-Order Systems with Exotic Nonlinearities," in review.
23. Geering, H.P., "Continuous Time Optimal Control Theory for Cost Functionals Including Discrete State Penalty Terms," *IEEE Trans. A.C.*, Vol. AC-21, 1976, pp. 866–869.
24. Mook, D.J., and Lew, J-H., "Multiple Shooting Algorithm for Jump Discontinuous Problems in Optimal Control and Estimation ," to appear in the *IEEE Trans. A. C.*.

Downdating a Time-Varying Square Root Information Filter

Ronald J. Muellerschoen †

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91109

Abstract

A new method to efficiently downdate an estimate and covariance generated by a discrete time Square Root Information Filter (SRIF) is presented. The method combines the QR factor downdating algorithm of Gill [3] and the decentralized SRIF algorithm of Bierman [4]. Efficient removal of either measurements or *a priori* information is possible without loss of numerical integrity. Moreover, the method includes features for detecting potential numerical degradation. Performance on a 300 parameter system with 5800 data points shows that the method can be used in real time and hence is a promising tool for interactive data analysis. Additionally, updating a time-varying SRIF filter with either additional measurements or *a priori* information proceeds analogously.

Introduction

A typical 24 hour data arc for the GPS demonstration on TOPEX/POSEIDON [6] will contain $\approx$ 30,000 data points. To process these data points on a VAX 8530, a sequential SRIF filter will require nearly 3 hours of CPU time. Upon processing, should outliers be discovered in the data, due to cycle slips, atmospheric fluctuations, multipath, *etc* ..., they must be removed from the data arc and another 3 hours of CPU time would then be needed for reprocessing. The method presented here permits efficient removal of these outlying data points without reprocessing the entire data set.

The discussion is organized in the following manner. First, a decentralized approach to updating a time-varying SRIF with a single measurement is presented. From this development, it is evident how to remove a measurement from a time-varying SRIF. Next, the one-component-at-a-time process noise methodology is applied in the actual implementation. The necessary linear combination of the data equation and the smoothing coefficients is then presented. Finally, the efficiency and the numerical integrity of the method is discussed.

Updating a Time-Varying SRIF with a Single Measurement

The goal is to find estimates $(x_0, x_1, \ldots, x_N)$ to minimize the the least-squares performance functional

$$J^N(x_0, x_1, \ldots, x_N) = \|\tilde{R}_x(0)x_0 - \tilde{z}_x(0)\|^2 + \sum_{j=0}^{N-1} \|R_w(j)w_j\|^2 + \sum_{j=0}^{N} \|A_j x_j - z_j\|^2 + \|H_k x_k - y_k\|^2 \quad (1)$$

† Member of Technical Staff, Navigation Systems

for the linear discrete-time dynamic model

$$x_{j+1} = \phi_j x_j + G w_j \qquad j = 0, N-1 \tag{2}$$

with measurements

$$z_j = A_j x_j + \nu_j \qquad j = 0, N \tag{3}$$

and the single measurement

$$y_k = H_k x_k + \nu_k \qquad 0 \le k \le N \tag{4}$$

where $\tilde{()}_x(j)$ represents *a priori* information between time updates j and j+1. The data noise ν_j and process noise w_j are independent mean zero white Gaussian noise processes with covariances I and $R_w(j)^{-1}R_w(j)^{-T}$, respectively. Of note, when a pseudo-epoch state formulation is used ϕ_j and G have a simplified structure.

$$\begin{pmatrix} x_{j+1} \\ p_{j+1} \end{pmatrix} = \begin{pmatrix} I & V_p(j) \\ 0 & M_j \end{pmatrix} \begin{pmatrix} x_j \\ p_j \end{pmatrix} + \begin{pmatrix} 0 \\ I \end{pmatrix} w_j \qquad j = 0, N-1 \tag{5}$$

Ignoring the additional data equation $y_k = H_k x_k + \nu_k$, the SRIF alternates between measurement updates

$$\begin{aligned} J^N(x_0, x_1, \ldots, x_N) = & \sum_{i=0}^{j-1} \|e_i\|^2 + \| \begin{pmatrix} \tilde{R}_x(j) \\ A_j \end{pmatrix} x_j - \begin{pmatrix} \tilde{z}_x(j) \\ z_j \end{pmatrix} \|^2 + \\ & \sum_{i=j+1}^{N} \|A_i x_i - z_i\|^2 + \|H_k x_k - y_k\|^2 + \\ & \sum_{i=0}^{j-1} \|R_w^*(i) w_i + R_{wx}^*(i+1) x_{i+1} - z_w^*(i)\|^2 + \sum_{i=j}^{N-1} \|R_w(i) w_i\|^2 \end{aligned} \tag{6}$$

$$\begin{aligned} = & \sum_{i=0}^{j} \|e_i\|^2 + \|\hat{R}_j x_j - \hat{z}_j\|^2 + \\ & \sum_{i=j+1}^{N} \|A_i x_i - z_i\|^2 + \|H_k x_k - y_k\|^2 + \\ & \sum_{i=0}^{j-1} \|R_w^*(i) w_i + R_{wx}^*(i+1) x_{i+1} - z_w^*(i)\|^2 + \sum_{i=j}^{N-1} \|R_w(i) w_i\|^2 \end{aligned} \tag{7}$$

for $j = 0, N$ and time updates

$$\begin{aligned} J^N(x_0, x_1, \ldots, x_N) = & \| \begin{pmatrix} R_w(j) & 0 \\ -\hat{R}_j \phi_j^{-1} G & \hat{R}_j \phi_j^{-1} \end{pmatrix} \begin{pmatrix} w_j \\ x_{j+1} \end{pmatrix} - \begin{pmatrix} 0 \\ \hat{z}_j \end{pmatrix} \|^2 + \\ & \sum_{i=0}^{j-1} \|R_w^*(i) w_i + R_{wx}^*(i+1) x_{i+1} - z_w^*(i)\|^2 + \sum_{i=j+1}^{N-1} \|R_w(i) w_i\|^2 + \end{aligned}$$

$$\sum_{i=j+1}^{N} \|A_i x_i - z_i\|^2 + \|H_k x_k - y_k\|^2 + \sum_{i=0}^{j} \|e_i\|^2 \tag{8}$$

$$= \|\tilde{R}_x(j+1)x_{j+1} - \tilde{z}_x(j+1)\|^2 +$$
$$\sum_{i=0}^{j} \|R_w^*(i)w_i + R_{wx}^*(i+1)x_{i+1} - z_w^*(i)\|^2 + \sum_{i=j+1}^{N-1} \|R_w(i)w_i\|^2 +$$
$$\sum_{i=j+1}^{N} \|A_i x_i - z_i\|^2 + \|H_k x_k - y_k\|^2 + \sum_{i=0}^{j} \|e_i\|^2 \tag{9}$$

for $j = 0, N-1$. The e_i term represents the residual sum of squares from the i'th mesurement update; the $()^*$ terms represent the smoothing coefficients. Additionally, the notation $\widehat{()}_j$ represents quantities that include data up to time $j+1$.

To update the filter with the additional data equation $y_k = H_k x_k + \nu_k$, it is necessary to express this equation in terms of x_{k+1} and w_k.

$$\|H_k x_k - y_k\|^2 = \| \begin{pmatrix} -H_k \phi_k^{-1} G & H_k \phi_k^{-1} \end{pmatrix} \begin{pmatrix} w_k \\ x_{k+1} \end{pmatrix} - y_k \|^2$$
$$= \| \begin{pmatrix} H_w^*(k) & H_{wx}^*(k+1) \end{pmatrix} \begin{pmatrix} w_k \\ x_{k+1} \end{pmatrix} - y^*(k) \|^2 \tag{10}$$

This equation is then merged with the smothing coefficients of the SRIF.

$$J^N(x_0, x_1, \ldots, x_N) = \sum_{j=0}^{N} \|e_j\|^2 + \|\hat{R}_N x_N - \hat{z}_N\|^2 +$$
$$\sum_{j=0}^{k-1} \|R_w^*(j)w_j + R_{wx}^*(j+1)x_{j+1} - z_w^*(j)\|^2 +$$
$$\sum_{j=k+1}^{N-1} \|R_w^*(j)w_j + R_{wx}^*(j+1)x_{j+1} - z_w^*(j)\|^2 +$$
$$\| \begin{pmatrix} R_w^*(k) & R_{wx}^*(k+1) \\ H_w^*(k) & H_{wx}^*(k+1) \end{pmatrix} \begin{pmatrix} w_k \\ x_{k+1} \end{pmatrix} - \begin{pmatrix} z_w^*(k) \\ y^*(k) \end{pmatrix} \|^2 \tag{11}$$

$$J^N(x_0, x_1, \ldots, x_N) = \sum_{j=0}^{N} \|e_j\|^2 + \|\hat{R}_N x_N - \hat{z}_N\|^2 +$$
$$\sum_{j=0}^{k-1} \|R_w^*(j)w_j + R_{wx}^*(j+1)x_{j+1} - z_w^*(j)\|^2 +$$
$$\sum_{j=k+1}^{N-1} \|R_w^*(j)w_j + R_{wx}^*(j+1)x_{j+1} - z_w^*(j)\|^2 +$$
$$\| \begin{pmatrix} R_w^{**}(k) & R_{wx}^{**}(k+1) \\ 0 & H_{k+1} \end{pmatrix} \begin{pmatrix} w_k \\ x_{k+1} \end{pmatrix} - \begin{pmatrix} z_w^{**}(k) \\ y_{k+1} \end{pmatrix} \|^2 \tag{12}$$

The ()** terms replace the old smoothing coefficients and $y_{k+1} = H_{k+1}x_{k+1} + \nu_{k+1}$ is the ensuing additional data equation. The time updates in equation (10) and the merges in equations (11) and (12) are then repeated for $k+2, k+3, \ldots, N$. In the end, the terminal data equation $y_N = H_N x_N + \nu_N$ is then merged with the terminal SRIF array.

$$J^N(x_0, x_1, \ldots, x_N) = \sum_{j=0}^{k-1} \|R_w^*(j)w_j + R_{wx}^*(j+1)x_{j+1} - z_w^*(j)\|^2 + \sum_{j=k}^{N-1} \|R_w^{**}(j)w_j + R_{wx}^{**}(j+1)x_{j+1} - z_w^{**}(j)\|^2 + \left\| \begin{pmatrix} \hat{R}_N \\ H_N \end{pmatrix} x_N - \begin{pmatrix} \hat{z}_N \\ y_N \end{pmatrix} \right\|^2 + \sum_{j=0}^{N} \|e_j\|^2 \quad (13)$$

$$J^N(x_0, x_1, \ldots, x_N) = \sum_{j=0}^{k-1} \|R_w^*(j)w_j + R_{wx}^*(j+1)x_{j+1} - z_w^*(j)\|^2 + \sum_{j=k}^{N-1} \|R_w^{**}(j)w_j + R_{wx}^{**}(j+1)x_{j+1} - z_w^{**}(j)\|^2 + \|\check{R}_N x_N - \check{z}_N\|^2 + \sum_{j=0}^{N} \|e_j\|^2 + \|\check{e}_N\|^2 \quad (14)$$

The notation $\check{()}_N$ represents quantities resulting from merging the terminal data equation with the terminal SRIF array. The filter estimate and covariance are then $\check{R}_N^{-1}\check{z}_N$ and $\check{R}_N^{-1}\check{R}_N^{-T}$, respectively. To obtain the smoothed estimates and covariances, the Dyer-McReynolds smoother [2] would then operate with the ()* terms for $j = 0, k-1$ and with the ()** terms for $j = k, N-1$.

Downdating a Time-Varying SRIF with a Single Measurement

The last lines of equations (11) and (12) contain the key to reversing this process. If the data equation $y_k = H_k x_k + \nu_k$ has been included in the SRIF, then the ()** terms of (12) are available. Additionally, the data equation $y^*(k) = H_w^*(k)w_k + H_{wx}^*(k+1)x_{x+1} + \nu_k$ in (11) can be generated from the measurement that is to be removed. It is possible, as shown below, to solve for a transformation to produce the upper ()* terms of (11) and the ensuing data equation $y_{k+1} = H_{k+1}x_{k+1} + \nu_{k+1}$ in (12). The upper ()* terms of (11) are then the downdated smoothing coefficients and replace the old smoothing coefficients. This process is repeated for $j = k+2, k+3, \ldots, N$. The terminal data equation $y_N = H_N x_N + \nu_N$ is then downdated from the terminal SRIF array [3].

Implementation

It is efficient to perform the process noise updates one component at at time. For example, in a four state filter with three process noise parameters, the smoothing coefficients from time update $j+1$ can be expressed as

$$\begin{matrix} R_{p1}^{**}(j) & R_{p1,x}^{**}(j+1) & R_{p1}^{**}(j+1) & R_{p1,p2}^{**}(j) & R_{p1,p3}^{**}(j) & z_{p1}^{**}(j) \\ R_{p2}^{**}(j) & R_{p2,x}^{**}(j+1) & R_{p2,p1}^{**}(j+1) & R_{p2}^{**}(j+1) & R_{p2,p3}^{**}(j) & z_{p2}^{**}(j) \\ R_{p3}^{**}(j) & R_{p3,x}^{**}(j+1) & R_{p3,p1}^{**}(j+1) & R_{p3,p2}^{**}(j+1) & R_{p3}^{**}(j+1) & z_{p3}^{**}(j) \end{matrix} \quad j = 0, N-1 \quad (15)$$

where the ()** terms are a reminder that all data has been processed in the SRIF. Note that the update order of the process noise parameters is p1, p2, p3.

The first step is to express the data equation to be removed in terms of x_{k+1} using equation (5).

$$\| \begin{pmatrix} H_{x_k} & H_{p1_k} & H_{p2_k} & H_{p3_k} \end{pmatrix} \begin{pmatrix} x_k \\ p1_k \\ p2_k \\ p3_k \end{pmatrix} - y_k \|^2 =$$

$$\| \begin{pmatrix} H_{x_k} & H_{p1_k} - H_{x_k} V_{p1}(k) & H_{p2_k} - H_{x_k} V_{p2}(k) & H_{p3_k} - H_{x_k} V_{p3}(k) \end{pmatrix} \begin{pmatrix} x_{k+1} \\ p1_k \\ p2_k \\ p3_k \end{pmatrix} - y_k \|^2 =$$

$$\| \begin{pmatrix} H_x^*(k+1) & H_{p1}^*(k) & H_{p2}^*(k) & H_{p3}^*(k) \end{pmatrix} \begin{pmatrix} x_{k+1} \\ p1_k \\ p2_k \\ p3_k \end{pmatrix} - y^*(k) \|^2 \tag{16}$$

The smoothing coefficients associated with $p1$ are then downdated with an orthogonal transformation.

$$T \begin{pmatrix} R_{p1}^{**}(k) & R_{p1,x}^{**}(k+1) & R_{p1}^{**}(k+1) & R_{p1,p2}^{**}(k) & R_{p1,p3}^{**}(k) & z_{p1}^{**}(k) \\ 0 & \bar{H}_{x_{k+1}} & \bar{H}_{p1_{k+1}} & \bar{H}_{p2_k} & \bar{H}_{p3_k} & \bar{y}_k \end{pmatrix}$$
$$= \begin{pmatrix} R_{p1}^{*}(k) & R_{p1,x}^{*}(k+1) & R_{p1}^{*}(k+1) & R_{p1,p2}^{*}(k) & R_{p1,p3}^{*}(k) & z_{p1}^{*}(k) \\ H_{p1}^{*}(k) & H_x^{*}(k+1) & 0 & H_{p2}^{*}(k) & H_{p3}^{*}(k) & y^*(k) \end{pmatrix} \tag{17}$$

For the sake of notation, the following variable (re)assignments are necessary:

$$\begin{aligned} H_x^*(k+1) &:= \bar{H}_{x_{k+1}} \\ H_{p1}^*(k+1) &:= \bar{H}_{p1_{k+1}} \\ H_{p2}^*(k) &:= \bar{H}_{p2_k} \\ H_{p3}^*(k) &:= \bar{H}_{p3_k} \\ y^*(k) &:= \bar{y}_k \end{aligned} \tag{18}$$

Next, the smoothing coefficients associated with $p2$ are then downdated.

$$T \begin{pmatrix} R_{p2}^{**}(k) & R_{p2,x}^{**}(k+1) & R_{p2,p1}^{**}(k+1) & R_{p2}^{**}(k+1) & R_{p2,p3}^{**}(k) & z_{p2}^{**}(k) \\ 0 & \bar{\bar{H}}_{x_{k+1}} & \bar{\bar{H}}_{p1_{k+1}} & \bar{\bar{H}}_{p2_{k+1}} & \bar{\bar{H}}_{p3_k} & \bar{\bar{y}}_k \end{pmatrix}$$
$$= \begin{pmatrix} R_{p2}^{*}(k) & R_{p2,x}^{*}(k+1) & R_{p2,p1}^{*}(k+1) & R_{p2}^{*}(k+1) & R_{p2,p3}^{*}(k) & z_{p2}^{*}(k) \\ H_{p2}^{*}(k) & H_x^{*}(k+1) & H_{p1}^{*}(k+1) & 0 & H_{p3}^{*}(k) & y^*(k) \end{pmatrix} \tag{19}$$

Again for the sake of notation, the following variable (re)assignments are necessary:

$$\begin{aligned} H_x^*(k+1) &:= \bar{\bar{H}}_{x_{k+1}} \\ H_{p1}^*(k+1) &:= \bar{\bar{H}}_{p1_{k+1}} \\ H_{p2}^*(k+1) &:= \bar{\bar{H}}_{p2_{k+1}} \\ H_{p3}^*(k) &:= \bar{\bar{H}}_{p3_k} \\ y^*(k) &:= \bar{\bar{y}}_k \end{aligned} \tag{20}$$

Finally, the smoothing coefficients associated with $p3$ are downdated.

$$T\begin{pmatrix} R^{**}_{p3}(k) & R^{**}_{p3,x}(k+1) & R^{**}_{p3,p1}(k+1) & R^{**}_{p3,p2}(k+1) & R^{**}_{p3}(k+1) & z^{**}_{p3}(k) \\ 0 & H_{x_{k+1}} & H_{p1_{k+1}} & H_{p2_{k+1}} & H_{p3_{k+1}} & y_{k+1} \end{pmatrix}$$
$$= \begin{pmatrix} R^{*}_{p3}(k) & R^{*}_{p3,x}(k+1) & R^{*}_{p3,p1}(k+1) & R^{*}_{p3,p2}(k+1) & R^{*}_{p3}(k+1) & z^{*}_{p3}(k) \\ H^{*}_{p3}(k) & H^{*}_{x}(k+1) & H^{*}_{p1}(k+1) & H^{*}_{p2}(k+1) & 0 & y^{*}(k) \end{pmatrix} \quad (21)$$

Note that all smoothing coefficients before time update $k+1$ remain unchanged. The $()^*$ terms are the downdated smoothing coefficients and replace the old smoothing coefficients. The ensuing data equation $y_{k+1} = H_{k+1}x_{x+1} + \nu_{k+1}$ is then used to continue this process to downdate the smoothing coefficients from time updates $k+2$ through N. In the end, the terminal data equation $y_N = H_N x_N + \nu_N$ is then downdated from the terminal SRIF array [3].

Constructing the Orthogonal Transformation

The orthogonal transformation in equations (17), (19), and (21) operates as

$$T\begin{pmatrix} R^{**}(k) & R^{**t} \\ 0 & H^t \end{pmatrix} = \begin{pmatrix} R^{*}(k) & R^{*t} \\ H^{*}(k) & H^{*t} \end{pmatrix} \quad (22)$$

where the $()(k)$ terms are scalars and the $()^t$ terms represent row vectors. As usual, let the orthogonally packed measurement be stored in the last column of the row vectors. To solve for the transformation T, multiply equation (22) by T^{-1} and take transposes.

$$\begin{pmatrix} R^{**}(k) & 0 \\ R^{**} & H \end{pmatrix} = \begin{pmatrix} R^{*}(k) & H^{*}(k) \\ R^{*} & H^{*} \end{pmatrix} T \quad (23)$$

Let $\tilde{A}^t = (A, \alpha)$, where $\alpha = \sqrt{1 - A^2}$ and A is a scalar to be determined. Note that the norm of $\tilde{A}$ is 1. The transformation T is then constructed as an elementary Givens reflection.

$$T = \begin{pmatrix} -\alpha & A \\ A & \alpha \end{pmatrix} \quad (24)$$

With T constructed as such, equation (23) is then post-multiplied by $\tilde{A}$.

$$\begin{pmatrix} R^{**}(k) & 0 \\ R^{**} & H \end{pmatrix}\begin{pmatrix} A \\ \alpha \end{pmatrix} = \begin{pmatrix} R^{*}(k) & H^{*}(k) \\ R^{*} & H^{*} \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (25)$$

The scalar A and vector H can now be determined.

$$A = \frac{H^{*}k}{R^{**}(k)} \quad (26)$$

$$H = \frac{H^{*} - AR^{**}}{\alpha} \quad (27)$$

If instead of downdating, it is desirable to update the filter solution with an additional measurement, an elementary Givens reflection would be applied to the right side of equation (22).

Efficiency

The discussed method was implemented on a VAX 8530. The filter test case required 30 minutes of CPU time to process 5800 data points in a 5.5 hour data arc of a typical GPS data processing scenario for the TOPEX/POSEIDON demonstration [6]; the filter estimated nearly 300 parameters. Downdating points from the resulting filter output required on average 10 seconds of CPU time per point. It proves that the linear combination of the data equation and the smoothing coefficients is relatively fast compared with downdating the terminal data equation from the terminal SRIF array.

Numerical Integrity

Stewart [5] has shown that downdating is stable in the presence of rounding errors. However, as he has also shown, if the spread of singular values is greater than half the computational precision, the precision of the smaller singular values may be lost. Fortunately, the downdating algorithm provides a way of detecting such ill-conditioned problems; the value of α (above and a similar quantity in the terminal downdating algorithm) is a reliable indication of trouble. Experience has shown that if α is less than 10^{-8}, on a machine with a 15 digits of precision, the results of particular estimates may be inaccurate. This situation occurs when all information is removed from an estimated parameter. However, in an operational environment there are generally sufficient measurements or *a priori* information to avoid such situations. In these cases, experience has shown that the estimates are accurate to better than 10 digits.

Applications

In spacecraft orbit determination, often critical measurements, such as Very Long Baseline Interferometry (VLBI) measurements, are not available until long after the usual radio-metric data (*e.g.* doppler, range) has been obtained. Using the discussed method, these critical measurements may be efficiently added after the usual data has been processed. When using optical data for orbit determination, camera pointing is often modeled as a white noise stochastic process. The discussed method permits the analyst to efficiently remove optical frames and replace them with others as desired. For large state systems with many stochastic parameters, such as GPS applications, outliers may be removed more efficiently and just as effectively by downdating rather than reprocessing the entire data set.

Conclusion

A new method to downdate a time-varying SRIF filter is presented. The method combines the algorithm of Gill [3] to downdate a matrix factorization and the decentralized SRIF algorithm of Bierman [4] to combine the results of independent time-varying SRIF filters. This method permits efficient removal of either measurements or *a priori* information. Additionally, updating a time-varying SRIF filter with either additional measurements or *a priori* information proceeds analogously. In both cases, a data equation is propagated by alternating between time mapping the equation and forming a particular linear combination with the SRIF's smoothing coefficients. The terminal data equation that results can then be merged with or removed from the terminal SRIF array. Smoothing then proceeds as usual [2]. For large state systems with many stochastic parameters, the discussed method is expected to be a critical component in the real-time reduction and analysis of large volumes of tracking data.

References

1. G. J. Bierman, Factorization Methods for Discrete Sequential Estimation, Academic Press, 1977.

2. G. J. Bierman, "Sequential Square Root Filtering and Smoothing of Discrete Linear Systems", Automatica, Vol. 10, pp. 147–158.

3. P. E. Gill, G. H.Golub, W. Murray, and M. A. Saunders, "Methods of Modifying Matrix Factorizatrions", Mathematics of Computation, Volume 28, Number 126, April 1974, pp. 505–535.

4. G. J. Bierman, "A Decentralized Square Root Information Filter/ Smoother", Proceedings of 24th Conference on Decision and Control, pp. 1902–1905, Ft. Lauderdale, FL., December 1985.

5. G. W. Stewart, "The Effects of Rounding Error on an Algorithm for Downdating a Cholesky Factorization", University of Maryland, Computer Science Technical Report TR-582, September 1977.

6. B. G. Williams, K. C. McColl, J. R. Guinn, "Navigation Accuracies for GPS Demonstration on TOPEX/POSEIDON", AIAA/AAS Astrodynamics Conference, Portland, OR., August 1990.

Acknowledgement

The work described in this paper was performed by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

N91-17091

THE NAVSTAR GPS MASTER CONTROL STATION'S KALMAN FILTER EXPERIENCE

Michael P. Scardera, Captain, USAF

2d Satellite Control Squadron
Falcon AFB, CO. 80912-5000

ABSTRACT

The Navstar Global Positioning System is a highly accurate, space based navigation system providing all weather, 24 hour-a-day service to both military and civilian users. The Navstar system provides a Gaussian position solution with four satellites, each providing its ephemeris and clock offset with respect to GPS time. Currently, GPS is building towards the full 24 satellite constellation.

The GPS Master Control Station (MCS) is charged with tracking each Navstar spacecraft and precisely defining the ephemeris and clock parameters for upload into the vehicle's navigation message. This paper briefly describes the Navstar system and the Kalman Filter estimation process used by the MCS to determine, predict, and quality control each satellite's ephemeris and clock states. Routine performance is shown. Kalman Filter reaction and response is discussed for anomalous clock behavior and trajectory perturbations.

Particular attention is given to MCS efforts to improve orbital adjust modeling. The satellite out of service time due to orbital maneuvering has been reduced in the past year from four days to under twelve hours. The planning, reference trajectory model, and Kalman Filter management improvements are explained.

Finally, this paper will summarize the future work to improve Kalman Filter and orbital adjust performance. Reccomendations will be made for future systems requiring precise ephemerides.

INTRODUCTION

The Navstar Global Positioning System (GPS) is a satellite based radionavigation system providing worldwide all weather coverage to both civilian and military users. The Navstar system promises a revolution in all activities requiring precise navigation or positioning. GPS signals will provide the precise positioning service to authorized users of 16 meters spherical error probable (SEP) and 100 meter circular error probable (CEP) to standard positioning service customers. The standard positioning service is subject to change according to United States national interests. GPS is made up of three segments. The space segment consists of the orbiting satellites and provides L-Band signals with modulated data to the world. The User segment represents the customers who receive and utilize the navigation data. Finally, the control segment comprises a system of L-Band monitor stations, S-band ground antennas, and a control center to monitor the satellites health and periodically upload new navigation parameters. GPS is currently the largest satellite constellation dedicated to a single purpose, and it's still growing to the planned 24 satellite constellation consisting of 21 satellites and three on-orbit spares. The Master Control Station (MCS), located at Falcon Air Force Base is responsible for the maintenance of the GPS satellites and their payload. The navigation payload of each satellite is regularly updated with orbital parameters, atomic frequency standard states, and almanac data for broadcast to the user community. This paper will briefly describe the Navstar system, the arrangement and organization of the MCS, and the Kalman Filter algorithms used by the MCS. The paper will go on to describe the ways in which the MCS manages those algorithms. Nominal performance as well as performance under anomalous conitions will be discussed. The improvement in orbital adjust modeling will be explained.

THE NAVSTAR SYSTEM

The GPS satellites operate at an altitude of approximately 20183 Km in a near circular orbit with a 12 hour semisynchronous orbit. Orbit inclination for the operational satellites is 55 degrees. Each satellite transmits two

L-Band signals. The L1 signal is a radio carrier frequency of 1575.42 MHz. The L1 signal is modulated by two pseudo random noise (PRN) codes known as the course acquisition (C/A) and precision (P) codes which are in phase quadrature with each other. The P code is a PRN stream of 10.23 Mbits per second compared to the C/A code's rate of 10.23 Kbits per second. The L2 carrier operates at a frequency of 1227.6 MHz and usually only contains the P code. Both the P and C/A PRN codes contain the modulated data representing the navigation message from the transmitting satellite. This navigation message is transferred at a rate of 50 bits per second. The navigation message contains the satellite's clock offsets from GPS time, the satellite's precise ephemeris, and course almanac information for the entire operational GPS constellation. Other parameters are included in the navigation message to indicate the health of the GPS satellite, give single frequency users ionospheric correction information, and provide Universal Time. Users can either receive the single C/A code, or use the C/A code to acquire the P codes on L1 and L2. All clocks in the Navstar System are synchronized to GPS time.

Once the user reads the navigation message from each of the visible tracking satellites, the user can synchronize his time to GPS time, calculate the apparent range to the satellite, and solve for position using the formula:

$$PR_i^I = \sqrt{(X_i - X_u)^2 + (Y_i - Y_u)^2 + (Z_i - Z_u)^2} + c\psi_u \qquad (1)$$

Where,

$(X_i, Y_i, Z_i)^T$ = Satellite Inertial Position

$(X_u, Y_u, Z_u)^T$ = User Inertial Position

PR_i^I = Ionospherically Corrected User Pseudorange Measurement From The i'th satellite

c = The Speed of Light

ψ_u = User Clock Offset from GPS Time

Since the satellite inertial positions at a given GPS time are derived from the navigation message and pseudorange measured from the PRN code time differences, the only unknowns are the user's inertial position and the user's clock actual offset from GPS time. Since their are four unknowns, four pseudorange measurements from four different satellites are required to find the user's inertial position and clock offset. Table I shows the specifications for GPS user accuracy compared with other navigation systems. Obviously for this system to work properly, knowledge of the satellite's ephemeris and clock offsets with respect to GPS time must be precise. The GPS Operational Control Segment (OCS) is charged with keeping the navigation message accurate and the spacecraft in good operating condition.

The Operational control segment consists of five L-Band Monitor Stations (MS), three S-Band Ground Antennas (with two additional antennas available on a part time basis), and the Master Control Station (MCS). The locations and coverage of these ground stations is shown in Figure I. The L-Band tracking data is transmitted from the monitor stations to the MCS. The MCS uses this information to generate new orbital elements and clock states which are transmitted through the ground antennas to the satellite. Information on the satellite's status is also gained from the ground antanna S-Band telemetry. This system is pictured in Figure II. The GPS MCS is the hub of all the activity in the command and control of the Navstar system.

THE MASTER CONTROL STATION MANNING

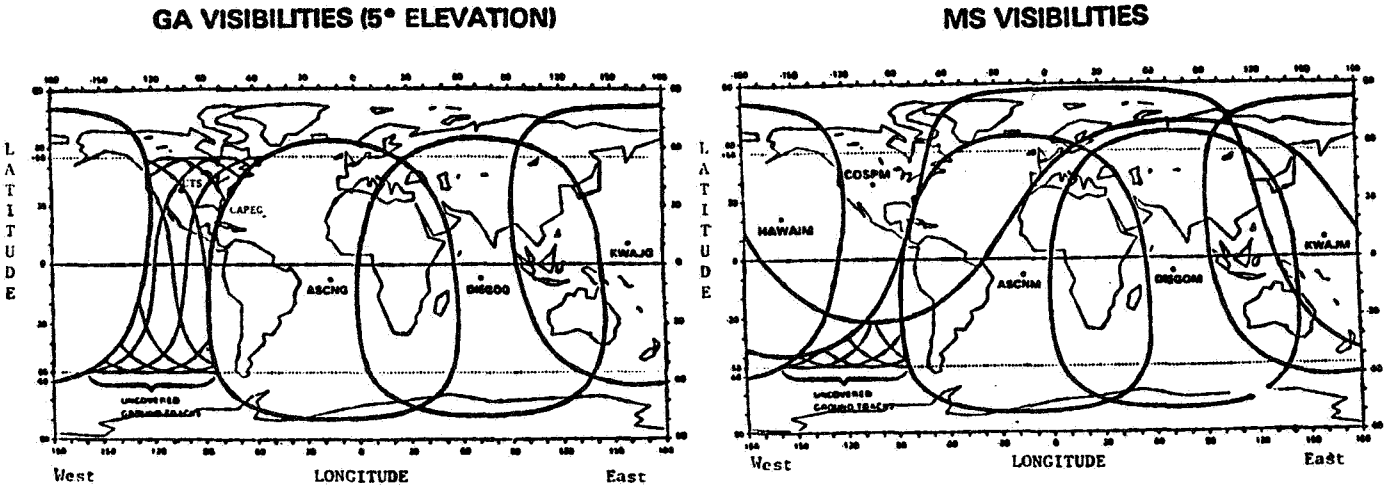
The GPS Master Control Station's operations center is staffed with seven members of Air Force Space Command's Second Satellilite Control Squadron. This group of seven people comprise the basic space operations crew. There are five crews, each working for a period of six days followed by four idle days. The off days provide for crew rest, as well as training and standby days. These operations crews maintain the Navstar constellation's state of health and navigation performance around the clock, every day of the year.

There are currently six positions in the operations center filled by the seven crew members. The Ground Controller is responsible for the communication links between the MCS and each of the four ground antennas and five monitor stations. The Satellite Operations Officer is the position tasked with making contact with a given satellite from the MCS through a ground antenna to monitor a given vehicle's state of health and transmit any necessary commands. The Satellite Engineering Officer is responsible for ensuring that a spacecraft is in good operating condition, all commands

TABLE I: Comparison Of GPS Specifications To Other Navigation Systems. Accuracies Are Given For The Precise Positioning Service.

System	Position Accuracy (m)	Velocity Accuracy (m/sec)	Range of Operation	Comments
GPS	16 (SEP) 3-D	≥ 0.1 (RMS per axis) (Note 3)	Worldwide	Operational worldwide with 24-hour all-weather coverage. Specified position accuracy available to authorized users.
Loran-C (Note 1)	180 (CEP)	No velocity data	U.S. Coast, Continental U.S., Selected Overseas areas	Operational with localized coverage. Limited by skywave interference.
Omega (Note 1)	2,200 (CEP)	No velocity data	Worldwide	Operational worldwide with 24-hour coverage. System is subject to VLF propagation anomalies.
Std INS (Note 2)	1,500 max after 1st hour (CEP)	0.8 after 2 hrs (RMS per axis)	Worldwide	Operational worldwide with 24-hour all-weather coverage. Degraded performance in polar areas.
TACAN (Note 1)	400 (CEP)	No velocity data	Line of sight (present air routes)	Position accuracy is degraded mainly because of azimuth uncertainty which is typically on the order of ±1.0 degree.
Transit (Note 1)	200 (CEP)	No velocity data	Worldwide	The interval between position fixes is about 90 minutes. For use in slow moving vehicles. Better position fix accuracy is available with dual frequency measurements.

NOTE:
1. Federal Radionavigation Plan, December 1984
2. SNU-84-1 Specification for USAF Standard Form Fit and Function (F^3) Medium Accuracy Inertial Navigation Set/Unit, October 1984
3. Dependent on integration concept and platform dynamics



CTS - Colorado Springs Tracking Station
CAPEG - Cape Canaveral Ground Antenna
ASCNG - Ascension Island Ground Antenna
DIEGOG - Diego Garcia Ground Antenna
KWAJG - Kwajalein Atoll Ground Antenna

COSPM - Colorado Springs Monitor Station
ASCNM - Ascension Island Monitor Station
DIEGOM - Diego Garcia Monitor Station
KWAJM - Kwajalein Atoll Monitor Station
HAWAIM - Hawaii Monitor Station

NOTE: CTS and CAPEG give coverage over the Continental United States on a part-time basis.

FIGURE I: The GPS Ground Antenna and Monitor Station Locations and Their Respective Visibilities

sent to the spacecraft are valid, and to respond to any anomalous behavior.

The Satellite Analysis Officer monitors the L-Band tracking data from each of the monitor stations and verifies the navigation performance of the GPS constellation. The Deputy Commander takes care of all the bookkeeping and upchannel reporting of the crew's activities required in an Air Force operation. Finally, there's the Crew Commander who manages the entire crew and performs real-time rescheduling of the operations center's activities.

At the present time, there are two Satellite Operations Officers while every other position is filled by one crew member. The Ground Controller is at this time the only enlisted person on crew. Every other position is staffed by junior officers. The next ten years will probably see the Satellite Operations Officer positions grow to three slots and transition to enlisted personnel. The Deputy Commander position is also likely to be taken over by an enlisted person.

In addition to the operations crews, the Second Satellite Control Squadron supports these crews with several other sections. The Space Operations section directly supports the operations crews. The Scheduling Section produces a daily schedule of events which the crew follows. The Training and Standardization/Evaluation sections guarantee crew proficiency in their jobs. There's also an engineering support group which is broken down into sections responsible for the analysis of the spacecraft bus, navigation payload, ground system, and system database. Finally there is a section dedicated to interfacing with various users and a Command Section. All in all, there are approximately 100 people in the Second Satellite Control Squadron responsible for the operation of the GPS constellation.

It's the Satellite Analysis Officer crew position and the navigation payload section of the engineering analysis group which work with the contents of this paper. This group of people manages the Kalman Filter algorithms and results, guaranteeing the end product to the GPS positioning/timing users is usable, accurate, and within all published specifications.

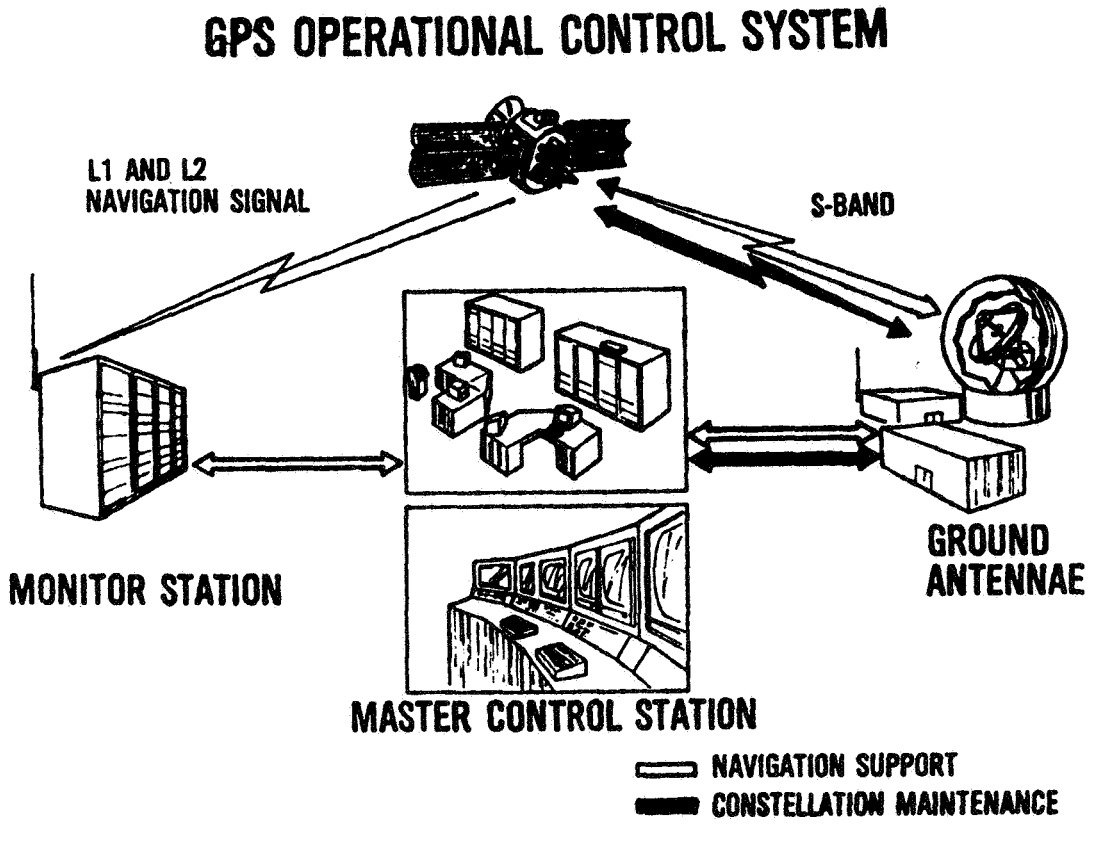


FIGURE II: The GPS Operational Control Segment. L-Band signals transmitted by each satellite are tracked by the Monitor Stations. Navigation Uploads, Telemetry, and Commands are routed through the Ground Antennas. The Master Control Station is the central processing location.

THE MCS KALMAN FILTER PROCESS

The two fundamental measurements made by the Navstar system are pseudorange (PR) and accumulated delta range (ADR). Pseudorange measures the distance between a satellite and a monitor station on the Earth by accounting for the total time between transmission and reception. This measurement is a "pseudo" range rather than a "true" range as a result of delays encountered on the signal's path to the Earth's surface. The equation used for pseudorange by the MCS is:

$$PR = c\left(t_{R_{MS}} - t_{T_{SV}}\right) \qquad (2)$$

Where,

- PR = Pseudorange Measurement
- c = Speed of Light
- $t_{R_{MS}}$ = Time Signal Received by Monitor Station
- $t_{T_{SV}}$ = Time Signal Transmitted By Satellite

The transmission time is represented by an integer number of z-counts where one z-count is equal to 1.5 seconds. The satellite and monitor station times are referenced to GPS time for synchronization.

Accumulated delta range (ADR) is a measurement of total phase accumulated between the transmision and reception of the carrier signal. Accumulated delta range is given by the expression:

$$ADR = \frac{c}{f}\left(\phi_{R_{MS}} - \phi_{T_{SV}}\right) \qquad (3)$$

Where,

- ADR = Accumulated Delta Range
- f = Carrier Frequency
- $\phi_{R_{MS}}$ = Monitor Station Phase Measurement
- $\phi_{T_{SV}}$ = Transmitted Satellite Phase

The only use for ADR's within the GPS control segment is to aid in the smoothing of pseudorange data. ADR's are available primarily for the GPS user community to calculate relative velocities.

The Master Control Station continuously tracks and collects pseeudorange data from each satellite visible to the monitor stations. A pseudorange measurement is made every z-count on both the L1 and L2 signals received. Data processing occurs at fifteen minute intervals. As shown in Figure III, the MCS begins to store data at each fifteen minute point. Over each fifteen minute interval the satellite transmissions are checked for correct parity and C/A to P code Handover Word. Each measurement must also be within limits for signal/noise ratio, code and phase slips, code and carrier lock, and first and second difference tests. At the end of the fifteen minute interval there is a five minute waiting period. This waiting period is used to retrieve any data which may reside in a monitor station's data buffer due to short communications outages. After the five minute wait, the data which passed all editing critertia for raw pseudorange measurements during the fifteen minute interval are now ready for processing.

Each L1 and L2 measurement passing the editing criteria is now used to determine the ionospherically corrected pseudorange. The layers of the Earth's atmosphere approximately between 60 and 640 Km altitude constitute the ionosphere. The ionosphere can vary greatly over the surface of the Earth for such reasons as solar activity, the effects of man, and the variation between day and night. The ionosphere affects GPS by bending the transient carrier signals. Fortunately, the L-Band signal frequencies used bt the Navstar system exhibit the property that the index of refraction is proportional to the inverse square of the frequency. Since the two different carrier frequencies, L1 and L2, measure the same distance between the satellite and the monitor station, the proportionality constant can be determined independent of the actual ionospheric conditions or the elevation of the satellites. The proporionality constant is given by:

$$K_T = \frac{PR_1 - PR_2}{\left(\frac{1}{2f_1^2} - \frac{1}{2f_2^2}\right)} \quad (4)$$

where

PR_1 = Direct L1 pseudorange measurement
PR_2 = Direct L2 pseudorange measurement
f_1 = L1 carrier frequency = 1575.42 MHz
f_2 = L2 carrier frequency = 1227.60 MHz

Using the L1 measurement, the ionospheric time delay now becomes:

$$\Delta t_i = \frac{K_T}{2cf_1^2} \quad (5)$$

So each pair of L1 and L2 pseudorange measurements received at every z-count is used to transform the raw pseudoranges to an ionospherically corrected pseudorange with the equation:

$$PR_I = c\left(t_{R_{M.S}} - t_{T_{SV}} - \Delta t_i\right) \quad (6)$$

The ionospherically corrected pseudorange is the only calculation which occurs to the signal before filter processing. Other phenomenon which affect the signal, such as troposheric delay, relativity, and the free space delay are accounted for within the filtering system.

The ionospherically corrected pseudoranges are now smoothed over the 15 minute collection interval to reduce the measurement noise before Kalman Filter processing. A polynomial fit is made to the ionospherically corrected pseudoranges for each satellite-monitor station measurement pair. The smoothed pseudorange is determined directly from the polynomial and is set at the beginning of the fifteen minute interval.

The vector containing smoothed pseudoranges for each satellite-monitor station pair is defined as the observation vector. This observation vector will be processed by the MCS Kalman Filter to produce a state vector for each satellite and monitor station as shown in Figure IV. The state vector for each satellite includes its inertial position and velocity, a scaling parameter (K1) and an acceleration parameter (K2) for a solar pressure force model, and clock bias, drift, and drift rate. The monitor station state vector includes clock bias and drift, and tropospheric height.

To reduce the magnitude of the estimation problem, the MCS uses an independent partitioning scheme. Each partition is made of state vectors for up to six spacecraft and includes all monitor station states. This scheme reduces the processing load of the MCS, allows for isolation of satellites with poor performance, and gives the operator insight into the ground stations' states between partitions.

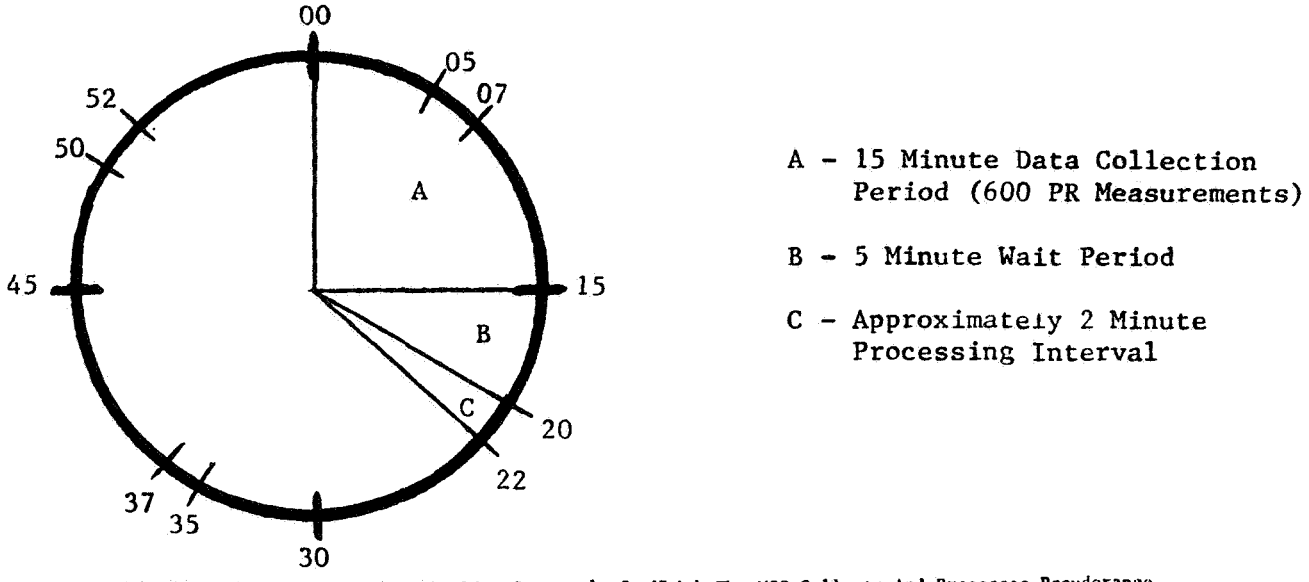


FIGURE III: Clock Representing The Time Intervals In Which The MCS Collects And Processes Pseudorange Data. The Cycle Shown Above Starts Every 15 Minutes.

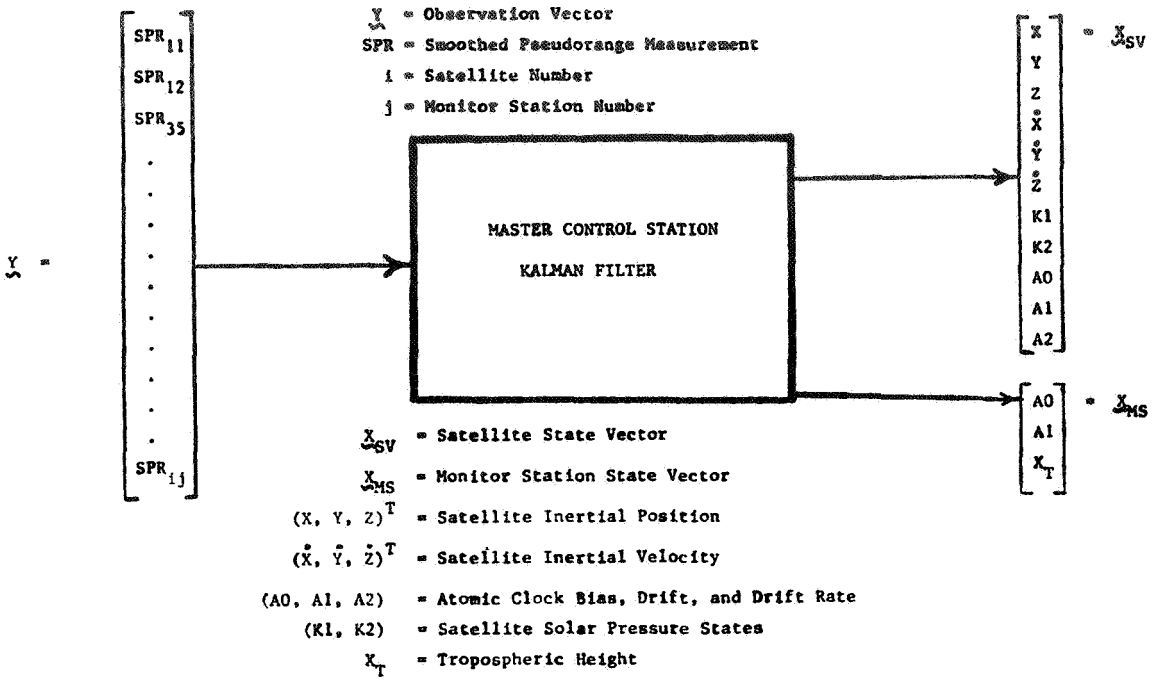


FIGURE IV: The MCS Kalman Filter Uses Smoothed Pseudoranges To Determine The Satellite And Monitor Station States Shown Above.

The purpose of the MCS Kalman Filter is to estimate the current states in a partition given the available measurements in the observation vector. In the MCS manifestation of the Kalman Filter, the current state process is linearized with respect to a reference system and then related to an epoch state residual process. The observation process is linearized about an apriori estimate of the state vector and likewise related to to the epoch state residual process. The formulation of the MCS Kalman Filter is described in the following paragraphs to give some context to the subsequent discussion of GPS Kalman Filter operations.

In the MCS implementation of the Kalman Filter, a predicted reference trajectory and a set of reference clock states exist. A simple relation is drawn between each state and the reference states at a specified time. For example, the current state residual is defined as the difference between the current state and the reference state at a current time:

$$\delta \underset{\sim}{x}_t = \underset{\sim}{x}_t - \underset{\sim}{x}_{ref(t)} \tag{7}$$

The epoch state residual is defined as the difference between the epoch state and the reference epoch state at the epoch time:

$$\delta \underset{\sim}{x}_0 = \underset{\sim}{x}_0 - \underset{\sim}{x}_{ref(0)} \tag{8}$$

The epoch states are mapped to the current states through a nonlinear state propagation model (F[...]) such that the reference state at current time can be found from the reference epoch state at epoch time:

$$\underset{\sim}{x}_{ref} = F[t, t_0, \underset{\sim}{x}_{ref(0)}] \tag{9}$$

The current state can be found from the epoch state and their associated times with:

$$\underset{\sim}{x}_t = F[t, t_0, \underset{\sim}{x}_0] \tag{10}$$

The relationship between the state residuals and their respective mapping functions is pictured in Figure V. Note that the reference state process shown in Equation 9 is a completely deterministic function.

For every state estimating partition, the state process is modelled in discrete time and given as:

$$\underset{\sim}{X}(T_{k+1}) = F[T_{k+1}, T_k, \underset{\sim}{X}(T_k)] + \underset{\sim}{w}(T_k) \tag{11}$$

Where,

$\underset{\sim}{X}$ = State Vector

$\underset{\sim}{w}$ = State Process Noise

T_k = Time of Current Data Interval

T_{k+1} = Time of Next Data Interval

The state and state process noise vectors are dimensioned by the number of states in a partition. This current state process is linearized by substituting Equation 7, at the discrete time, into Equation 11 and then performing a first degree Taylor Series expansion on the current state residual term. With the substitution of Equation 9 at discrete time into the expansion, the current state residual is found to be:

$$\delta_{\underset{\sim}{X}(T_{k+1})} = \left.\frac{\partial F[T_{k+1}, T_k, \lambda]}{\partial \lambda}\right|_{\lambda = \underset{\sim}{X}_{ref}} \underset{\sim}{X}(T_k) + \underset{\sim}{w}(T_k) \tag{12}$$

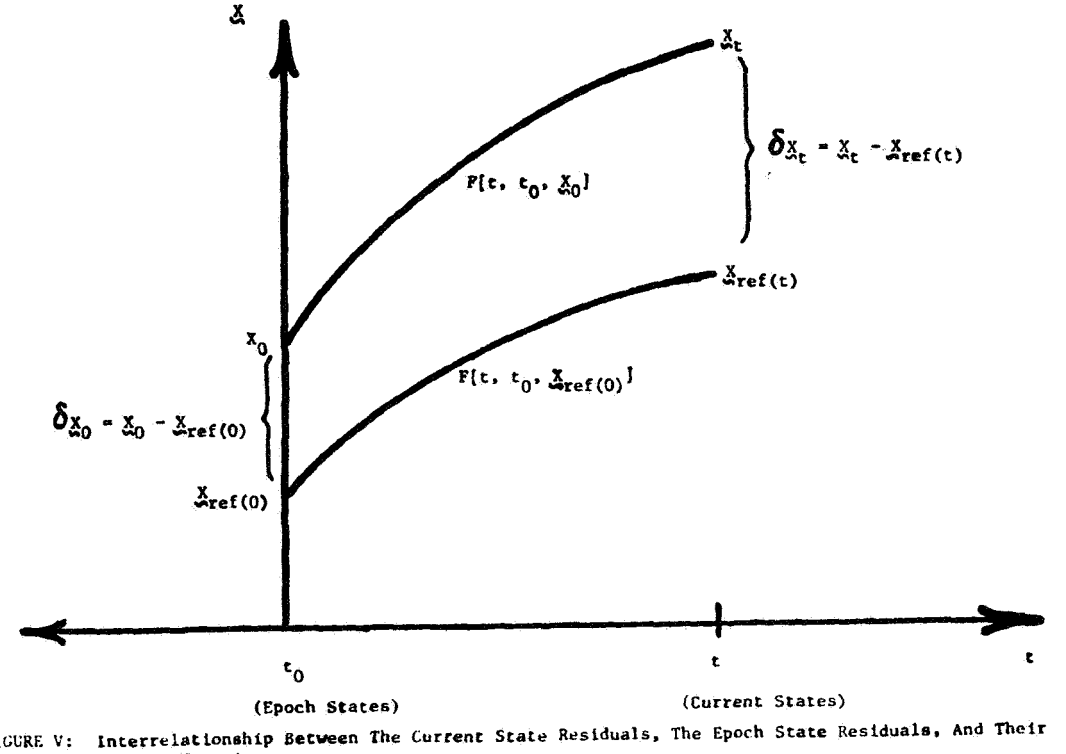


FIGURE V: Interrelationship Between The Current State Residuals, The Epoch State Residuals, And Their Mapping Functions.

By following the same logic, the current state residual is related to the epoch state residual by linearizing the mapping function with respect to the reference epoch state. Substituting Equation 8 into 10, performing a Taylor Series expansion on the epoch state residual term, and substituting Equation 7 results with the equation:

$$\delta_{\underset{\sim}{X}(T_k)} = \left.\frac{F[T_k, T_0, \lambda]}{\partial\lambda}\right|_{\substack{\underset{\sim}{X}_0(T_k) \\ \lambda=\underset{\sim}{X}_0}} \tag{13}$$

If the two discrete time forms of Equation 13 are substituted back into Equation 12 and algebraically manipulated, the final result is the epoch state residual process used by the MCS:

$$\delta_{\underset{\sim}{X}_0(T_{k+1})} = \delta_{\underset{\sim}{X}_0(T_k)} + \left[\frac{\partial_{\underset{\sim}{X}_{ref}(T_{k+1})}}{\partial_{\underset{\sim}{X}_0}}\right]^{-1} \underset{\sim}{w}(T_k) \tag{14}$$

The state process noise is assumed to be a white gaussian process with zero mean such that the expected values of the state process noise have the form:

$$E[\underset{\sim}{w}(T_k)] = 0 \qquad \text{and} \qquad E[\underset{\sim}{w}(T_k)\underset{\sim}{w}(T_L)] = Q(T_k)\,\delta_{kL} \tag{15}$$

δ_{kL} = Kronecker Delta
T_L = Any Other Data Interval

where Q represents the covariance matrix for each state. The covariances have a constant value defined in the MCS database, and are added at each Kalman Filter data interval.

The observation process is modelled in discrete time with the nonlinear equation:

$$\underset{\sim}{y}_t = h(\underset{\sim}{x}_t) + \underset{\sim}{v}_t \tag{16}$$

The observation vector and process noise are dimensioned by the total number of observations made by a partition. This equation relates the current state vector to the smoothed pseudoranges with the state transition matrix, h[] with the addition of measurement process noise. This observation process is linearized about an apriori state estimate and then expressed in terms of the epoch state residual. The resulting linearized equation is:

$$\underset{\sim}{y}_t = h(\underset{\sim}{\tilde{x}}_t) + \left[H(\underset{\sim}{\tilde{x}}_t)\right](\delta_{\underset{\sim}{x}_0} - \delta_{\underset{\sim}{\tilde{x}}_0}) + \underset{\sim}{v}_t \tag{17}$$

where

$\left[\frac{\partial h}{\partial \underset{\sim}{\tilde{x}}_t}\right]$ = The matrix containing the partial derivatives of the pseudorange measurement function with respect to each apriori state

$[\phi] = \frac{\partial_{\underset{\sim}{x}_{ref}}}{\partial_{\underset{\sim}{x}_{ref(0)}}}$ = Transition matrix

$$\left[H(\underset{\sim}{\tilde{x}}_t)\right] = \left[\frac{\partial h}{\partial \underset{\sim}{\tilde{x}}_t}\right][\phi]$$

As shown in Equation 2 the pseudorange measures the time delay between the time a monitor station received the given signal. This time difference can be broken down further to its components such that the observation model takes the form:

$$h(\underset{\sim}{X}(T_k)) = c\left[\Delta T_T(T_k) + \Delta T_{FS}(T_k, T_R) + \psi_{MS}(T_R) - \psi_{SV}(T_k) - \Delta T_r(T_k)\right] \quad (18)$$

Where,

ΔT_T = Tropospheric Delay
ΔT_{FS} = Free Space Delay
ΔT_r = Relativistic Delay
ψ_{MS} = Monitor Station Clock Offset
ψ_{SV} = Satellite Clock Offset
T_k = Signal Transmit Time = Data Interval Time
T_R = Signal Reception Time

The last three terms inside the brackets give the timing offsets between the monitor station and satellite clocks. The first two terms represent the propogation delay between the satellite and monitor station. Remember that the ionospheric delay has already been accounted for in the dual frequency corrections made before smoothing. The clock offsets are simply the time bias between either the satellite or monitor station clock and GPS time. The free space, tropospheric and relativistic delays will briefly be described.

The free space delay is experienced by a satellite's transmitted signal due to the physical propogation path length in vacuum. The free space delay is given as:

$$\Delta T_{FS}(T_k, T_R) = \frac{\left|R_{SV}(T_k) - R_{MS}(T_R)\right|}{c} \quad (19)$$

R_{SV} = Satellite Inertial Range

The monitor station reception time and the position vector for the monitor station phase center are found through an iterative process given the known transmit time.

The troposphere is the portion of the atmosphere which fills the volume between the Earth's surface to an altitude of about 12 kilometers. The major effect of the troposphere is to give an apparent increase in the signal path length. To help in the evaluation of tropospheric delay, three atmospheric parameters are measured at the monitor station recieving the signal. These parameters are atmospheric pressure, temperature, and dew point. Since different locations around the world can have widely varying conditions of the troposphere, the tropospheric height is modeled as a state variable in the estimation process of the Kalman filter. The tropospheric height is the actual altitude of the troposphere at the monitor station. The tropospheric delay is found with the equation:

$$\Delta T_T = \frac{1}{c}\left\{\frac{(0.02312)P_s\left[T_k - 4.11 + \frac{5(r_{mm} - r_{ma})}{148.98}\right]}{T_k\sqrt{1 - \left[\frac{r_{ma}\cos(\theta)}{r_{mm} + (1-l_c)(r_d - r_{mm})}\right]^2}} + \frac{(0.0746)E_0 X_T\left[1 + \frac{5(r_{mm} - r_{ma})}{X_T}\right]}{T_k^2\sqrt{1 - \left[\frac{r_{ma}\cos(\theta)}{r_{mm} + (1-l_c)X_T}\right]^2}}\right\} \quad (20)$$

P_s = Monitor station barometric pressure converted to Kilopascals
T_k = Monitor station measured temperature converted to the Kelvin scale
r_{mm} = The radial distance from the Earth's center to the meteorological sensors.
r_{ma} = The radial distance from the Earth's center to the monitor station antenna
r_d = Tropospheric dry radius
θ = Satellite elevation
X_T = Tropospheric height
E_0 = Estimate of the partial water vapor pressure

$$= (6.11)(10)^{\left[\frac{(7.5)T_d}{T_d + 237.3}\right]}$$

T_d = Dew point temperature
l_c = Assumed integration constant
= 0.95 when $\theta > 5.0°$

A satellite in orbit will experience the effect of special relativity due to its speed with respect to an observer on the Earth. To the ground observer, the satellite clock will appear to be behind. The same satellite will also feel the effect of general relativity because of the orbit's lower gravitational potential. To the same ground observer, the satellite clock will appear to be ahead. The effect of general relativity is greater so to compensate, the satellite's clock is built to run at 10.2999999545 MHz rather than the 10.23 MHz of the monitor station clocks. With the satellite in orbit the two clocks will appear to be synchronized.

Setting the satellite clock behind the monitor station clock would completely account for the relativistic effects if the satellite's orbit were perfectly circular. Since the satellite is never in a perfectly circular orbit, a periodic relativistic variance needs to be included with the equation:

$$\Delta t_r = \frac{-2e\sqrt{\mu a}}{c^2} \sin(E) \tag{21}$$

where

- e = Orbit eccentricity
- μ = Universal gravitational parameter.
- a = Semi-major axis of orbit
- c = The speed of light
- E = Eccentric Anomaly From Kepler's Equation

At the end of each data collection interval, a Kalman Filter solution is used to find the current state vector with the linearized whitened epoch state process (Equation 14) and observation process (Equation 17). A time update occurs first. In the time update, the aposteriori state estimates from the previous data interval become the apriori state estimates for the current data interval. Process noise is added to the aposteriori covariance matrix to define the apriori covariance matrix for the current data interval. The equations dictating the time update are:

$$\delta\bar{\underline{x}}_0(T_{k+1}) = \delta\hat{\underline{x}}_0(T_k)$$

$$\bar{\underline{P}}_0(T_{k+1}) = \hat{\underline{P}}_0(T_k) + \left[B(T_k)\right] Q(T_k) \left[B(T_k)\right]^T \tag{22}$$

Where,

- $\delta\bar{\underline{x}}_0(T_k)$ = Apriori estimate of epoch state residuals given past observations.
- $\delta\hat{\underline{x}}_0(T_k)$ = Aposteriori estimate of epoch state residuals given past and current observations.
- $\underline{P}_0(T_k)$ = Covariances of the epoch state residuals.
- $Q(T_k)$ = Process Noise
- $\left[B(T_k)\right] = \left[\Phi(T_{k+1},T_0)\right]^{-1}\left[G(T_k)\right]$ where: $\left[G(T_k)\right]$ is a wieghting matrix.

After the time update occurs, a Kalman Gain is calculated using the apriori states and expected covariance matrix:

$$\underline{K} = \bar{\underline{P}}_0 \left[H(\bar{\underline{x}}_t)\right]^T \left[\left[H(\bar{\underline{x}}_t)\right] \bar{\underline{P}}_0 \left[H(\bar{\underline{x}}_t)\right]^T + \left[I\right]\right]^{-1} \tag{23}$$

The aposteriori estimate of the epoch state residuals for the current data interval is now found from the equation:

$$\delta\hat{\underline{x}}_0 = \delta\bar{\underline{x}}_0 + \underline{K}\left[\underline{y}_t - h(\bar{\underline{x}}_t)\right] \tag{24}$$

$$\hat{\underline{P}}_0 = \bar{\underline{P}}_0 - \underline{K}\left[H(\bar{\underline{x}}_t)\right] \bar{\underline{P}}_0$$

Using Equations 8 and 10, the current states can be found from the epoch state residuals.

With the linearized epoch state process, if the current state deviation from the reference states becomes too large, the quality of the MCS Kalman Filter solution can degrade. To avoid this problem, the MCS updates reference trajectories every four weeks and whenever linearity limits are exceeded.

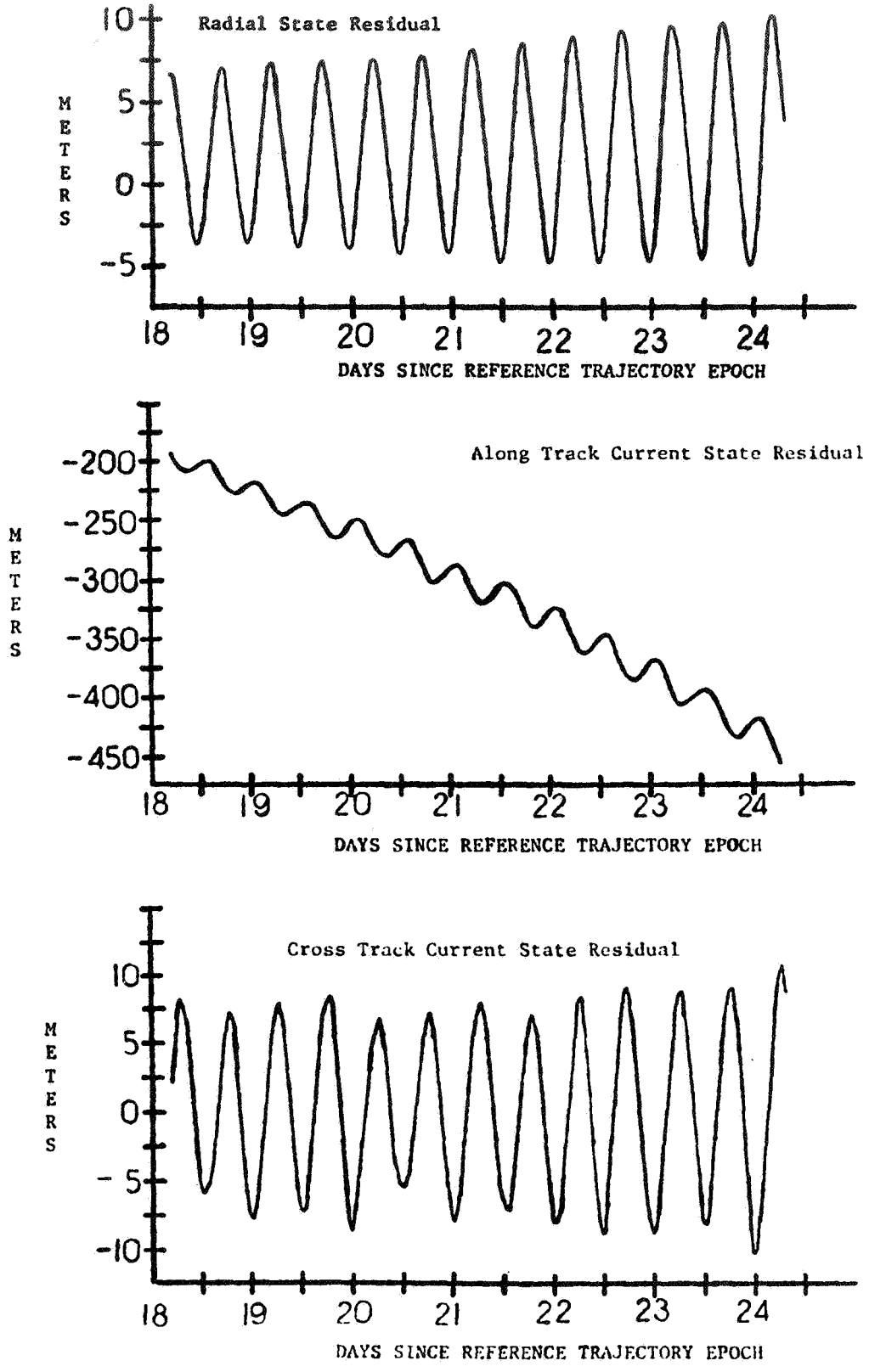


FIGURE VI-A: Typical Ephemeris State Performance. Inertial Position Offsets Have Been Transformed Into Body Radial, Along-Track, and Cross Track Deviations. The Offsets at Reference Epoch are Zero. On the 28th Day, The Current Reference Trajectory is Replaced with a New Reference Trajectory Based on the Current State Estimates at the New Epoch.

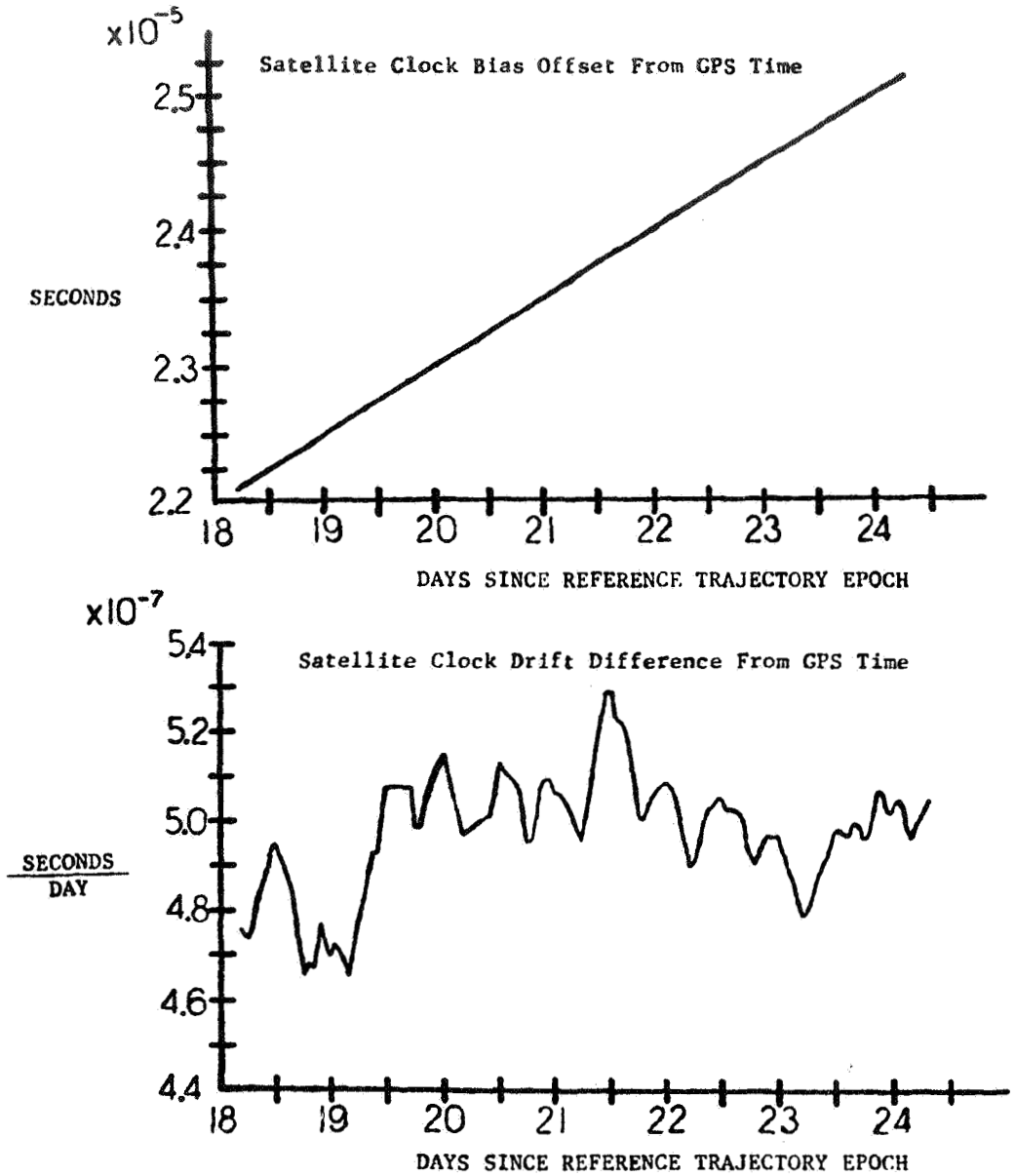


FIGURE VI-B: Satellite Cesium Atomic Clock State Bias and Drift Offsets From GPS Time. These Plots are also Typical of Normal Performance. The Reference States for Clocks in the GPS System are Determined from GPS Time. GPS Time is Derived from Either a Single Designated "Master Clock" or a Composition of All the Clocks in the GPS System. Note That There isn't any Apparent Influence on Clock States From Ephemeris States or Vice Versa.

MCS KALMAN FILTER OPERATIONS

Figure VI-A indicates the typical performance of current state residuals over a period of one week, approximately three weeks after a reference epoch. Aging of the reference trajectory states is evident over this period. The aging occurs because only major orbital perturbations are modeled by the Master Control Station. Figure VI-B shows typical cesium atomic clock performance compared to GPS time over the same period as the ephemeris states.

Epoch State Residuals, Current State Residuals, Epoch States, and Current States, as well as plots over a twenty-four hour period are all available for viewing to the Satellite Analysis Officer (SAO) at the end of each data interval. Though always available to the SAO, this information is usually only accessed when additional data on a given situation is needed.

The difference between the ionospherically corrected pseudorange and the pseudorange calculated from the apriori state estimates is called the pseudorange residual. The SAO observes the pseudorange residuals at the end of each data processing interval. The value of the pseudorange residual is usually below three meters. Values much above six meters are usually the first, and best indicator that there is a monitor station problem, satellite problem, or possibly a poor Kalman Filter model. If the pseudorange residual is above a dynamic limit based on the covariance matrix, the measurement obtained during the data interval is not used in the update of the Kalman Filter model.

When there are problems with the state estimates of a satellite or monitor station, the SAO has relatively few options to remedy the given situation. One option is to allow the process noise to add over time such that the measurement falls within limits some time in the future. This option generally isn't acceptable due to either the total time involved or the size of the rejected measurement. If the error can be attributed to a satellite or monitor station clock state, the SAO is capable of specifically modifying the clock state based on the observed pseudorange residuals. Another option left to the SAO is to change the value of the covariance matrix [P] in Equation 24. This action has been inappropriately called the "Q-Bump". It's possible to modify either monitor station, satellite, clock, ephemeris, or solar pressure state covariances. Any subset of these state covariances can be "Q-Bumped" at the same time. The value of the "Q-Bumped" covariances is prescribed in the MCS database and can vary with different situations. It's also possible to reinitialize the estimation problem, but this action is usually a drastic step. In some situations, various MCS database parameters can be changed which affect the behavior of the Kalman Filter algorithms.

There are two other tools used by the SAO. The first tool is the capability to mask the Kalman Filter. Masking prevents the update of the Kalman Filter state estimates with new data, thereby forcing only a time update to occur each data interval. The second tool is the capability, on a non-interference basis, to reprocess the Kalman Filter over a twenty four hour period with varying conditions. This reprocessing capability is quite powerful and chiefly responsible for the resolution of many problems.

The navigation data providing the GPS user community the precise ephemerides needed for accurate positioning are based on the Kalman Filter State estimates. When a navigation data upload is generated for transmission to a Navstar spacecraft, the satellite's ephemeris is predicted by using the aposteriori state estimates to differentially correct the reference trajectory. The clock states are predicted merely by propagating the aposteriori clock state estimates into the future. These ephemeris and clock state predictions are transmitted to the spacecraft with S-Band and, through the satellite, transmitted to the user community L-Band at the appropriate time.

The Master Control Station's insight into the quality of the upload prediction is performed with two statistical measures using the Kalman Filter state estimates. The first measure is the Estimated Range Deviation (ERD). The ERD is the difference between the aposteriori estimate of pseudorange and the pseudorange derived from the upload prediction for a particular vehicle. The ERD is calculated for a mathematically determined set of locations around the world and each GPS satellite visible to those locations. Normally ERD's grow over time indicating a slow degradation of the upload prediction. Figure VII shows this aging process. The need to upload each Navstar spacecraft with new navigation data on a daily basis is apparent in Figure VII to maintain the specified navigation performance. ERD's usually vary a little more than shown in Figure VII due to such aspects as monitor station visibility, timing of the upload creation, and clock movement. The second measure of performance is the Observed Range Deviation (ORD). The ORD is the difference the ionospherically corrected smoothed pseudorange and the aposteriori pseudorange determined by the navigation upload state prediction. An ORD is produced for each satellite visible to a monitor station. The ORD can be a good indicator of navigation performance. Should the ERD's and ORD's rise above a tolerance, set by the MCS as 10 Meters, a contingency navigation data upload will be accomplished to maintain user specified accuracy. Equations describing the Pseudorange residual, ERD, and ORD are given below:

$$PRR = SPR_{ij} - PR(\tilde{\underset{\sim}{X}})_{ij} = \text{Pseudorange Residual}$$

$$ERD = PR(\hat{\underset{\sim}{X}})_{ij} - PR(\underset{\sim}{X}_{\text{upload prediction}})_{ij} = \text{Estimated Range Deviation} \qquad (25$$

$$ORD = SPR_{ij} - PR(\hat{\underset{\sim}{X}}_{\text{upload prediction}})_{ij} = \text{Observed Range Deviation}$$

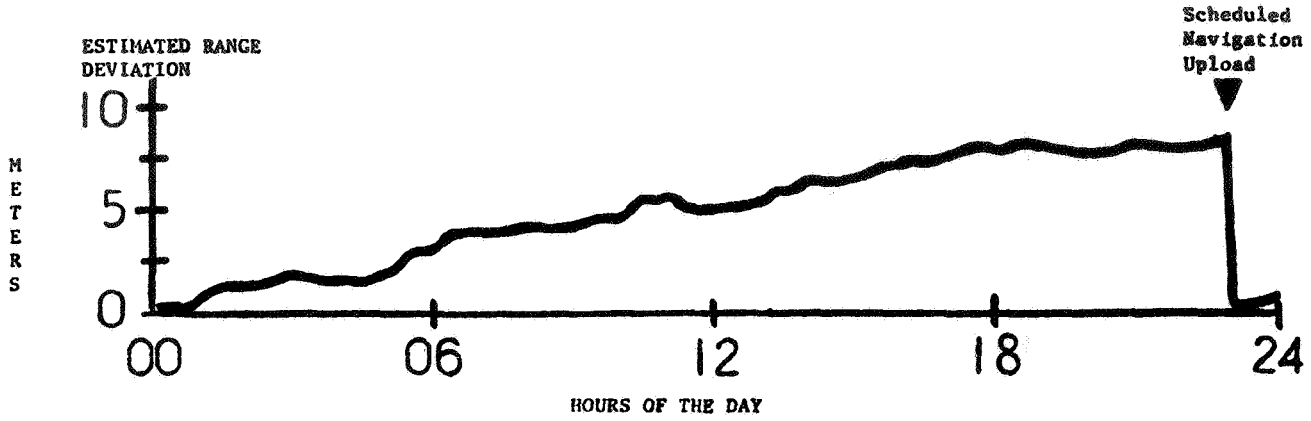


FIGURE VII: Typical Upload Prediction Performance as Reported by ERD's. Performance is Measured Against Range Based on Current State Estimates. Navigation Uploads Contain Ephemeris and Clock State Predictions Based on the Most Current State Estimates.

RESPONSE TO ANOMALOUS EVENTS

In January 1989, a Cesium atomic clock aboard Navstar vehicle number 11 failed. This failure necessitated a clock switch on board the spacecraft to a Rubidium atomic clock. A "Q-Bump" on the satellite's clock states was performed after the hardware failed. The Kalman Filter estimates had stabilized to new clock states in approximately six hours. This performance is shown in Figure VIII. The transient during this twelve hour period is readily observed. One month prior to the clock switch, the spacecraft bus section of the 2SCS engineering support group, predicted the cesium clock failure. As the time of the predicted failure approached the bad clock's performance was deteriorating. This deterioration is barely discernible in the clock drift plot shown in Figure VIII. Even with the deteriorated clock, performance of the satellite's navigation payload was maintained through a series of Kalman Filter clock state modifications and contingency navigation data uploads. On the fifth day of the year, maintaining performance through this means became impractical, so the satellite was set "Unhealthy" to the user community and the atomic clock was switched. At this time, the Kalman Filter was masked to prevent update of the satellite's clock and ephemeris states and the new Rubidium clock was allowed to "warm up". On the sixth day of the year, the clock state "Q-Bump" was performed. Clock performance was observed over the next several days and Navstar 11 was set back "Healthy". on the tenth day of the month.

Figure IX illustrates the impact of a satellite trajectory perturbance. In this case, Navstar 10 came out of a solar eclipse with an attitude error. This error was sensed by the spacecraft's Attitude and Velocity Control Subsystem (AVCS), and compensated with an attitude thruster firing. The impact on the trajectory in this instance was readily followed by the Kalman Filter State estimation process. In the situation shown in Figure IX, only a contingency upload was required to maintain navigation performance and a new reference trajectory was built approximately one week after the event to prevent linearity failures.

Satellite Vehicle thruster momentum dumps look very similar to Figure IX, but may be more severe. Normally the satellite will reduce reaction wheel momentum through variable setting magnets which torque the vehicle. Occasionally, momentum will saturate the satellite's reaction wheels at which time the vehicle's AVCS will command a thruster momentum dump. Sometimes this event may not be seen by the MCS due to visibility constraints. The momentum dump will manifest itself with growing ERD's and ORD's with respect to the ephemeris estimate and after twenty four hours, the pseudorange residuals will be rejected since the resulting states are above the calculated limits, indicating a mismodelling problem. In this situation, the Kalman Filter is reprocessed with an ephemeris covariance "Q-Bump" near the expected time of the thruster momentum dump. Usually only a contingency upload is needed to maintain navigation performance followed up later with a new reference trajectory based on the changed state estimates. If the pseudorange residuals, ERD's and ORD's are over tolerable limits the satellite is set "Unhealthy" to the user community.

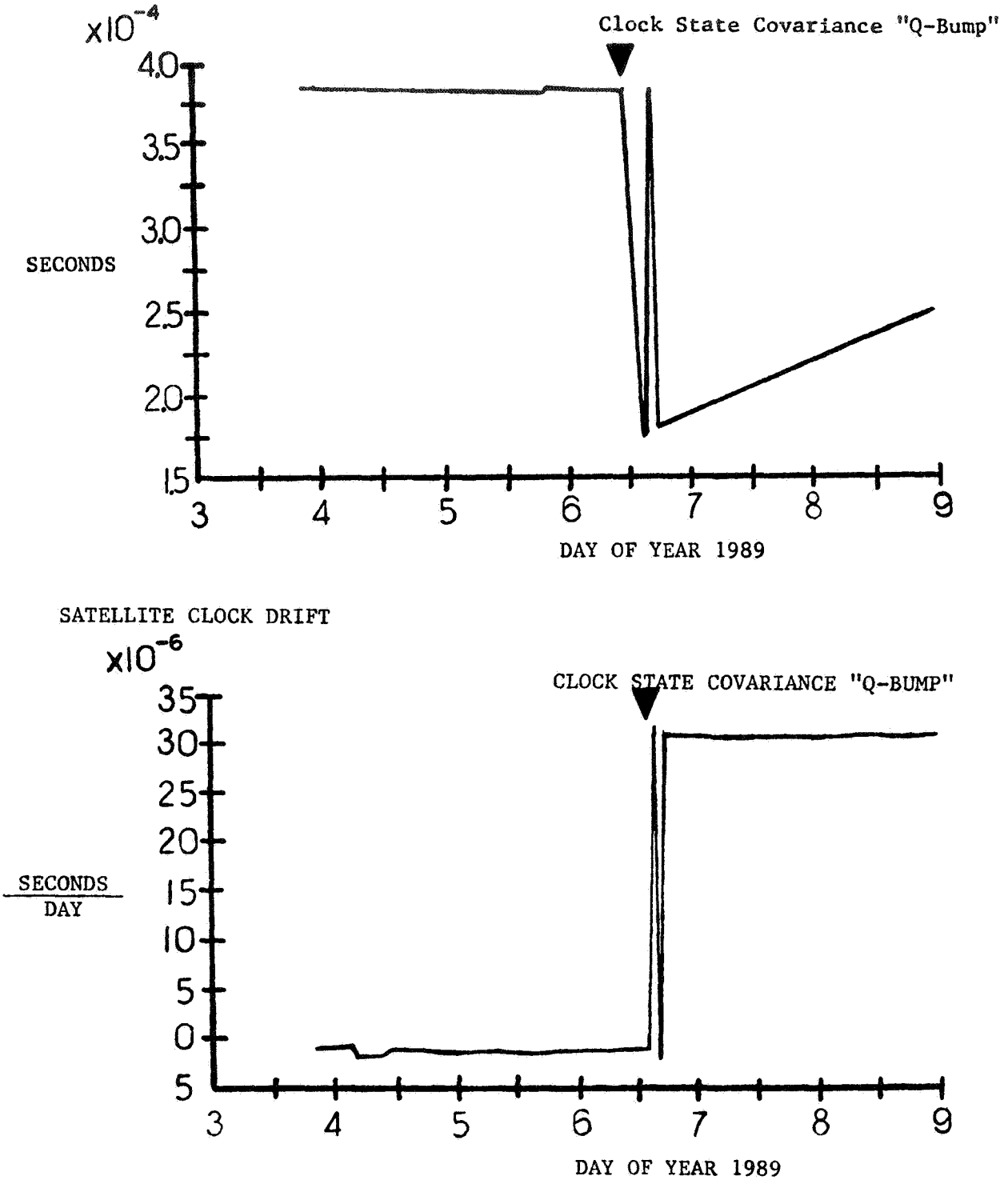


FIGURE VIII: The Satellite Clock State Reaction to a "Q-Bump" After a Clock Failure and Switch on NAVSTAR 11. The Determination of the New Clock States was Relatively Quick. Very Small Perturbations Affected the Ephemeris States. A Full Description of This Anomaly is in the Text.

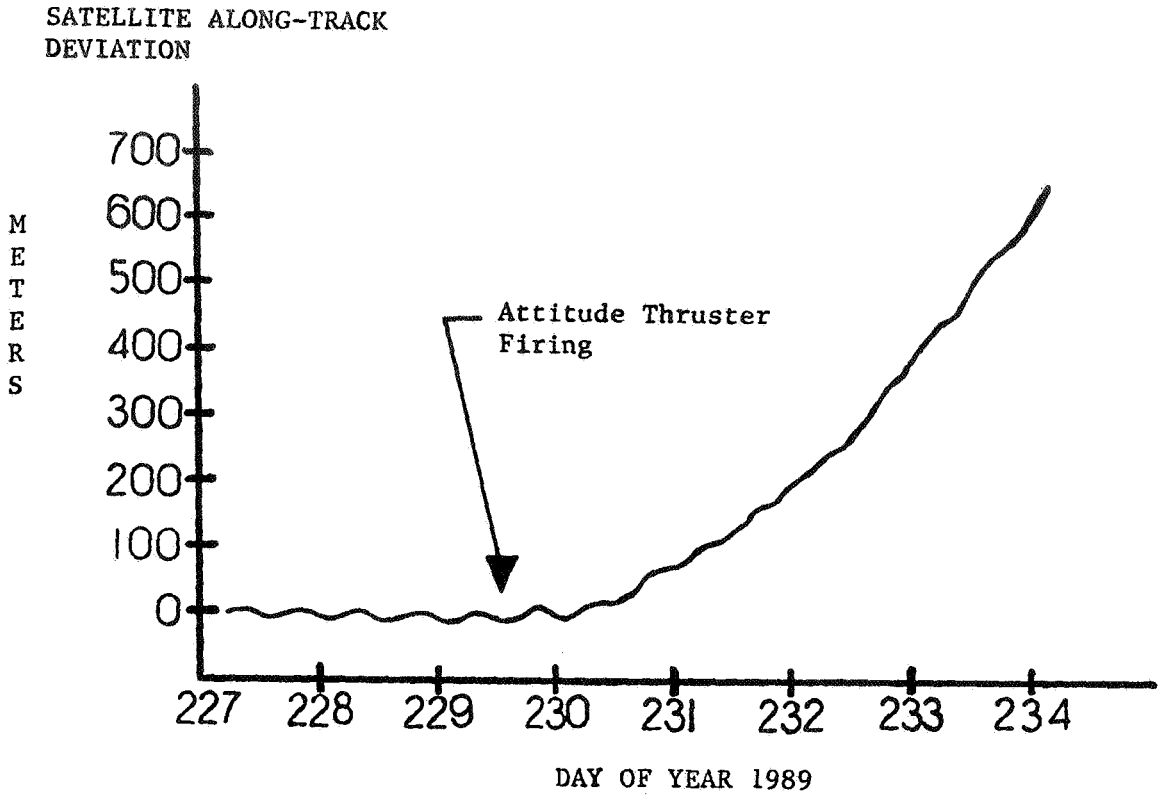


FIGURE IX: The Affect on Ephemeris State Estimates After an Attitude Thruster Firing on NAVSTAR 10. Satellite Thruster Momentum Dumps and Well-Modeled Delta-V Manuevers Show the Same Characteristics as this Plot. Radial and Cross-Track Deviations are not Shown for Brevity. There wasn't any Noticeable Clock State Perturbations due to this Ephemeris Change.

DELTA-V PLANNING AND MODELING IMPROVEMENTS

Prior to 1989, satellites undergoing Delta-V orbital maneuvers were usually set "Unhealthy" to the user community for a period of four days after the trajectory change. The engineers and the analysts in the 2 SCS didn't believe that this type of performance was acceptable. There were several efforts made to shorten the total "Unhealthy" time due to a Delta-V.

The first improvements were made in the area of Delta-V planning. The location of a GPS satellite is required to be at a specific longitude of ascending node with a tolerance of +/- 2 degrees. When a satellite originally approached this tolerance boundary, a station maintenance Delta-V was performed to return the satellite to its targeted geographic node. Since longitudinal acceleration always has a constant direction with respect to a specific geographic node, it was recognized that targeting for the far limit boundary would increase the time between station maintenance maneuvers. Figure X pictures the difference between these two targeting schemes. But there was one complication. Due to the configuration of the first generation GPS spacecraft, Delta-V's had to occur within specific beta-angle windows. This led into research directed at characterizing the performance of the 0.1 lb thrusters used for station maintenance maneuvers in GPS. Previous Delta-V planning tools modeled the thrust produced by the 0.1 lb rocket engine as an initial thrust followed by a linear decrease in force. Data from the manufacturer of the rocket engines showed a thrust build up to a peak followed by a decreasing force. The difference between these two thrust models is shown in Figure XI. Since small errors in the force model of a Delta-V can translate into large position errors several months down the road, a program was written incorporating this new thruster model. Aiming for maximum time between station keeping maneuvers with accurate beta-angle window placement reduced the number of Delta-V's and the total "Unhealthy" time due to station keeping.

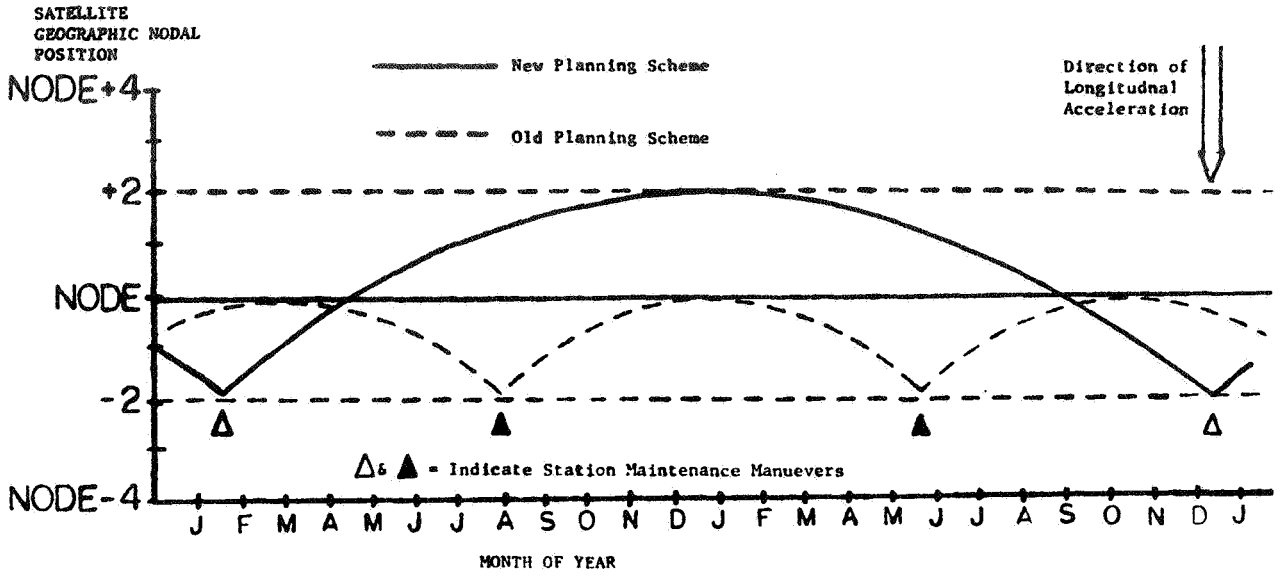


FIGURE X: Pictoral Representation Showing The Difference Between the Old and New Methods of Station Maintenance Manuever Planning. For The New Method To Work, Accurate Knowledge of Thruster Characteristics is Needed To Precisely Target a Beta-Angle Window.

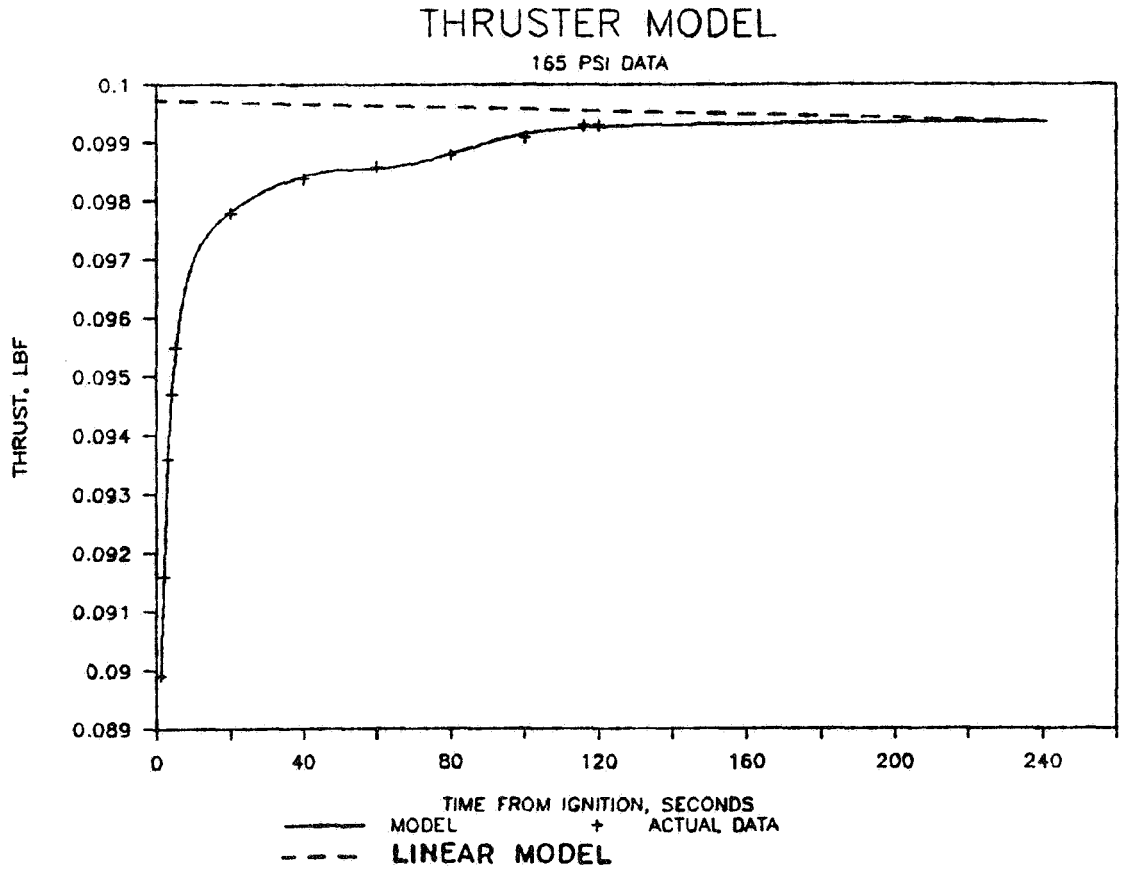


FIGURE XI: The Linear Thrust Model Compared to a Model Based On The Manufacturer's Test Data for The NAVSTAR Spacecraft 0.1 lb Hydrazine Thrusters. Incorporating The New Model Into The Planning Process for Station Maintenance Maneuvers has Improved Performance.

The MCS defines total Delta-V efficiency as the observed change in orbit semi-major axis compared to the expected change in the semi-major axis. The first generation, research and development (Block I), Navstar spacecraft tend to have Delta-V efficiencies of 92%. The second generation, operational (Block II) satellites generally have Delta-V efficiencies of 99%. Many factors can affect Delta-V efficiency including rocket exhaust body impingement, variable thruster efficiencies due to temperature, catalyst, engine misalignment differences, or attitude motion and attitude thruster firings. Figure XII shows a Block II satellite with major axis body reference frame. The 0.1 lb thrusters are located on a thruster pod as indicated in this figure. Block I satellites have a similar configuration. One major difference between the two types of vehicles is that the Block I spacecraft only have thrusters located along the Y body axis. Block II spacecraft have thrusters aligned along both the Y and X body axes. Given that the Y axis rocket engine exhaust impinges against the satellite body, and Block II Delta-V's are performed using the X body axis thrusters, the majority of the Block I Delta-V efficiency loss can probably be attributed to the thrust impingement. The Delta-V efficiency has been trended for each satellite and the MCS compensates for inefficiencies during the planning of a Delta-V. Unfortunately, Delta-V efficiency doesn't only vary from vehicle to vehicle, but between Delta-V's as well.

Delta-V maneuvers are modeled in the Kalman Filter's reference trajectory for the satellite undergoing the maneuver. Before 1989, an impulse model was used to account for the Delta-V in the reference trajectory. The impulse model instantaneously changes the reference trajectory at the midpoint time of the orbital maneuver. Though the impulse model was adequate for short duration maneuvers, Delta-V's of varying, and greater lengths than a few minutes were not well modeled. A thrust model was implemented which integrates the force of the rocket engines over the duration of the burn. This thrust model still uses the old linear force model mentioned above, but the prediction accuracy was still better than the impulse model. The thrust model had another advantage; the effects of attitude motion could be approximated with the available thruster misalignment inputs. When rocket engine characterization work began, a database of attitude motion during maneuvers also was started. This data could now be incorporated into the thrust model to better improve its prediction accuracy.

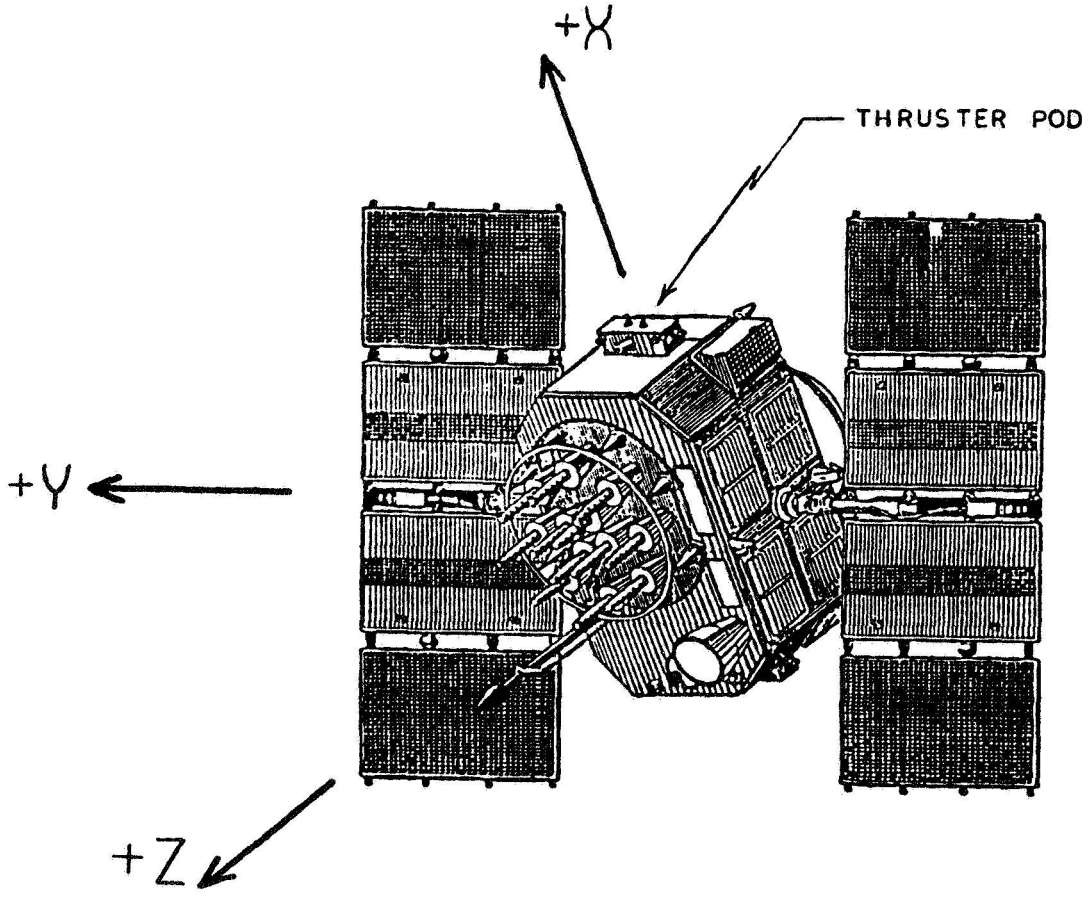


FIGURE XII: Picture of the Block II NAVSTAR Spacecraft with the Thruster Pod Indicated. On Block II Satellites, Station Maintenance Thrusters are Located Along Both The X and Y Axes, Though Usually Only The X Direction Thrusters are Used. Block I Satellites Only Have The Y Axis Thrusters.

Until very recently, immediately after a thruster firing, a "Q-Bump" on the satellite's ephemeris states was used to account for the errors in the Delta-V model. The parameters used in the covariance matrix for the "Q-Bump" were excessively large given the available thrust model. Work had begun on optimizing the values to set the covariance matrix after a Delta-V until it was discovered that the process noise expression in Equation 14 could be modified with values based on the Delta-V duration and expected error. Between the new thrust model, approximation of attitude errors, and the addition of maneuver process noise, large Delta-V's are beginning to look like the attitude thruster firing in Figure X. Before 1989, satellites were unavailable to the user community for up to four days after a maneuver. Today the same maneuver will cause under six hours of unavailability.

But improvements can still be made. An effort is under way to include the same rocket engine thrust curve pictured in Figure XI and used with maneuver planning in the reference trajectory thrust model. A rather simple relation is used currently to calculate the amount of process noise. Though this relation performs well, determination whether it's optimal needs to be made. In the near future, attempts to approximate the attitude thruster firing which occurs during a typical Delta-V will be made.

CONCLUSION AND RECOMMENDATIONS

Overall, The MCS Kalman Filter is quite robust. The bottom line is that highly accurate navigation data is routinely transmitted to a worldwide community of military and civilian users with minimal interruption in service. The online monitoring and control of the Kalman Filter is relatively simple. However, some improvements in the operation of the MCS Kalman Filter could be made. For example, many parameters which reside in the MCS database, such as the [P] matrix value for "Q-Bumps" or Delta-V process noise would yield more control if they were updatable on the online system. The aging of the reference trajectory would be less and ephemeris upload prediction would be better if more orbital perturbations were modeled. Better clock predictions should be possible.

The MCS should have a capability to monitor user accuracy that is independent of the Kalman Filter. The way ERD's and ORD's are calculated assumes that the Kalman Filter solution is truthful. It's easily possible for the MCS operators to see large ERD's and ORD's on a satellite which is a result of Kalman Filter corruption. Thanks to the MCS partitioning scheme this error can be caught rather quickly as the corruption starts to affect other states in the partition. Unfortunately it's still possible to be fooled and mistakenly upload satellites with corrupted data, therefore an independent check is needed.

For Delta-V's, it's apparent with the proper engineering data on a vehicle system and associated rocket engines, highly accurate trajectory predictions are possible. Though relatively late for the GPS system, satellite attitude motion and attitude thruster firings during a maneuver could have been accounted for real time with data from rate gyroscopes on board the satellite. This data could have been used in a specialized Kalman Filter routine designed for maneuvering spacecraft. To counter the variability in Delta-V efficiencies, it is highly recommended to have thrusters located on a satellite such that their thrust is directed outwards, minimizing any potential body impingement.

Finally, designers should be careful not to oversimplify the data available to the operator of systems incorporating a Kalman Filter, while not making the operation of the Filter overly complex. The full potential of a system should also be realized given the design constraints if at all possible. I believe the MCS Kalman Filter experience shows what operations personnel can pull off given the relative freedom to make improvements.

REFERENCES

1) Information on the general nature and history of the Global Positioning System can be found in:

"Global Positioning System" Volumes I, II, and III, Published by The Institute of Navigation, Washington DC., 1980.

2) Information on the design of the MCS Kalman Filter was obtained from:

"Computer Program Development Specification for the MCS Ephemeris/Clock Computer Program: Appendix A" USAF Contract F04701-80-C-0011, Specification Number CP-MCSEC-302C.

ACKNOWLEDGMENTS

Many thanks to Dr. Ming Chien, Dr. David Winfield, and Dr. John Berg for not only teaching me most of what I know about GPS, but also for playing nursemaid to myself and others through many of the activities mentioned in this paper.

FLIGHT MECHANICS/ESTIMATION THEORY SYMPOSIUM

MAY 22-24, 1990

SESSION 4

GAMMA RAY OBSERVATORY (GRO) OBC ATTITUDE ERROR ANALYSIS

by
R. R. Harman
NASA - Goddard Space Flight Center
Greenbelt, MD 20771

ABSTRACT

This analysis involves an in-depth look into the OBC attitude determination algorithm. A review of TRW error analysis and necessary ground simulations to understand the onboard attitude determination process are performed. In addition, a plan is generated for the in-flight calibration and validation of OBC computed attitudes. Pre-mission expected accuracies are summarized and sensitivity of onboard algorithms to sensor anomalies and filter tuning parameters are addressed.

1.0 INTRODUCTION

The Gamma Ray Observatory (GRO) (see Reference 10) is a three axis stabilized spacecraft scheduled to be launched into a 350-450 Km orbit 1990 by the Space Transportation System (STS). The GRO science instruments study gamma ray sources between 0.1 to 30000 mega-electron-volts (MeV) before they are absorbed by the Earth's atmosphere. The spacecraft is designed to stay inertially pointed, using reaction wheel control, for two weeks at a time before maneuvering to the next gamma ray target.

GRO has an onboard attitude determination accuracy requirement of 86.4 arcseconds per axis (3 sigma) during the normal science observation mode. This accuracy is accomplished by the use of two Fixed Head Star Trackers (FHSTs) and an Inertial Reference Unit (IRU). Both of these attitude sensors have been used on the Solar Maximum Mission (SMM), LANDSAT 4, and LANDSAT 5 spacecraft. As a backup, the Fine Sun Sensor (FSS) can take the place of a FHST with the resultant attitude accuracy of 167.5 arcseconds/axis (3 sigma). In both cases, the attitude is propagated using the IRU data and updated after a FHST or FSS measurement by using an extended Kalman Filter.

2.0 GRO ONBOARD ATTITUDE ESTIMATION (Reference 1)

Time Propagation

In GRO, the attitude computations are contained in two modules: kinematic integration module and attitude estimation module. The kinematic integration routine uses the previous cycle OBC quaternion and the current gyro output to update the OBC quaternion. The kinematic equation for updating a quaternion is (Reference 2):

$$q(t_{n+1}) = \left\{ \cos(wT/2)\ I + 1/w\ \sin(wT/2) \begin{vmatrix} 0 & w_z & -w_y & w_x \\ -w_z & 0 & w_x & w_y \\ w_y & -w_x & 0 & w_z \\ -w_x & -w_y & -w_z & 0 \end{vmatrix} \right\} q(t_n) \quad \ldots 2.1$$

where $w = (w_x^2 + w_y^2 + w_z^2)^{1/2}$
T = time interval
I = 4x4 identity matrix.

Since the gyro output consists of three angles θ_x, θ_y, θ_z the following substitutions can be made:

$$\theta_x = w_x T, \quad \theta_y = w_y T, \quad \theta_z = w_z T$$

Equation 2.1 then becomes:

$$q(t_{n+1}) = \left\{ \cos(\theta/2)\, I + 1/\theta \sin(\theta/2) \begin{vmatrix} 0 & \theta_z & -\theta_y & \theta_x \\ -\theta_z & 0 & \theta_x & \theta_y \\ \theta_y & -\theta_x & 0 & \theta_z \\ -\theta_x & -\theta_y & -\theta_z & 0 \end{vmatrix} \right\} q(t_n) \quad \ldots 2.2$$

where $\theta = (\theta_x^2 + \theta_y^2 + \theta_z^2)^{1/2}$

Every 32.768 seconds, the attitude estimation routine (ATTEST) generates roll, pitch, and yaw errors. These errors are fed into the kinematic integration routine in place of the normal gyro data that is used between 32.768 second updates.

The attitude estimation routine (ATTEST) itself consist of an extended Kalman filter (KF). Reference 1 contains an outline of ATTEST. The KF is implemented in two steps. First, the propagation of the internal statistics based on the Dynamics Model and second, updating the state vector based on the Observation Model, the measurements, and the internal statistics. ATTEST alternates between the two sensors (FHST/FHST or FHST/FSS) being used for attitude estimation every 32.768 seconds.

Dynamics Model

The gyro rate measurement is assumed to have the following form:

$$\dot{\underline{\theta}} = \underline{w} + \underline{b}_o + \underline{b} - \underline{n}_v$$

$$\dot{\underline{b}} = \underline{n}_u$$

where, $\dot{\underline{\theta}}$ - gyro rate measurement
$\underline{w}$ - true spacecraft rate
$\underline{b}_o$ - gyro bias error
$\underline{b}$ - gyro random walk error
$\underline{n}_v$ - float torque noise (gaussian white noise)
$\underline{n}_u$ - float torque derivative noise (gaussian white noise)

The gyro drift error, $\dot{e}$, is defined as follows:

$$\dot{\underline{e}} = \underline{w} - \dot{\underline{\theta}}$$

It then becomes the following equation:

$$\dot{\underline{e}} = -\underline{b}_o - \underline{b} + \underline{n}_v$$

The gyro bias $\underline{b}_o$ is assumed to be known and can be taken out of the above equation. Therefore:

$$\dot{\underline{e}} = -\underline{b} + \underline{n}_v$$

$$\dot{\underline{b}} = \underline{n}_u$$

The attitude error, $\underline{\Psi}$, is computed as follows:

$$\dot{\underline{\Psi}} + \underline{w} \times \underline{\Psi} = \dot{\underline{e}}$$

However since $\underline{w}$ is neglible, the dynamic model is reduced to the following form:

$$\dot{\underline{\Psi}} = -\underline{b} + \underline{n}_v$$

$$\dot{\underline{b}} = \underline{n}_u$$

If these two equations are put into a linear state space formulation equations (2.3) and (2.4) are derived:

$$\dot{\underline{X}}(t) = F\,\underline{X}(t) + \underline{W}(t) \tag{2.3}$$

$$\dot{\underline{X}}(t) = \begin{vmatrix} \dot{\underline{\Psi}} \\ \dot{\underline{b}} \end{vmatrix} = \begin{vmatrix} 0_{3x3} & -I_{3x3} \\ 0_{3x3} & 0_{3x3} \end{vmatrix} \begin{vmatrix} \underline{\Psi} \\ \underline{b} \end{vmatrix} + \begin{vmatrix} \underline{n}_{v(3x1)} \\ \underline{n}_{u(3x1)} \end{vmatrix} \tag{2.4}$$

where, $\underline{\Psi}$ - attitude error
$\underline{b}$ - gyro random walk error
$\underline{n}_v$ - float torque noise (Gaussian)
$\underline{n}_u$ - float torque derivative noise (Gaussian)

The state equation is discretized to the following form:

$$\underline{X}(t_k) = \phi_k\,\underline{X}(t_{k-1}) + \underline{W}(t_k)$$

where $\phi_k = e^{AT_k}$ and $T_k = t_k - t_{k-1}$.

The two characteristics of $\underline{W}(t)$ are the mean:

$$E\,[\,\underline{W}(t)\,] = 0$$

And the covariance:

$$E\,[\,\underline{W}(t)\,\underline{W}^T(t')\,] = \begin{vmatrix} \underline{n}_v\underline{n}_v^T & 0_{3x3} \\ 0_{3x3} & \underline{n}_u\underline{n}_u^T \end{vmatrix} \int (t-t') \tag{2.5}$$

where "T" denotes the transpose. Note that the off diagonal elements in (2.5) are zero since it is assumed that there is no correlation between $\underline{n}_u$ and $\underline{n}_v$.

The Spectral Density Matrix is defined as follows:

$$Q(t) = E\,[\,\underline{W}(t)\,\underline{W}^T(t)\,] \tag{2.6}$$

Thus the covariance is given as

$$Q(t)\,\delta(t-t') \tag{2.7}$$

The discrete Dynamics Noise Covariance matrix, Q_k, is obtained using the state transition matrix, ϕ_k, and the Spectral Density matrix, $Q(t)$, in the following manner:

$$Q_k = \int_{t_{k-1}}^{t_k} \phi(t_k,t')\, Q(t')\, \phi^T(t_k,t')\, dt' \tag{2.8}$$

Once Q_k is computed, it is used to propagate the state covariance matrix as follows:

$$P_k\,(-) = \phi_k\, P_{k-1}(+)\, \phi_k^T + Q_k \tag{2.9}$$

where $P_k\,(-)$ is the Propagated Covariance Matrix at time k and,
$P_{k-1}\,(+)$ is the updated Covariance Matrix at time k-1.

Observation Model

FHST Model

In the GRO Flight Software, the FHST measurements are used to create an observed star unit vector, $\underline{OS}$, in the Sensor Coordinate frame. The identified star position in the star catalog is used to create an expected or computed unit star vector, $\underline{CS}$, in the Sensor Coordinate Frame. We then define

$$Z_k(i) = OS_k(i) - CS_k(i) \quad \text{for } i=x \text{ and } y \tag{2.10}$$

where Z_k is the measurement residual.

From this definition of Z_k , H_k is shown to be

$$H_k = \begin{bmatrix} (\underline{X} \times \underline{S}_k)^T & 0_{1x3} \\ (\underline{Y} \times \underline{S}_k)^T & 0_{1x3} \end{bmatrix} \tag{2.11}$$

where $\underline{S}_k$ is the observed star vector in the spacecraft body frame,
$\underline{X}$ is the X axis of the FHST in the spacecraft body frame,
and $\underline{Y}$ is the Y axis of the FHST in the spacecraft body frame.

In the observation model

$$\underline{Z}_k = H_k \underline{X} + \underline{V}_k \tag{2.12}$$

where $\underline{Z}_k$ is the observation defined in (2.9), and $\underline{V}_k$ is the sensor noise (Gaussian).

The sensor noise characteristics are the following:

$$E[\ \underline{V}_k\ ] = 0 \tag{2.13}$$

$$R_k = E[\ \underline{V}_k\ \ \underline{V}_1^T] = \begin{bmatrix} R_{11} & 0 \\ 0 & R_{22} \end{bmatrix} \tag{2.14}$$

It is further assumed that the initial state vector $\underline{X}_o$ is Gaussian and $\underline{X}_o$, $\underline{W}$,and $\underline{V}_k$ are independant of each other. Since all are assumed zero mean and Gaussian, this is equivalent to assuming they are uncorrelated with each other.

FSS Model (Reference 10)

As with the FHST, the FSS model uses an observed Sun position, $\underline{OS}$, and a computed Sun position, $\underline{CS}$, to compute measurement residuals, $\underline{Z}$, as follows:

$$Z_k(i) = OS_k(i) - CS_k(i) \text{ for } i=x \text{ and } y.$$

The measurement equation is the same used for the FHST (2.12). For the FSS, the H_k is shown to be:

$$H_k = \begin{bmatrix} (\ \underline{X}_{MP} \times \underline{S}_k\ )^T & 0_{1x3} \\ (\ \underline{Y}_{MP} \times \underline{S}_k\ )^T & 0_{1x3} \end{bmatrix} \tag{2.15}$$

where $\underline{S}_k$ is the computed sun vector. $\underline{X}_{MP}$ and $\underline{Y}_{MP}$ are as follows:

$$\begin{aligned} X_{MPx} &= (\ X_{Fx} - Z_{Fx}\ XP\ )\ /\ (\ \hat{\underline{Z}}_F \cdot \underline{S}_k\ ) \\ X_{MPy} &= (\ X_{Fy} - Z_{Fy}\ XP\ )\ /\ (\ \hat{\underline{Z}}_F \cdot \underline{S}_k\ ) \\ X_{MPz} &= (\ X_{Fz} - Z_{Fz}\ XP\ )\ /\ (\ \hat{\underline{Z}}_F \cdot \underline{S}_k\ ) \\ Y_{MPx} &= (\ Y_{Fx} - Z_{Fx}\ YP\ )\ /\ (\ \hat{\underline{Z}}_F \cdot \underline{S}_k\ ) \\ Y_{MPy} &= (\ Y_{Fy} - Z_{Fy}\ YP\ )\ /\ (\ \hat{\underline{Z}}_F \cdot \underline{S}_k\ ) \\ Y_{MPz} &= (\ Y_{Fz} - Z_{Fz}\ YP\ )\ /\ (\ \underline{Z}_F \cdot \underline{S}_k\ ) \end{aligned}$$

where $\underline{X}_F$ is the FSS X-coordinate axis in the spacecraft frame,
$\underline{Y}_F$ is the FSS Y-coordinate axis in the spacecraft frame,
$\underline{Z}_F$ is the FSS Z-coordinate axis in the spacecraft frame,
XP and YP are the FSS expected measurements.

The FSS noise characteristics are the same as those for the FHST.

Update Algorithms

The state vector is updated by processing the following equation with the inputs $P_k(-)$ (2.9), H_k (2.11 and 2.15), R_k (2.14), and the observation vector Z_k (2.12):

$$K_k = P_k(-)\ H_k^T\ [\ H_k\ P_k(-)\ H_k^T + R_k]^{-1} \tag{2.16}$$

$$P_k(+) = (\ I - K_k\ H_k)\ P_k(-) \tag{2.17}$$

$$\hat{\underline{X}}_k(+) = \hat{\underline{X}}_k(-) + K_k(\ \underline{Z}_k - H_k\ \hat{\underline{X}}_k(-)\ ) \tag{2.18}$$

where $P_k(+)$ is the updated Covariance Matrix.
K_k is the Kalman Gain Matrix.
$\hat{\underline{X}}_k(+)$ is the updated State Vector.

The GRO Flight Software employees a scaler implementation method which requires the sequence of equations (2.16-2.18) to be executed twice. In the first pass the following substitutions are made for the FHSTs (similarly for the FSS):

$$H_k = H_{k,1} = [\ (\ \underline{X} \times \underline{S}_k\ )^T \quad 0_{1x3}\ ]$$
$$R_k = R_{k,1} = R_{11}$$

The resulting Kalman Gain Matrix $K_{k,1} = K_k$ is used to update the covariance matrix (2.17) where $P_{k,1} = P_k$ and the update (2.18) where $\hat{\underline{X}}_k(-) = 0$. The equations are listed below:

$$K_{k,1} = P_k(-)\ H_{k,1}^T\ /\ [\ H_{k,1}\ P_k(-)\ H_{k,1}^T + R_{k,1}\ ]$$
$$P_{k,1}(+) = [\ I - K_{k,1}\ H_{k,1}]\ P_k(-)$$
$$X_{k,1}(+) = K_{k,1}\ Z_{k,1}$$

where $Z_{k,1}$ is the X component of $\underline{Z}_k$.

In the second pass, there are the following substitutions:

$$H_k = H_{k,2} = [\ (\ \underline{Y} \times \underline{S}_k\ )^T \quad 0_{1X3}\ ]$$
$$R_k = R_{k,2} = R_{22}$$
$$K_k = K_{k,2}$$
$$\hat{P}_k(-) = \hat{P}_{k,1}(+)$$
$$\hat{\underline{X}}_k(-) = \hat{\underline{X}}_{k,1}(+)$$

where $\hat{\underline{X}}_{k,1}(+)$ is the state vector update from the first pass.

The final Kalman Gain Matrix $K_k = K_{k,2}$ is used to update the covariance matrix and the state vector. The equations are listed below:

$$K_{k,2} = P_k(-)\ H_k^T\ /\ [\ H_k\ P_k(-)\ H_k^T + R_k]$$
$$P_k(+) = [\ I - K_k\ H_k]\ P_k(-)$$
$$\hat{\underline{X}}_k(+) = \hat{\underline{X}}_k(-) + K_k[\ Z_{k,2} - H_k\ \hat{\underline{X}}_k(-)\ ]$$

where $Z_{k,2}$ is the Y component of $\underline{Z}_k$.

3.0 ADEAS

The Attitude Determination Error Analysis System (ADEAS) was the attitude tool used in this analysis. It models state estimation using either a batch filter or a Kalman filter. The ADEAS Kalman filter is the same two pass filter implementation as described for GRO in Section 2.0. Thus, ADEAS provides a convient method for GRO onboard attitude error analysis.

ADEAS models an attitude system by use of consider and solve-for parameters. The solve-for parameters are those the onboard filter uses in its state vector. For GRO, the solve-for state components are the three attitude errors and the three gyro drift errors. The consider parameters are those that the onboard filter does not take in account such as misalignments and scale factor errors.

4.0 ONBOARD ATTITUDE ACCURACY

4.1 Attitude Accuracy Requirement

Using the above attitude estimation algorithm, the overall attitude determination requirement and the TRW stated capabilities are as follows (Reference 3):

	Requirement (arc-sec) (3-sigma)	Capability (arc-sec) (3-sigma)
Attitude determination accuracy using two FHSTs	86.4	71.2
Attitude determination accuracy using 1 FHST and 1 FSS	167.5	143.8

The FHST/FHST algorithm errors were the largest single contributer to the error budget at 53.3 arcseconds/axis. The FSS alignment error of 97.5 arcseconds/axis was the largest contributer to the FHST/FSS error budget. According to TRW error budgets, the absolute attitude determination requirement is met for both cases with a reserve.

ADEAS simulations were conducted to independantly verify that the Onboard attitude determination requirement could be met. According to TRW analysis, the update filter required 3 hours to converge (Reference 3). The simulation length consisted of the three hour convergence time plus one orbit of data. The primary error parameters used in these simulations are listed below:

Input parameters

Dynamic Noise (Reference 3)
White 4.2459E-2 arcseconds/sec$^{1/2}$ per axis
Random walk 4.4413E-5 arcseconds/sec$^{3/2}$ per axis
Misalignments
FHST # 1 32 arcseconds/axis (Table 1)
FHST # 2 32 arcseconds/axis (Table 1)
FSS 97.5 arcseconds/axis (Table 2)
Gyro 56 arcseconds/axis (Reference 4)
Measurement Noise
FHST #1 32.3 arcseconds (Note 1)
FHST #2 32.3 arcseconds (Note 1)
FSS 82.1 arcseconds (Note 2)
Gyro Scale Factor Error (3000 PPM) (Reference 5)
Initial Attitude Error (1800 arcseconds/axis)
Initial Gyro Drift Error (0.5 arcseconds/second/axis)
Systematic Calibration Errors (FHST #1 and FHST #2)
H and V measurements - 7 arcseconds (Reference 6)
Attitude Stabilization Errors 0.096 degrees (Reference 3)
Simulation Length (273.58 minutes)

For each simulation, the algorithm uncertainties and the jitter due to reaction wheel disturbances were RSS'd with the resultant simulation attitude errors in order to compare the simulation results to the error budget.

4.2 FHST/FHST CASE

This case is the nominal configuration for attitude estimation onboard and the most accurate. After the three hour convergence period, the maximum filter roll, pitch, and yaw attitude errors were 64.11, 64.74, and 60.80 arcseconds respectively. The gyro drift errors for the roll, pitch, and yaw axes were 3.335×10^{-3}, 3.356×10^{-3}, 3.899×10^{-3} arcseconds/second respectively. The onboard attitude accuracies are well within the 86.4 arcsecond/axis requirement. From Figure 4.1, steady state convergence occurs approximately 6200 seconds into the simulation.

4.3 FHST/FSS Case

This case is used only if one FHST fails, and the resultant accuracy degrades considerably. After the three hour convergence period, the maximum roll, pitch, and yaw attitude errors were 126.85, 120.13, and 80.12 arcseconds respectively. The roll, pitch, and yaw gyro drift errors were 4.046×10^{-3}, 3.379×10^{-3}, and 6.376×10^{-3} arcseconds/second respectively. The attitude accuracies are well within the 167.5 arcsecond/axis requirement. From Figure 4.2, the steady state convergence occurs at approximately 8800 seconds.

4.4 1 FHST With Two Guide Stars Case

A simulation of 56000 seconds was made using FHST #2 and two guide stars within one degree of the FHST boresight. The stars were measured alternately every 32 seconds. The attitude estimation errors were smaller for this case than for the FHST/FSS case. From Figure 4.3, the convergence time for this simulation was approximately 48000 seconds which is about eight times that of the FHST/FHST case and six times that of the FHST/FSS case. The longer convergence time is understandable from observability reasons alone. The maximum roll, pitch, and yaw attitude estimation errors over the last 8000 seconds of the simulation were 100.34, 100.03, and 64.38 arcseconds respectively. The roll, pitch, and yaw gyro drift error were 5.378×10^{-3}, 5.375×10^{-3}, and 3.382×10^{-3} arcseconds/second respectively.

5.0 SENSITIVITY ANALYSIS

For the sensitivity analysis, the consider parameters used for the ADEAS simulations were increased by a factor of two. The resulting ADEAS attitude errors are RSS'd with the algorithm implementation error and jitter due to the reaction wheels. The case designations were as follows:

[1]From reference 7, the FHST calibration error is 30 arcseconds (3-sigma), and the noise equivalent angle (NEA) is 24 arcseconds (3-sigma). The NEA is reduced to 12 arcseconds by data averaging onboard. The resultant measurement error is the RSS of 30 and 12 arcseconds.

[2]From reference 8, the calibration error is 79.2 arcseconds (3-sigma), and the noise equivalent angle is 21.6 arcseconds. The resultant measurement error is the RSS of 79.2 and 21.6 arcseconds.

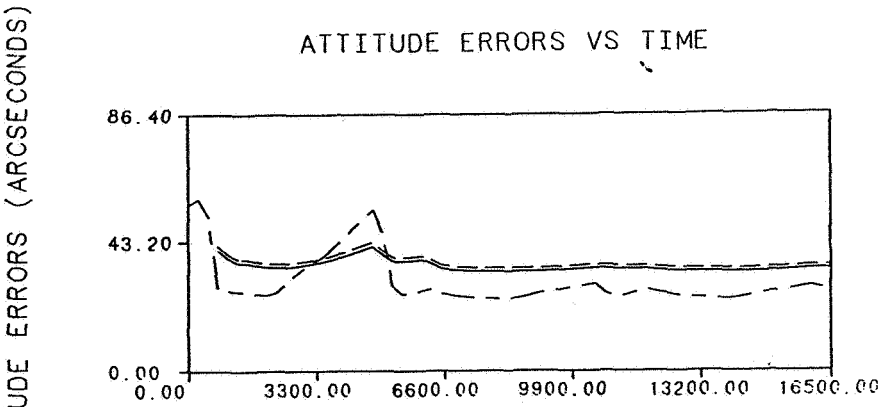


FIGURE 4.1 FHST #1 AND FHST #2

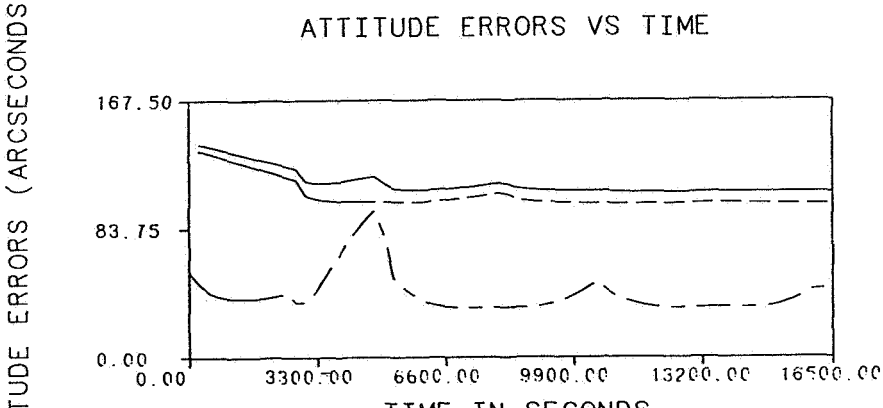


FIGURE 4.2 FHST #1 AND THE FSS

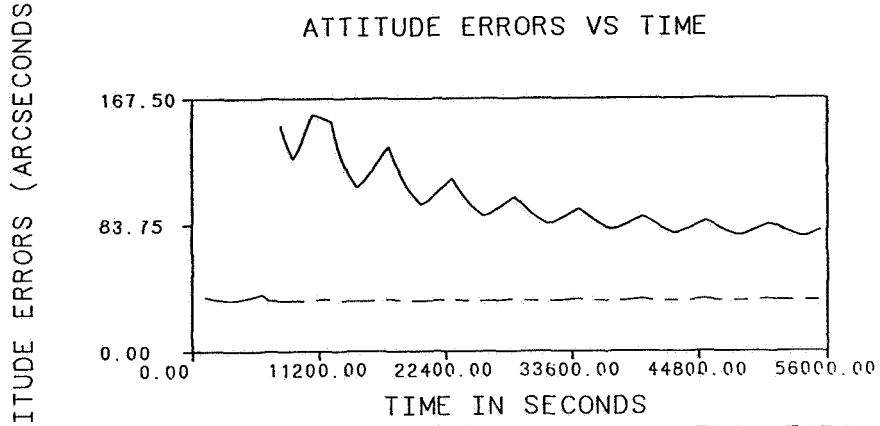


FIGURE 4.3 FHST #2 AND TWO GUIDE STARS

FHST/FHST Configuration

Case	Error Description
1A	Baseline
2A	2x Gyro white noise about each gyro axis
3A	2x Gyro random walk noise about each gyro axis
4A	2x Gyro scale factor error about each gyro axis
5A	2x Gyro misalignment about each gyro axis
6A	2x FHST #1 misalignment about each FHST #1 axis
7A	2x FHST #2 misalignment about each FHST #2 axis
8A	2x FHST noise

FHST/FSS Configuration

Case	Error Description
1B	Baseline
2B	2x Gyro white noise about each gyro axis
3B	2x Gyro random walk noise about each gyro axis
4B	2x Gyro scale factor error about each gyro axis
5B	2x Gyro misalignment about each gyro axis
6B	2x FHST #1 misalignment about each FHST #1 axis
7B	2x FSS misalignment about each FSS axis
8B	2x FHST noise
9B	2x FSS noise
10B	2x FHST and FSS noise

1 FHST with Two Guide Stars Configuration

Case	Error Description
1C	Baseline
2C	2x Gyro white noise about each gyro axis
3C	2x Gyro random walk noise about each gyro axis
4C	2x Gyro scale factor error about each gyro axis
5C	2x Gyro misalignment about each gyro axis
6C	2x FHST #2 misalignment about each FHST axis
7C	2x FHST noise

The attitude error results for the FHST/FHST case are as follows:

Attitude Determination Errors (arcseconds)

Case	Axis	Roll	Pitch	Yaw
1A		64.11	64.74	60.80
2A	x	64.24	64.75	60.80
2A	y	64.23	64.86	60.80
2A	z	64.11	64.74	60.92
3A	x	65.39	64.80	61.22
3A	y	64.18	66.04	61.22
3A	z	64.11	64.74	61.74
4A	x	64.11	64.74	60.80
4A	y	64.11	64.74	60.80
4A	z	64.11	64.74	60.80
5A	x	64.11	64.74	60.80
5A	y	64.11	64.74	60.80
5A	z	64.11	64.74	60.80
6A	x	74.80	76.02	61.22
6A	y	64.00	64.74	74.35
6A	z	64.12	64.74	60.80

7A	x	74.77	76.05	61.22
7A	y	64.12	64.73	75.42
7A	z	64.12	64.74	60.81
8A		65.39	66.10	60.47

The gyro drift error results for the FHST/FHST case are as follows:

Gyro Drift Estimation Errors (10^{-3} arcseconds/second)

Case	Axis	Roll	Pitch	Yaw
1A		3.335	3.356	3.899
2A	x	3.414	3.352	3.899
2A	y	3.332	3.432	3.899
2A	z	3.335	3.356	3.982
3A	x	6.228	3.404	4.414
3A	y	3.385	6.260	4.414
3A	z	3.335	3.356	6.185
4A	x	3.335	3.356	3.899
4A	y	3.335	3.356	3.899
4A	z	3.335	3.356	3.899
5A	x	3.335	3.356	3.899
5A	y	3.335	3.356	3.899
5A	z	3.335	3.356	3.899
6A	x	3.356	3.374	4.414
6A	y	3.165	3.201	6.250
6A	z	3.335	3.356	3.899
7A	x	3.356	3.374	4.414
7A	y	3.356	3.376	5.774
7A	z	3.335	3.356	3.899
8A		3.805	3.859	3.420

The increased gyro white noise and random walk noise about an axis primarily affects that axis as expected. There is some correlation between the X and Y axes but not enough to be significant. Of the two errors, the random walk component proves to affect the attitude errors the most. As expected, the random walk errors contribute the most to the gyro drift estimation errors (see equations 2.4-2.9). Between updates, this higher gyro drift error would degrade the attitude solution since the gyro data would compensated with an incorrect gyro drift estimate. The gyro scale factor errors and misalignments have no significant affect on the attitude solution since the spacecraft is inertially pointed and has no significant angular rates. As expected, the FHST misalignments have the largest affect on attitude accuracy. For both FHST #1 and #2, the misalignment of the X and Y tracker axes result in attitude estimation errors of over 70 arcseconds. Both FHSTs had a maximum attitude estimation error of 76 arcseconds when their Y-axis was misaligned. The increased FHST noise has only a small effect on the attitude error since there are sufficient measurements to reduce the scope of the error, and the system has good observability.

The attitude error results for the FHST/FSS cases are as follows:

Attitude Determination Errors (arcseconds)

Case	Axis	Roll	Pitch	Yaw

1B		127.21	121.30	74.65
2B	x	127.24	121.30	74.65
2B	y	127.24	121.40	74.65
2B	z	127.21	121.30	74.69
3B	x	128.04	120.13	80.12
3B	y	127.43	122.02	74.65
3B	z	127.21	121.30	74.91
4B	x	127.21	121.30	74.65
4B	y	127.21	121.30	74.65
4B	z	127.21	121.30	74.65
5B	x	127.21	121.30	74.65
5B	y	127.21	121.30	74.65
5B	z	127.21	121.30	74.65
6B	x	149.19	121.30	74.66
6B	y	127.21	121.30	91.91
6B	z	127.21	121.30	74.65
7B	x	208.19	209.49	74.76
7B	y	127.24	121.36	74.85
7B	z	127.21	121.33	74.65
8B		128.35	120.32	94.66
9B		131.64	126.30	74.93
10B		132.30	126.36	74.97

The gyro drift error results for the FHST/FSS case are as follows:

Gyro Drift Estimation Errors (10^{-3} arcseconds/second)

Case	Axis	Roll	Pitch	Yaw
1B		4.093	4.187	4.169
2B	x	4.147	4.183	4.169
2B	y	4.108	4.262	4.169
2B	z	4.093	4.187	4.248
3B	x	7.139	3.399	6.376
3B	y	4.277	6.653	4.169
3B	z	4.093	4.187	5.962
4B	x	4.093	4.187	4.169
4B	y	4.093	4.187	4.169
4B	z	4.093	4.187	4.169
5B	x	4.093	4.187	4.169
5B	y	4.093	4.187	4.169
5B	z	4.093	4.187	4.169
6B	x	4.093	4.187	4.169
6B	y	4.093	4.187	4.486
6B	z	4.093	4.187	4.169
7B	x	4.093	4.187	4.172
7B	y	4.093	4.187	6.534
7B	z	4.093	4.187	4.172
8B		5.195	3.838	5.767
9B		6.073	6.268	2.968
10B		6.343	6.329	3.533

As with the FHST/FHST simulations, an increase of gyro white noise or random walk about an axis primarily affects that axis. Not only does a correlation of the X and Y axes exist as in the FHST/FHST case, but a X and Z correlation exits. Of the two gyro noises, the random walk error has the largest affect on the attitude errors due

to attitude computations between measurement updates as mentioned in the previous case. The gyro scale factors and misalignments have no significant affect on the attitude or gyro bias errors due to the spacecraft being inertially pointed. The FHST #2 and FSS misalignments are the greatest contributors to attitude errors as expected since they define the attitude. The FSS affects the attitude errors more due to the greater alignment errors as compared to FHST # 2. Increasing the FHST noise results in the roll error increasing by 1 arcsecond, the pitch error decreasing by 1 arcsecond, and the yaw estimation error increasing by 20 arcseconds. The increased FHST noise to 64 arcseconds is much closer to the FSS measurement noise of 82 arcseconds. Thus, the FSS measurements would be weighted almost as much as the FHST measurements. The large FSS alignment uncertainties are then fed into the yaw estimation errors. The above maximum error occurs when the FHST is occulted which further degrades the yaw solution. Increasing the FSS noise increases the roll, pitch, and yaw estimation errors due to the limited memory of the filter to measurements. Thus, the measurement noise cannot be averaged out totally. The yaw error is only slightly higher than the baseline since any FSS is weighted much less than in the baseline while the FHST measurements are weighted the same. When the noise on the FSS and FHST are increased, the total error is due totally to the increased noise and the limited memory of the filter mentioned above.

The attitude error results for the 1 FHST with 2 Guide Stars are:

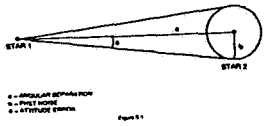
	Attitude Determination Errors (arcseconds)			
Case	Axis	Roll	Pitch	Yaw
1C		100.34	100.03	64.38
2C	x	100.43	100.07	64.38
2C	y	100.40	100.13	64.38
2C	z	100.34	100.03	64.56
3C	x	109.61	107.65	64.38
3C	y	108.02	109.23	64.38
3C	z	100.34	100.03	65.98
4C	x	100.34	100.03	64.38
4C	y	100.34	100.03	64.38
4C	z	100.34	100.03	64.38
5C	x	100.34	100.03	64.38
5C	y	100.34	100.03	64.38
5C	z	100.34	100.03	64.38
6C	x	107.77	107.46	64.38
6C	y	100.34	100.03	84.97
6C	z	107.77	107.43	64.38
7C		155.73	154.99	65.19

The gyro drift error results for the 1 FHST with 2 Guide Stars is:

	Gyro Drift Estimation Errors (10^{-3} arcseconds/second)			
Case	Axis	Roll	Pitch	Yaw
1C		5.378	5.375	3.382
2C	x	5.407	5.393	3.382

2C	y	5.396	5.400	3.382
2C	z	5.378	5.375	3.503
3C	x	8.856	6.653	3.382
3C	y	6.660	8.845	3.382
3C	z	5.378	5.375	6.394
4C	x	5.378	5.375	3.382
4C	y	5.378	5.375	3.382
4C	z	5.378	5.375	3.382
5C	x	5.378	5.375	3.382
5C	y	5.378	5.375	3.382
5C	z	5.378	5.375	3.382
6C	x	5.378	5.375	3.382
6C	y	5.378	5.375	3.382
6C	z	5.378	5.375	3.382
7C		6.811	6.804	3.694

As in the FHST/FHST case, an increases in the white and random walk noise about an axis primarly affects that axis. The correlation between the X and Y axes still exists. Of the two noises, the random walk has the largest affect on the attitude due to the increased gyro drift estimation error as mentioned above. The gyro scale factor error and misalignments as expected have no affect on the attitude estimation error. Also as expected, the added FHST alignment errors have a significant affect on the attitude estimation errors. The maximum attitude estimation error from the FHST misalignments is 107 arcseconds. The primary source of attitude estimation error is from the FHST noise. This results from the small separation of the guide stars in the FHST (2.8 degrees). The attitude estimation errors due to the measurement noise is an arctangent relationship seen in the following diagram (Reference 9):



For this simulation, this error was approximately 1290 arcseconds. The RSS attitude estimation error for the noise simulation is approximately 229 arcseconds. The filter was able to improve the solution by 1061 arcseconds. The filter's memory for the measurements was not long enough to average out the measurement noise further.

Of the three cases simulated, the FHST/FHST case is by far the most accurate as expected. The choice for the backup case is not as easy to choice. The disadvange of the FHST/FSS case is the FSS alignment errors. Unless the FSS alignment is updated often, alignment errors due to thermal effects could degrade the attitude accuracy significantly. From an alignment point of view, the 1 FHST case is the most preferable due to the FHST being mounted on a fairly stable platform which is beside the gyros. From a noise point of view, the 1 FHST case is worse than the FHST/FSS case. The accuacy could be improved if more than two guide stars were available or the angular separation was larger. A possible solution to the problem is to use the 1 FHST case as a backup provided at least two guide stars are available and the FHST noise has not increased significantly from

launch. The alignment errors would be less than that for the FHST/FSS case and more of the sky would be open for viewing due to the limited FOV of the FSS. Otherwise, if the FHST noise has increased significantly since launch or there are not enough guide stars, use the FHST/FSS case.

6.0 OBC ATTITUDE ESTIMATION CALIBRATION AND VALIDATION

Nominal Operations

After the spacecraft has completed inorbit checkout, there are two phases for validating the onboard calculated attitude: post normal maneuver mode phase and normal pointing mode phase. Once the normal maneuver is performed, the onboard filter convergence needs to be validated. This convergence is defined as steady state operation and should occur within 3 hours of filter initialization (Reference 3). Once the filter has converged, the mission is in the normal pointing mode phase. According to Teledyne documentation, the acceleration insensitive drift rate (AIDR) peak over six hours is 0.0006 arcseconds/second (Reference 4). Assuming the update filter was disabled, the attitude errors due to the AIDR alone could be 51.84 arcseconds in 24 hours. This requires the onboard attitude estimation to be checked a minumum of once/day.

Before the procedure for validating the onboard attitude estimation process is discussed, the error comparison limits need to be determined. The GRO requirement for attitude determination is an absolute requirement. This absolute requirement references the attitude to the spacecraft body. An attitude sensor alignment can be determined relative to an optical cube on the ground. However once the spacecraft is placed on orbit, this alignment is unknown due to launch shocks. A ground system can align the attitude sensors relative to a reference attitude sensor. The attitude can be determined relative to the reference attitude sensor. The resulting attitude estimation error would be a function of the attitude sensors and attitude determination algorithm accuracies. This "relative" attitude estimation is what will be checked by the ground system since the alignment relative to body necessary for an "absolute" attitude estimate is unknown on the ground as well as onboard.

The portion of the TRW error budget devoted to algorithm errors is 53.3 arcseconds for FHST/FHST case and 66.3 arcseconds for the FHST/FSS case. The FHST and FSS noise allocations in the error budget are 10 and 24.1 arcseconds respectively. After RSSing the appropriate sensor noises with the associated algorithm errors for each case, the following comparison accuracy limits are obtained:

GRO ONBOARD ATTITUDE ESTIMATION ERROR COMPARISON LIMITS	
FHST/FHST	55.1 arcseconds/axis
FHST/FSS	71.2 arcseconds/axis

Figure 6.1

The onboard attitude estimate will be compared to the attitude

estimate from the Code 550 Fine Attitude Determination System (FADS). An outline for the comparison procedure is as follows:

1) Select an orbit of data.
2) Process the sensor data through a batch least squares filter to obtain a ground attitude estimate.
3) Compute errors between the ground attitude estimate and the OBC estimate.
4) Compare these errors to the comparison numbers in Figure 6.1.
5) If the attitude errors are less than those in Figure 6.1, then the OBC attitude estimation function is operating properly, and the validation process is complete.
6) If the attitude errors are greater than those in Figure 6.1, follow steps 7-9 since the the OBC may not be functioning properly.
7) Check to see if the attitude estimate is diverging from the ground solution. If not, then the filter needs tuning.
8) Check to see if any of the following attitude sensors have failed or if the sensor data is degraded:
    a) Gyros
    b) FHSTs
    c) FSS
9) Check to see if the FHST(s) are tracking stars for less than 32.768 seconds. If this is happening, then there could be a star match problem. This could mean a bad onboard covariance (filter tuning probably required), failure of the FHST, or a bad uplink of guide stars. The update filter validates an observed star only if matches with <u>one</u> uplinked guide star. A lack of a guide star match or a match with more than guide star causes the update filter to send a break track command to search for a new star.

At the time these tests are being performed, all update filter data base parameters need to be checked for the previous 24 hour period. Bad data base updates could easily upset onboard attitude estimation.

<u>Contingencies</u>

If the attitude estimation errors equal or exceed comparison limits and the possible problems discussed above have been eliminated, OBC attitude determination calibration may be required. The Onboard filter can be calibrated by the following methods:

a) tuning parameter adjustment
b) changing update frequency of covariance matrix
c) changing measurement frequency

Onboard filter calibration can most easily be performed by tuning parameter adjustment. Short of modifying the Onboard filter, the tuning parameters are the following:

a) Initial Attitude errors along the covariance diagonal
b) Initial Gyro Drift errors along the covariance diagonal
c) FHST noise variance
d) FSS noise variance
e) Gyro white noise estimate.
f) Gyro random walk estimate.

The initial attitude and gyro drift errors in the state covariance matrix occupy the first six diagonal elements. These are data base constants that are used to initialize the state covariance matrix when the filter is initialized. The more accurate these numbers are, the faster the filter will converge to the correct solution. Since the largest value of the upper lefthand 3x3 matrix are used for star selection criteria, a good estimate of the initial attitude and gyro drift errors will decrease the possibility of a misidentification of a star.

The sensor noise variances are used in determining the weight of a particular measurement. This can be seen in the GRO Kalman gain equation:

$$K_{k,i} = P_{k,i}(-)\ H_{k,i}^T\ /\ [\ H_{k,i}\ P_{k,i}(-)\ H_{k,i}^T + R_{k,i}\ ] \qquad (6.1)$$

where $R_{k,i}$ is the sensor noise for a particular measurement. If $R_{k,i}$ is increased, then the Kalman gain K_k will be decreased and more emphasis will be placed on the estimate $X_k(-)$:

$$\hat{\underline{X}}_k(+) = \hat{\underline{X}}_k(-) + K_k\ (\ \underline{Z}_k - H_k\ \hat{\underline{X}}_k(-)\ ) \qquad (6.2)$$

This allows the filter to place emphasis on more accurate measurements. In the FHST/FSS case, the FHST measurements would be given more confidence by the filter than the FSS measurements since FHSTs are more accurate. The estimated dynamic noise for the filter (white noise and random walk) come into play with the propagation of the state covariance matrix. The spectral density matrix is defined as follows:

$$Q(t) = E\ [\ \underline{W}(t)\ \underline{W}^T(t)\ ] \qquad (6.3)$$

where $\underline{W}(t)$ is white noise vector in the state equation (2.3). $Q(t)$ is used to form the discrete dynamics noise covariance matrix as follows:

$$Q_k = \int_{t_{k-1}}^{t_k} \phi(t_k,t')\ Q(t')\ \phi^T(t_k,t')\ dt' \qquad (6.4)$$

which is used in the propagation of the covariance matrix as follows

$$P_k(-) = \phi_k\ P_{k-1}(+)\ \phi_k^T + Q_k \qquad (6.5)$$

Calibration Method	Table Parameters	FDF Control?	Comments
TUNING PARAMETERS			
FHST Noise Var.	Table 37	YES	Useful if FHST data is noisy or to add weight to the FSS measurements
FSS Noise Var.	Table 42	NO	Useful if FSS data is noisy or to add weight to the FHST measurements
Initial attitude errors	Table 42	NO	Useful for the convergence of the filter
Initial gyro drift errors	Table 42	NO	Useful for the convergence of the filter
White noise	Table 36	YES	Needed if gyro white noise increases or if less emphasis on the dynamic model is desired
Random walk	Table 36	YES	Needed if gyro random walk noise increases or if less emphasis on the dynamic model is desired
Related Parameters			
Estimated attitude errors	Table 42	NO	Useful for small onboard attitude quaternion corrections.
Estimated gyro drift errors	Table 42	NO	Needed to insure the onboard attitude is propagated correctly.
Attitude quaternion	Table 59	NO	Needed if onboard quaternion is bad
OTHER METHODS			
Covariance Update Frequency Change	N/A	N/A	Requires code changes in ATTEST and probably the EXEC routine. Not an easy modification.
Measurement Update Frequency Change	N/A	N/A	Requires changes in the EXEC routine. Probable OBC loading problems.

Table 6.1 Update Filter Calibration Method Summary

The progagation of the covariance matrix is used in the computation of the Kalman gain as can be seen in equation (6.1) as well as the star identification process mentioned above. Since the Kalman gain is used to determine the updated state vector (6.2), $P_k(-)$ has potential to undermine the state estimation process if poorly computed.

The second possible calibration method is changing the update frequency of the state covariance matrix (6.5). Currently, the onboard filter updates $P_k(-)$ every 32.768 seconds before the star identification process occurs. Increasing the update frequency could assist in state estimation as well as star identification. The state covariance matrix update is embedded in ATTEST which makes this possibility difficult to implement due to the software mods needed.

The last calibration method mentioned deals with measurement updates. Stellar updates very accurately pin down attitude errors and gyro drift errors. Gyro drift errors affect the attitude between measurement updates. During periods of no measurement updates, the onboard quaternion would be updated in the kinematic integration routine which uses slowly degrading gyro data. The gyro data used requires an accuate estimate of the gyro drift error to compensate the incremental angles. Without accurate gyro drift compensation, the onboard attitude quaternion would slowly diverge from the proper attitude. Now if measurement updates were made more frequently, the attitude would be compensated before it could degrade significantly. This modification would be easier to perform than changing the propagation frequency of the state covariance matrix since little coding changes would be required. The resulting OBC loading would need to be studied to determine whether this modification is viable.

7.0 CONCLUSIONS

For both the FHST/FHST and FHST/FSS configurations, the GRO onboard attitude determination accuracies can be met with significant margins (assuming nominal on-orbit conditions). For the FHST/FHST configuration, the roll, pitch, and yaw predicted attitude estimation errors are 64.11, 64.74, and 60.80 arcseconds respectively. The requirement is 86 arcseconds/axis. For the FHST/FSS configuration, the roll, pitch, and yaw predicted attitude estimation errors are 127.21, 121.30, and 74.65 arcseconds respectively. The requirement is 167.5 arcseconds/axis. For the 1 FHST with 2 guide star case, the attitude estimation errors were 100.34, 100.03, and 64.38 arcseconds respectively.

As expected for the FHST/FHST and FHST/FSS cases, the attitude estimation accuracies are most sensitive to FHST and FSS alignment errors. The gyro drift errors are most sensitive to random walk errors in both cases, but alignment errors provide the second largest component of gyro drift error. These results emphasize the importance of proper attitude sensor alignment determination.

The single FHST case demonstrated that onboard attitude estimation

rivalling that of the FHST/FSS combination is possible. However, this choice is highly sensitive to measurement noise. If the measurement noise is nominal and two guide stars are available, it is recommended to use this case over the FHST/FSS case due to high FSS alignment errors and the FSS FOV limitation on target attitudes.

A procedure was outlined for GRO OBC attitude estimation validation. It was pointed out that the accuracy check is a relative check and not an absolute check. The attitude error comparison for the FHST/FHST case is 55.1 arcseconds/axis and 71.2 arcseconds/axis for the FHST/FSS case.

When and if problems with the onboard attitude estimation process are detected, three onboard filter calibration techniques are available: filter tuning, increased state covariance matrix update frequency, and increased ATTEST frequency. Of the three techniques, filter tuning parameter adjustment is the easiest. In extreme situations where filter turning is not sufficient, an increased measurement frequency for ATTEST would require the least software modifications. OBC loading would need to be studied though since ATTEST is the largest function in the Attitude Control and Determination (ACAD) portion of the OBC flight software.

10.0 REFERENCES

1. Snow, Harman, Garrick, "An Analysis of the Kalman Filter in the Gamma Ray Observatory (GRO) Onboard Attitude Determination Subsystem," NASA/GODDARD SPACE FLIGHT CENTER FLIGHT MECHANICS/ ESTIMATION SYMPOSIUM 1988, May 10-11.
2. Wertz, James, edit., SPACECRAFT ATTITUDE DETERMINATION AND CONTROL, D. Reidel Publishing Company, 1985.
3. TRW, 40420-86-322-001, GRO OPERATIONS DATA BOOK PART 3 - ACAD SUBSYSTEM VOLUME II, 2 Sept. 1986.
4. Kulp, D., "IRU calibration Analysis", SEAS QUICK NOTE NO.: EUVE-413-90003, October 27, 1989.
5. Teledyne Systems Company, "NASA Standard High Performance Inertial Reference Unit DRIRU II", January 1988.
6. Jayaraman, C. and Braunstein, K., "Explorer Platform Star Tracker Calibration", Fairchild Interoffice Communication, ACS:EP:89:080, July 19, 1989.
7. Torgow, Joan S., USERS' GUIDE TO THE STANDARD STAR Configuration, Performance, and Use, TM86-02, Ball Aerospace Division, April 1986.
8. TRW, EQ7-145, EQUIPMENT SPECIFICATION FINE SUN SENSOR ASSEMBLY, 22 August 1983.
9. Convesation with F. Landis Markley, NASA/GSFC Code 712.1, December 13, 1989.
10. Harman, R, GRO MISSION FLIGHT DYNAMICS ANALYSIS REPORT: Onboard Computer (OBC) Attitude Error Analysis (Analysis Item No. A1.2), Mission Report 90004N.

TWO AXIS POINTING SYSTEM (TAPS)
Attitude Acquisition, Determination, and Control

John D. Azzolini
David E. McGlew

Guidance and Control Branch
Space Technology Division
Engineering Directorate
Goddard Space Flight Center

ABSTRACT

The TAPS is a 2 axis gimbal system designed to provide fine pointing of STS borne instruments. It features center-of-mass instrument mounting and will accommodate instruments of up to 1134 kg (2500 pounds) which fit within a 1.0 by 1.0 by 4.2 meter (40 by 40 by 166 inch) envelope. The TAPS system is controlled by a microcomputer based Control Electronics Assembly (CEA), a Power Distribution Unit (PDU), and a Servo Control Unit (SCU). A DRIRU-II inertial reference unit is used to provide incremental angles for attitude propagation. A Ball Brothers STRAP star tracker is used for attitude acquisition and update.

The theory of the TAPS attitude determination and error computation for the Broad Band X-ray Telescope (BBXRT) are described. The attitude acquisition is based upon a 2 star geometric solution. The acquisition theory and quaternion algebra are presented. The attitude control combines classical position, integral and derivative (PID) control with techniques to compensate for coulomb friction (bias torque) and the cable harness crossing the gimbals (spring torque). Also presented is a technique for an adaptive bias torque compensation which adjusts to an ever changing frictional torque environment. The control stability margins are detailed, with the predicted pointing performance, based upon simulation studies, presented. Finally the TAPS user interface, which provides high level operations commands to facilitate science observations, is outlined.

NOTATION AND NOMENCLATURE

Matrices will be written as bold faced capital letters, e.g. **A**. Vectors will be written as a lower case letter e.g. $\vec{v}$. The vector inner product will be written as $\langle \vec{a}, \vec{b} \rangle$ (i.e. $\vec{a}$ *dot* $\vec{b}$). Quaternions will be written as lower case barred letters, e.g. $\bar{q}$. A quaternion is used to represent a the TAPS coordinate system with respect to an inertial coordinate system. It may be expressed as a vector part and a scaler part. If we define a rotation about a unit vector, $\hat{\epsilon}$, of angle ϕ, it may be written as:

$$\bar{q} = (\sin\frac{\Phi}{2}\,\hat{\varepsilon}, \cos\frac{\Phi}{2})$$

The inverse rotation is denoted $\bar{q}'$ and is formed by negating the vector part of $\bar{q}$. For every attitude quaternion, $\bar{q}$, there is a unique equivalent attitude direction cosine matrix, **A**. For our purposes, $\bar{q}$ and **A** represent inertial to TAPS coordinate transformations. The column vectors of **A** are the inertial basis vectors expressed in TAPS coordinates. It follows that the row vectors of **A** are the TAPS basis vectors expressed in inertial coordinates.

ATTITUDE DETERMINATION

When the TAPS is in parked or stowed position, its coordinate system is defined to be nominally parallel to the STS coordinate system. The x axis is the shuttle roll axis, positive forward. The z axis points downward through the floor of the orbiter bay. The y axis points to starboard, forming a right handed cartesian frame. The inertial reference frame is the geocentric inertial coordinate system, 1950 mean (M50). The components of the TAPS and its coordinate system are shown in figure 1.

The attitude quaternion is initialized from the STS state vector. An attitude acquisition sequence is performed to remove the misalignments and uncertainties in the STS state vector. This sequence is discussed in the section on attitude acquisition.

The quaternion is then propagated using a closed form solution to the kinematic equations of motion.

$$\bar{q}(t+T) = e^{\frac{\Omega T}{2}}\bar{q}(t) \qquad (1)$$

where, T is the propagation sample period, and Ω is the skew symmetric matrix,

$$\Omega = \begin{bmatrix} 0 & \omega_z & -\omega_y & \omega_x \\ -\omega_z & 0 & \omega_x & \omega_y \\ \omega_y & -\omega_x & 0 & \omega_z \\ -\omega_x & -\omega_y & -\omega_z & 0 \end{bmatrix}$$

This solution assumes that the rate vector over the propagation interval is constant. By using the average rate over this interval, this approximation is close to ideal. The equivalent attitude direction cosine matrix, **A**, is computed each control cycle (16 Hz). The input to these computations are the scaled, drift corrected gyro rates.

ATTITUDE ERROR COMPUTATION

The attitude error is computed using the attitude direction cosine matrix, $\mathbf{A}$, and the desired inertial target unit vector, $\hat{s}$. The target vector is transformed into TAPS coordinates.

$$\hat{s}_{TAPS} = \mathbf{A}\hat{s} \tag{2}$$

The required rotation vector is computed by taking the cross product:

$$\vec{e} = \hat{s}_{TAPS} \times \hat{b} \tag{3}$$

where $\hat{b}$ is the reference unit vector in the TAPS coordinate frame, (the vector we are trying to point). In normal operations, $\hat{b} = \hat{b}_i$ the BBXRT instrument boresight unit vector. During acquisition, $\hat{b} = \hat{b}_s$ the star tracker boresight unit vector. In all cases $\hat{b}$ is nominally aligned with the $-z_{TAPS}$ axis.

The magnitude of $\vec{e}$ is given by:

$$\|\vec{e}\| = \|\hat{s}_{TAPS}\| \|\hat{b}\| \sin\theta = \sin\theta$$

where θ is the angle between $\hat{s}_{TAPS}$ and $\hat{b}$. We want the projection of θ on the x and y axes. One way to do this is to normalize $\vec{e}$ as:

$$\hat{e} = \frac{\vec{e}}{\|\vec{e}\|} = (e'_x, e'_y, e'_z)$$

Now, $\hat{e}$ is the unit rotation vector. The projection of θ on the x axis is then:

$$x_e = \sin^{-1}\|\vec{e}\| \, e'_x = \frac{\theta \, e_x}{\sin\theta}$$

and for small θ, $\sin\theta \sim \theta$ and:

$$x_e \sim e_x \tag{4}$$

similarly,

$$y_e \sim e_y \tag{5}$$

For TAPS, θ can be as large as 56 degrees. However, the control law limits the position error to a much smaller value which allows us to use this approximation. This is discussed in the attitude control section.

ATTITUDE ACQUISITION

The attitude acquisition sequence is designed to eliminate the alignment and other errors inherent in using the STS state vector to initialize the TAPS attitude quaternion. The sequence assumes that the TAPS has been deployed to parked index position, nominally along the shuttle -z axis, and that the shuttle state vector has been used to initialize the attitude quaternion. From that point the following steps are taken.

For each pair of stars:

1) Slew on gyros to the first star.
2) Store attitude information $\bar{q}_1$, $\hat{z}_1$, where $\hat{z}_1$ is the inertial TAPS z axis vector (the third row of the **A** matrix).
3) Null on the star using star tracker position.
4) Replace the attitude quaternion by the stored quaternion.
5) Slew on gyros to the second star.
6) Store attitude information $\bar{q}_3$, $\hat{z}_3$.
7) Null on the star using star tracker position.
8) Store attitude information $\bar{q}_4$, $\hat{z}_4$.
9) Compute updated attitude quaternion from the stored information.

If there are more than two stars, we may repeat the sequence using subsequent pairs. In this case, we repeat steps 1) through 4) for the last star to remove residual roll and pitch errors caused by numerical computational errors.

The mathematics of this acquisition sequence will now be explored. This analysis assumes perfect control, alignment ($\hat{b} = -\hat{z}_{TAPS}$), and sensors. We will use the subscript, t, to represent true attitude.

At the start of the sequence,

Let $\bar{q}_{t0}$ be the true attitude and let $\bar{q}_0$ be our on-board attitude, such that

$$\bar{q}_{t0} = \bar{q}_0 \, \bar{q}_{e0}$$

where $\bar{q}_{e0}$ is the error in initializing $\bar{q}_0$. The slew of step 1) of our attitude sequence defines an attitude rotation, $\bar{q}_{s1}$, after which,

$$\bar{q}_1 = \bar{q}_0 \, \bar{q}_{s1}$$

and,

$$\bar{q}_{t1} = \bar{q}_{t0} \, \bar{q}_{s1} = \bar{q}_0 \, \bar{q}_{e0} \, \bar{q}_{s1} = \bar{q}_1 \, \bar{q}_{e1}$$

The values of $\bar{q}_1$ and $\hat{z}_1$ are stored in step 2).

It is important to note that the first slew results in,

$$\vec{e} = 0 \Rightarrow \mathbf{A}\,\hat{s} \times \hat{b} = 0 \Rightarrow$$

$$\mathbf{A}(\bar{q}_1)\,\hat{s} = -\hat{z}_{TAPS} \qquad (6)$$

Step 3) of our sequence produces the rotation $\bar{q}_{n2}$, after which,

$$\bar{q}_2 = \bar{q}_1 \, \bar{q}_{n2}$$

and,

$$\bar{q}_{t2} = \bar{q}_{t1} \, \bar{q}_{n2} = \bar{q}_1 \, \bar{q}_{e1} \, \bar{q}_{n2} = \bar{q}_2 \, \bar{q}_{e2}$$

Again note that the star tracker null results in,

$$\mathbf{A}(\bar{q}_{t2})\ \hat{s} = -\hat{z}_{TAPS} \tag{7}$$

From equations (6) and (7),

$$\mathbf{A}(\bar{q}_1)\ \hat{s} = \mathbf{A}(\bar{q}_{t2})\ \hat{s}$$

so,

$$\hat{s} = \mathbf{A}'(\bar{q}_1)\ \mathbf{A}(\bar{q}_{t2})\ \hat{s} = \mathbf{A}(\bar{q}_{t2}\,\bar{q}_1')\ \hat{s}$$

and,

$$\bar{q}_{t2}\,\bar{q}_1' = \bar{q}_s = \bar{q}_{-z}$$

where $\bar{q}_{-z}$ is some rotation about the $\hat{s}$ vector or, equivalently, the $-\hat{z}_{TAPS}$ vector expressed in inertial coordinates.

We can thus write,

$$\bar{q}_{t2} = \bar{q}_{-z}\,\bar{q}_1$$

Step 4) replaces the on-board quaternion $\bar{q}_2$ by the stored quaternion $\bar{q}_1$.

Proceeding with step 5),

$$\bar{q}_{t3} = \bar{q}_{t2}\,\bar{q}_{s3} = \bar{q}_{-z}\,\bar{q}_1\bar{q}_{s3} = \bar{q}_{-z}\,\bar{q}_3$$

The quaternion, $\bar{q}_3$, and $\hat{z}_3$ are stored in step 6).

Step 7) yields,

$$\bar{q}_{t4} = \bar{q}_{t3}\,\bar{q}_{n4} = \bar{q}_{-z}\,\bar{q}_3\,\bar{q}_{n4} = \bar{q}_{-z}\,\bar{q}_4$$

In step 8) we store $\bar{q}_4$ and $\hat{z}_4$.

We can compute the correction quaternion, $\bar{q}_{-z}$, from the stored attitude information at the three corners of the spherical triangle defined by our sequence. Figure 2 shows the acquisition sequence. The angle Γ is the gyro slew from the star 1 to star 2 and the angle Γ' is the angle between $\hat{z}_1$ and $\hat{z}_4$. The angle β is the slew produced by nulling on star 2. The angle α is the error about $\hat{z}_1$ in $\bar{q}_1$.

The quaternion $\bar{q}_{-z}$ is computed as follows. From figure 2 the angle α is given by:

$$\cos\alpha = \frac{\cos\beta - \cos\Gamma\cos\Gamma'}{\sin\Gamma\sin\Gamma'}$$

By assuming that $\Gamma \approx \Gamma'$ we can make the simplification,

$$\cos\alpha \approx \frac{\cos\beta - \cos^2\Gamma}{\sin^2\Gamma}$$

where,

$$\cos\Gamma = \langle \hat{z}_1, \hat{z}_3 \rangle$$

and,

$$\cos\beta = \langle \hat{z}_3, \hat{z}_4 \rangle$$

We compute the quaternion $\bar{q}_{-z}$ as:

$$\bar{q}_{-z} = (\pm\sin\frac{\alpha}{2}\,\hat{z}_1, \cos\frac{\alpha}{2}) \qquad (8)$$

where,

$$\cos\frac{\alpha}{2} = \sqrt{\frac{1+\cos\alpha}{2}}$$

and,

$$\sin\frac{\alpha}{2} = \sqrt{\frac{1-\cos\alpha}{2}}$$

The sign ambiguity in (8) is resolved by setting the sign equal to:

$$sgn\{\langle \hat{z}_1, (\hat{z}_4 - \hat{z}_1)\times(\hat{z}_3 - \hat{z}_1)\rangle\} = sgn\{\langle \hat{z}_1, \hat{z}_4\times\hat{z}_3\rangle\}$$

Note that the order of the cross product is chosen to provide the correct sign, since we are pointing the -z axis but storing the +z axis for use in these computations. The on-board attitude quaternion is then computed as:

$$\bar{q} = \bar{q}_{-z}\,\bar{q}_4 \qquad (9)$$

This computation eliminates the z axis error.

The effects of misalignments, gyro and star tracker quantization and noise, control errors, and STS inertial hold limit cycle motion have all been evaluated in simulation studies. These studies show that the worst case attitude acquisition error is approximately 2 arcminutes.

ATTITUDE CONTROL -- BASIC

In the x axis, the classic second order control law is,

$$T_x = -(K_p x_e + K_r \omega_x) \qquad (10)$$

where T_x is the x axis gimbal motor torque, and ω_x is the drift corrected, measured gyro rate about the x axis. Here, K_p is the position gain and K_r is the rate gain. They are chosen to achieve the desired control bandwidth, ω, and damping, ζ. The relations are:

$$\omega = \sqrt{\frac{K_p}{I_{xx}}}$$

and,

$$\zeta = \frac{K_r}{(2\omega I_{xx})}$$

where I_{xx} is the inertia about the x axis.

For large rotations, we wish to limit the rate, ω_x, to 1 degree/sec ($\frac{\pi}{180}$ rad/sec). From equation (10) notice that when

$$x_e = -\frac{K_r \omega_x}{K_p},$$

$$T_x = 0$$

We can exploit this by limiting the magnitude of x_e as

$$\tilde{x}_e = sgn[min(x_{lim}, |x_e|), x_e]$$

where,

$$x_{lim} = \frac{K_r}{K_p}\frac{\pi}{180}$$

For typical values of K_r and K_p, x_{lim} is less than 1 degree and our small angle approximation is acceptable.

When we are in the normal pointing mode, and the position and rate errors are small, we add an integral control term to compensate for torque hangoff effects. The TAPS basic control law is then,

$$T_x = -(K_p \tilde{x}_e + K_r \omega_x + K_i \tilde{x}_\Sigma) \qquad (11)$$

where,

$$\tilde{x}_\Sigma = \sum \tilde{x}_e$$

Note that K_i must be adjusted from the continuous case by multiplying by the control interval, T, in seconds. This is because of the way we are approximating the integral of position.

$$\int_0^{nT} \tilde{x}_e(\tau)\, d\tau = \sum_{i=1}^{n} \int_{(i-1)T}^{iT} \tilde{x}_e(\tau)\, d\tau$$

By the mean value theorem,

$$= T\sum_{i=1}^{n} \tilde{x}_e(\xi_i) \approx T\sum_{i=1}^{n} \tilde{x}_e(iT) = T\tilde{x}_\Sigma$$

The development for the y axis follows the above exactly.

In order to allow slewing and to avoid stability problems, the integral compensation must only be added when we are truly holding. To accomplish this we compute the boolean variable:

$$HOLD := (|\tilde{x}_e| < p_{hold}) \wedge (|\tilde{y}_e| < p_{hold}) \wedge (|\omega_x| < \omega_{hold}) \wedge (|\omega_y| < \omega_{hold})$$

where, p_{hold} and ω_{hold} are the position and rate values we choose to consider the limits of hold mode pointing. When HOLD is true, we add in the integral compensation. When HOLD is false, we set $\tilde{x}_\Sigma$ and $\tilde{y}_\Sigma$ to zero and do not add the integral term.

During attitude acquisition and update, we must be able to control the TAPS pointing based upon the star tracker error signals. To accomplish this x_e and y_e in equations (4) and (5) are replaced by x_t and y_t, the star tracker error signals. The rest of the control law is unchanged.

The basic control law is also modified to limit the range of motion. We must avoid gimbal contact with limit stops. We achieve this by providing a software limit of 19 degrees from the gimbal index position. When the position error exceeds 19 degrees minus our current gimbal encode position, in the direction we are moving, we replace the position error by this difference. For the x axis the logic is,

```
if g_rate_x > 0 then
  if (g_limit - g_position_x) < x_e then
    x_e := g_limit - g_position_x
else
  if (-g_limit - g_position_x) > x_e then
    x_e := -g_limit - g_position_x
```

where g_rate_x, g_limit, and $g\_position_x$ are the x axis tachometer rate, the gimbal software limit, and the x axis encoder angle respectively.

ATTITUDE CONTROL -- FRICTION AND SPRING TORQUE COMPENSATION

The modelling and early operation of the TAPS gimbal provided insight into the frictional torque characteristics of the gimbal. The gimbal physically exhibits friction, which may be approximated by a coulomb friction model, and a spring torque, due to the bending of the cable harnesses, which may be approximated by an angular hook's law spring.

Early simulation studies demonstrated that we could not maintain the required 0.5 arcminute pointing jitter, during vernier thruster firings, with the basic control law alone. In order to improve this transient performance, 3 additional compensation terms are optionally added to the basic control law. These terms were a constant torque bias in the direction of motion, a spring torque term proportional to the gimbal encoder angle, and an adaptive bias torque term to correct for errors in the modelling and parameters of the other two terms.

The bias torque term is given by,

$$T_{xbias} = sgn(K_{xbias}, g\_rate_x) \quad (12)$$

The spring torque term is given by,

$$T_{xspring} = K_{xspring}\ (g\_position_x - g\_zero_x) \quad (13)$$

where g_zero_x is the encoder reference point for zero spring torque.

The adaptive bias torque term is computed as follows,

$$T_{xadap_{i+1}} = K_{adap}\, T_{xadap_i} + K_\tau\, T_{xbasic_i} \quad (14)$$

which is a first order difference equation in T_{xadap}. In this equation, K_{adap} is chosen to be less than one for stability, and K_τ if chosen as the inverse of the desired time constant to reach steady state. The torque, T_{xbasic_i}, is the torque computed by the basic TAPS control law. A separate adaptive term is computed for each direction of motion, for each axis. If we assume a constant motion in one direction,

$$T_{xbasic_i} = T_{friction} - T_{adap_i}$$

where $T_{friction}$ is a constant frictional torque to be overcome. The steady state solution of equation (13) is then given by,

$$T_{xadap_{ss}} = K_f\, T_{friction}$$

where,

$$K_f = \frac{K_\tau}{1 - K_{adap} + K_\tau}$$

By proper specification of the gains, K_f, the fraction of the frictional torque which will be removed by the adaptive torque term, can be selected.

The final control law is then given by,

$$T_x = -(K_p x_e + K_r \omega_x + K_i x_\Sigma) + T_{xbias} + T_{xspring} + T_{xadaptive} \quad (15)$$

ATTITUDE CONTROL -- ANALYSIS

Both the roll and pitch control loops are analyzed with appropriate rate, position and integral gains selected to provide at least 6 db of gain margin and approximately 30 degrees of phase margin. The analysis was performed using the Interactive Controls Analysis (INCA) program. The Nyquist frequency responses of the system are shown in figure 3 and 4. From the plots, the gain and phase margins are,

Roll Axis

Upper Gain Margin:	7.89 db at 11.7 radians/second
Lower Gain Margin:	11.5 db at 1.83 radians/second
Phase Margin:	28 degrees at 5.05 radians/second

Pitch Axis

Upper Gain Margin: 7.76 db at 11.7 radians/second
Lower Gain Margin: 11.6 db at 1.84 radians/second
Phase Margin: 28 degrees at 5.12 radians/second

Once the gains were selected, planer simulations representing the shuttle and experiment interconnected via the TAPS were developed. The desire was to show that the pointing stability requirements of maintaining a peak pointing error of less than 0.5 arcminutes could be achieved in the presence of vernier jet firings.

The MODEL translator was used to generate the FORTRAN rigid body simulations to evaluate the performance of the control laws developed in the above analysis. The x-z planer plant model equations, cast into matrix form, is given by,

$$M\frac{d\vec{\dot{X}}}{dt} = \vec{T}$$

In this equation,

$$\vec{X} = \begin{bmatrix} X \\ Z \\ \theta_o \\ \theta_g \end{bmatrix}$$

where,

X ≙ x axis translational coordinate of shuttle motion
Z ≙ z axis translational coordinate of shuttle motion
θ_o ≙ Shuttle y axis rotational coordinate
θ_g ≙ Instrument y axis rotational coordinate

and,

$$\vec{T} = \begin{bmatrix} F_X \\ F_Z \\ T_o \\ T_g \end{bmatrix}$$

where,

F_X ≙ x axis component of vernier thrust
F_Z ≙ z axis component of vernier thrust
T_o ≙ Shuttle y axis torque
T_g ≙ Instrument y axis torque

and,

$$M = \begin{bmatrix} M_o+M_g & 0 & M_g r_{cm}\cos\theta_{o_1} & M_g r_g \cos\theta_{g_1} \\ 0 & M_o+M_g & -M_g r_{cm}\sin\theta_{o_1} & -M_g r_g \sin\theta_{g_1} \\ M_g r_{cm}\cos\theta_{o_1} & -M_g r_{cm}\sin\theta_{o_1} & M_g r_{cm}^2+I_o & M_g r_{cm} r_g(\cos\theta_{o_1}\cos\theta_{g_1}+\sin\theta_{o_1}\sin\theta_{g_1}) \\ M_g r_g \cos\theta_{g_1} & -M_g r_g \sin\theta_{g_1} & M_g r_{cm} r_g(\cos\theta_{o_1}\cos\theta_{g_1}+\sin\theta_{o_1}\sin\theta_{g_1}) & M_g r_g^2+I_g \end{bmatrix}$$

where,

M_o ≙ Orbiter mass

M_g ≙ Gimbal and instrument mass
I_o ≙ Orbiter moment of inertia
I_g ≙ Gimbal moment of inertia
r_{cm} ≙ Distance from orbiter center of mass to gimbal axis
r_g ≙ Distance from gimbal axis to center of mass of load
θ_{o_i} ≙ Initial shuttle y axis angle
θ_{g_i} ≙ Initial instrument y axis angle

A similar set was developed for the Y-Z plane.

For the simulation, a simple orbiter limit cycle was developed to illustrate the performance of the controller when the relative rate between orbiter and experiment changed sign.

The initial conditions were established so that the orbiter was in a limit cycle at a rate of 0.02 degrees/sec with a dead-zone of ± 0.5 degrees. The experiment was held inertially fixed while the orbiter was rotated underneath it. As can be observed in Figure 5, the experiment position error (top channel) exceeded 4 arcminutes, which is well beyond the 0.5 arcminutes pointing requirement, during the vernier jet firings.

The same initial conditions as above were established for the results shown in Figure 6 with the adaptive bias logic enabled. Initially the position error peaks well beyond the 0.5 arcminutes pointing requirement whenever a vernier jet fires. However, after several jet firings the bias term has adjusted to the point where the friction level is compensated by the bias term and not the integrator.

Illustrated in Figure 7 is the effect of a bias level nearly twice as high as required. The result is a temporary oscillation of the position error at the very beginning of the simulation run. The bias term being too high overdrives the experiment resulting in a switching between the positive and negative bias levels until they are adjusted to the appropriate value.

The curves shown in Figure 8 illustrate the adaptive changes that are occurring in the positive and negative bias terms. Since the bias levels were selected to be nearly twice as high as required, the positive adaptive term must adjust to decrease the positive term. Similarly the negative bias term must adjust to decrease the negative bias. In order to avoid this temporary oscillation an initial bias level lower than the anticipated friction should be selected and allowed to adjust upward.

Ideally the adaptive bias term is used to generate a torque that is equal to the friction level opposing the motion. Any torque opposing the motion that appears as a spring contaminates the bias capability. If such a spring torque exists in orbit, it will be possible to compensate for this torque by using the spring torque

control term. This will maintain the capability of the adaptive bias to effectively handle the friction level in the bearings.

In the above linear analysis it has been shown that the basic control law is stable with adequate margins while the simulation has shown that the pointing requirement of 0.5 arcminutes can be met with the additions to the basic control law.

TAPS USER INTERFACE

One of the design goals of the TAPS system was to provide science experimenters a user interface which allows them to control the pointing of their instrument. The TAPS Ground Support Equipment (TGSE) provides a complete command and telemetry interface to the user. Commands may be composed and processed in real time, processed from a command disk file, or stored in on-board command timelines for timed, sequenced execution. Telemetry is captured, processed, and displayed in real time.

The commands, called TAPS Mission Operations Commands (TMOCs), are of high level with a descriptive, English-like syntax. A sampling of commands which are of interest to an experimenter illustrates this.

Perform and inertial slew to a given M50 target vector:

```
OPER  ISLEW  S1,S2,S3
```

Perform an inertial slew to a catalogued target vector:

```
OPER  CSLEW    CAT_NO
```

Note that experimenters need not be concerned with slew sequences about gimbal axes to acquire a target. Rather, they may specify targets in inertial coordinates. The TAPS takes care of the rest.

CONCLUSIONS

The TAPS has been designed to provide pointing of the BBXRT instrument to an absolute accuracy of better than 4 arcminutes, with a pointing stability of better than 0.5 arcminutes. The design of the attitude error computation provides for simplified user control of observations. The control law provides rate limiting to avoid gyro saturation and position limiting to keep the gimbals within an acceptable range of motion.

It is our belief that the TAPS provides a flexible pointing capability which may be used for a variety of instruments. The system software has been designed in a highly modular fashion to allow the TAPS to accommodate other instruments and mission profiles. In particular, the architectural and algorithmic design required to point the Wide Angle Michelson Doppler Imaging Interferometer (WAMDII) instrument, which is an Earth limb pointer, had been completed prior to the suspension of the WAMDII program.

ACKNOWLEDGEMENTS

This paper provides a description of the TAPS system for the BBXRT-ASTRO mission. Much of the theory used to develop this system may be found, or derived from material, within Wertz (1). We have found this reference to be extremely valuable for any development work in this field and recommend it highly.

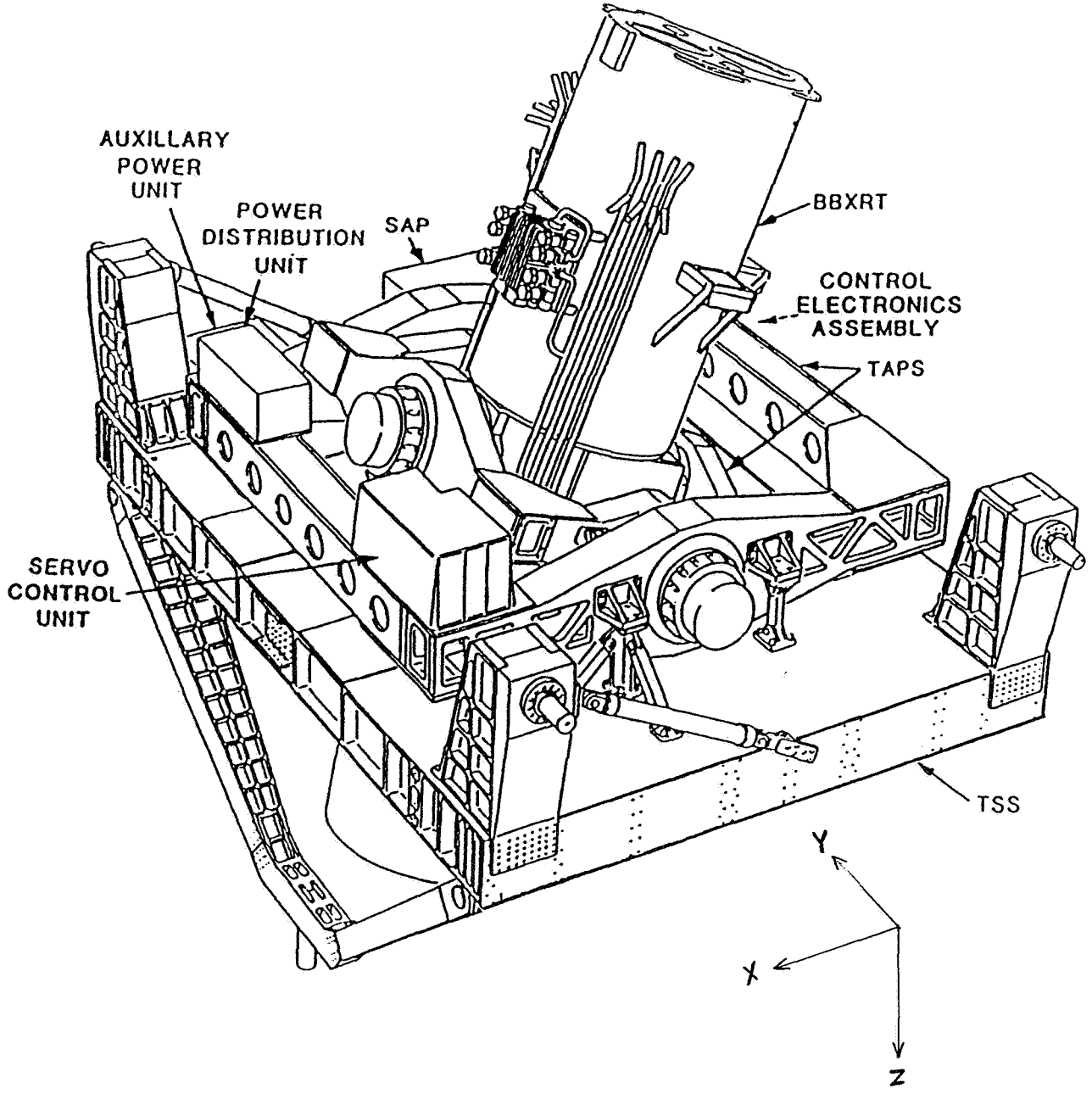
The authors gratefully acknowledge the members of the Guidance and Control Branch for helpful discussions and independent verification of the work described. In particular, we recognize Thomas Flatley for his MACSYMA based, independent verification of the attitude acquisition algorithm, and F. Landis Markley and Michael Femiano for their critical review of this paper and helpful suggestions.

REFERENCES

(1) Wertz, James R. Ed., Spacecraft Attitude Determination and Control, Holland: D. Reidel Publishing Company, 1980

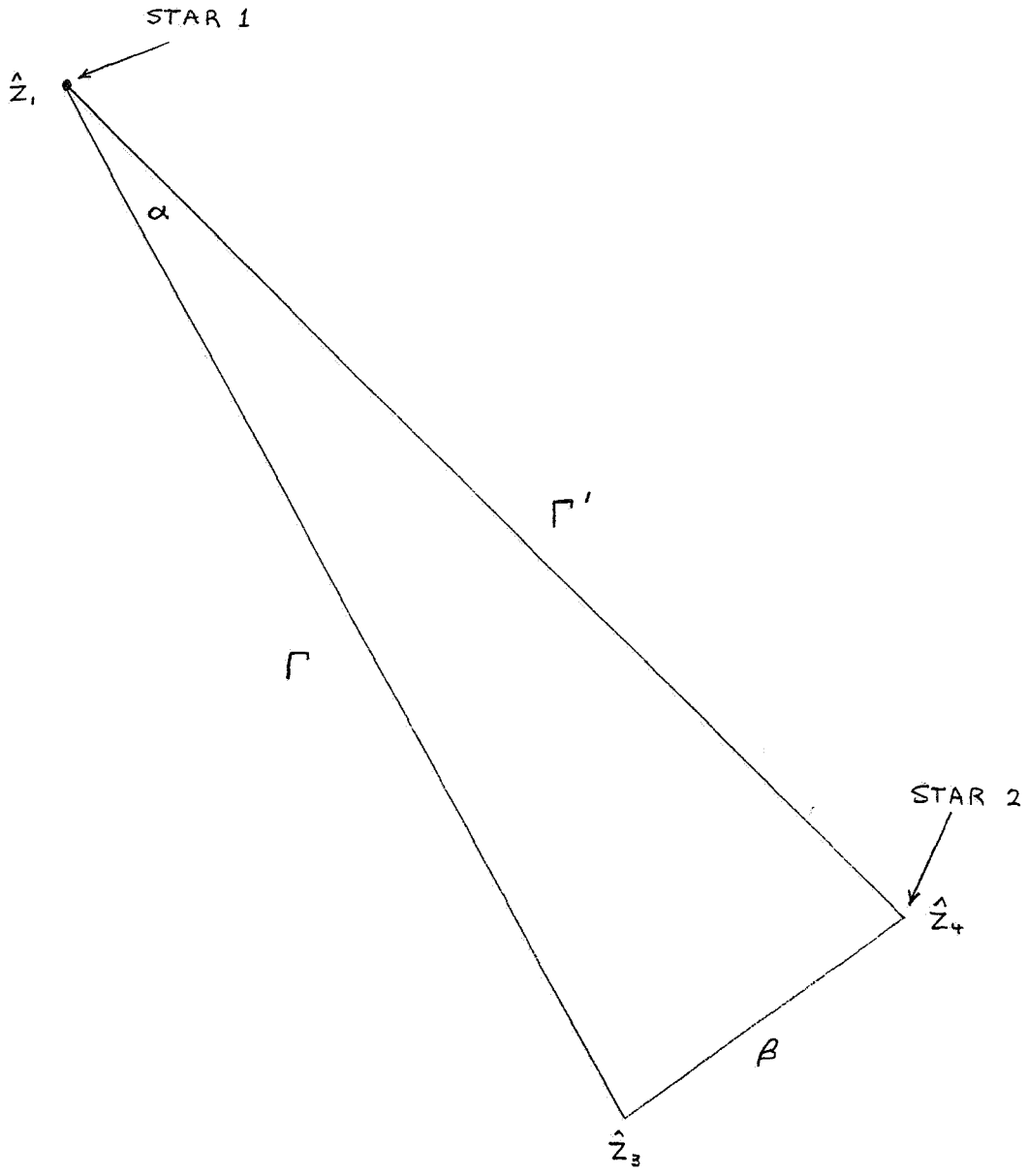
(2) Memorandum, John Azzolini and Frederick Hager to Walter Nagel, "TAPS Baseline Attitude Determination, Control, and Acquisition for the BBXRT SHEAL-II Mission, 1986

(3) Memorandum, David McGlew to John Azzolini, "TAPS-BBXRT Controller Gains and Adaptive Feedback",1989



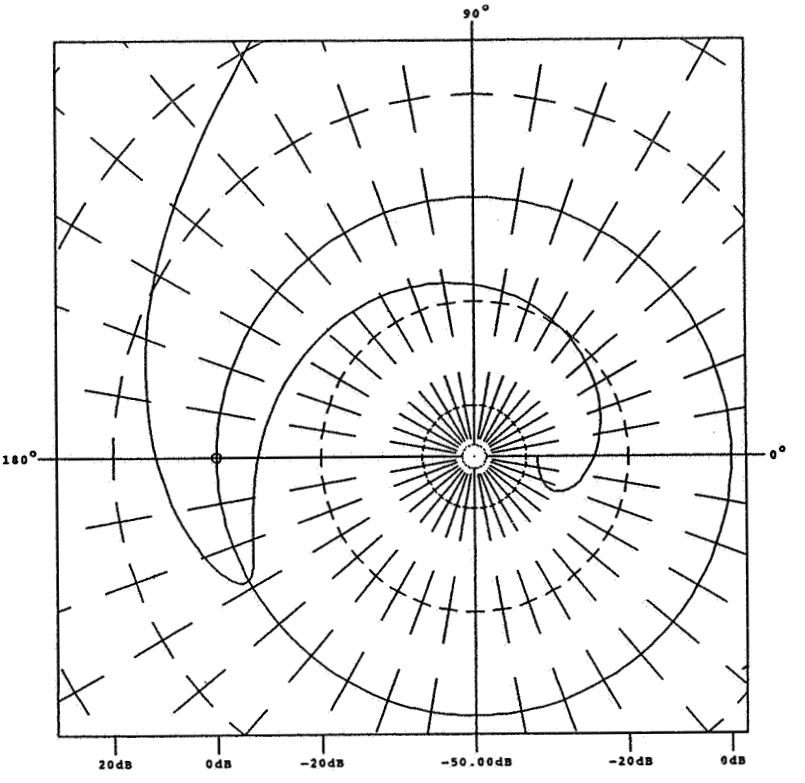
TWO AXIS POINTING SYSTEM
BROAD BAND X-RAY TELESCOPE

FIGURE 1

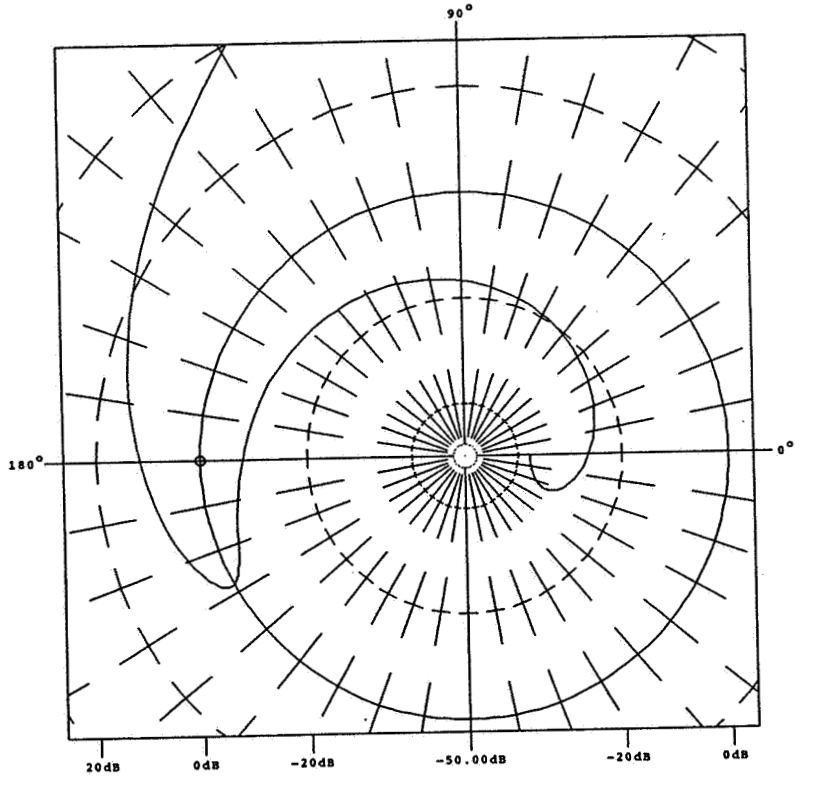


ATTITUDE ACQUISITION GEOMETRY

FIGURE 2



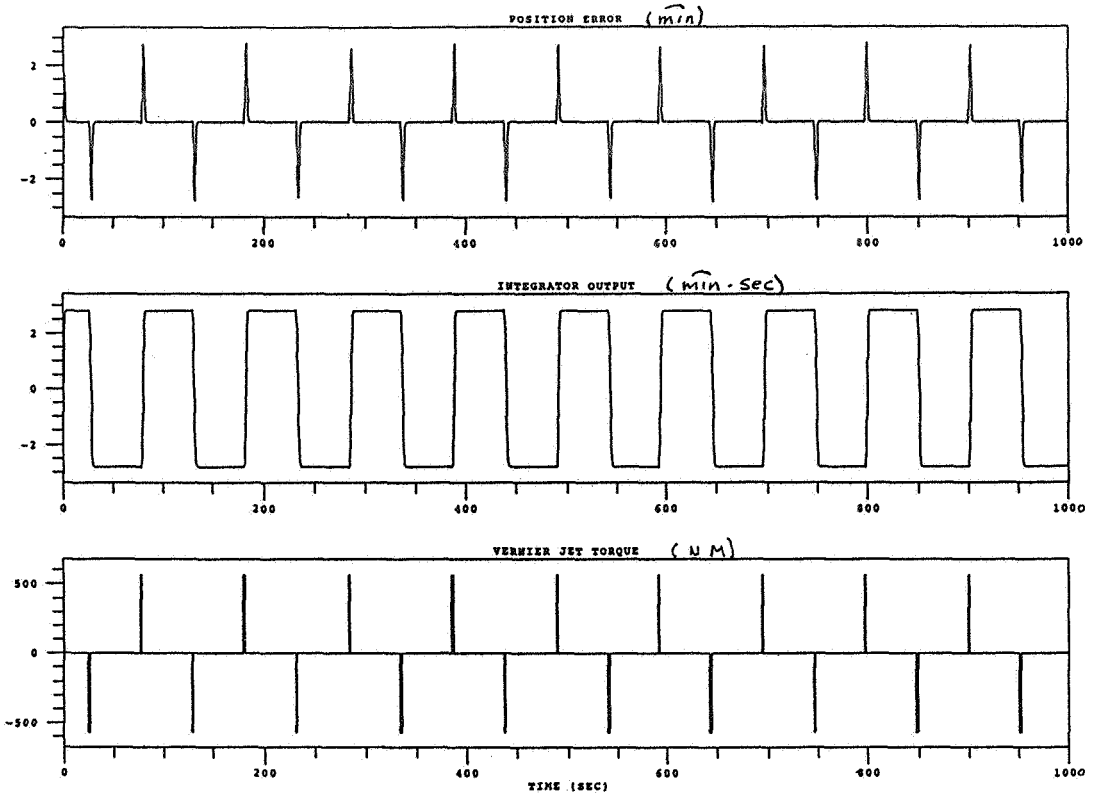
ROLL AXIS



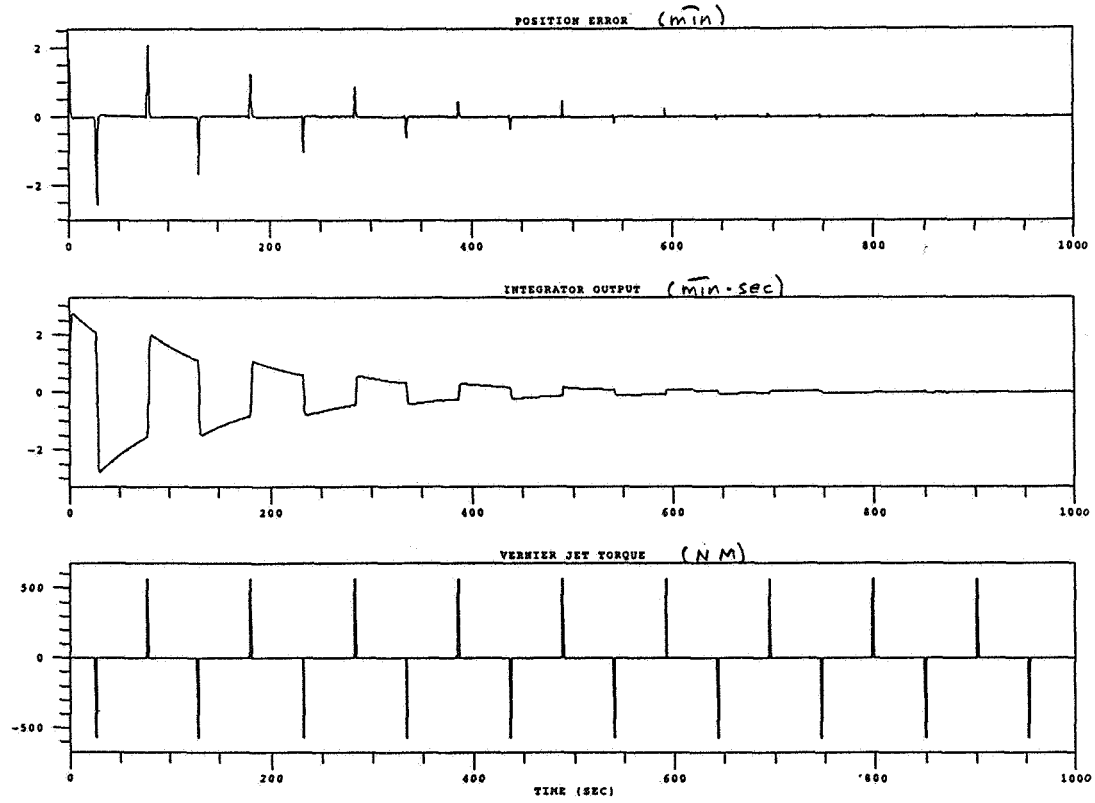
PITCH AXIS

NYQUIST STABILITY DIAGRAMS

FIGURES 3 & 4



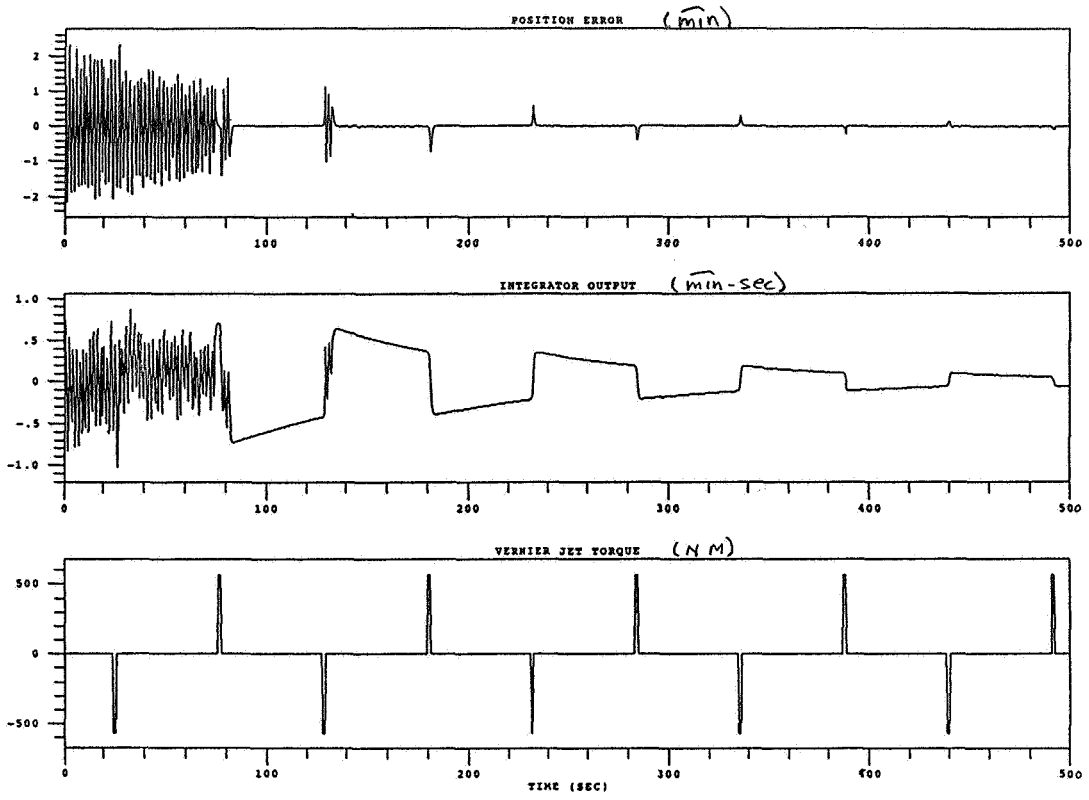
WITHOUT ADAPTIVE BIAS



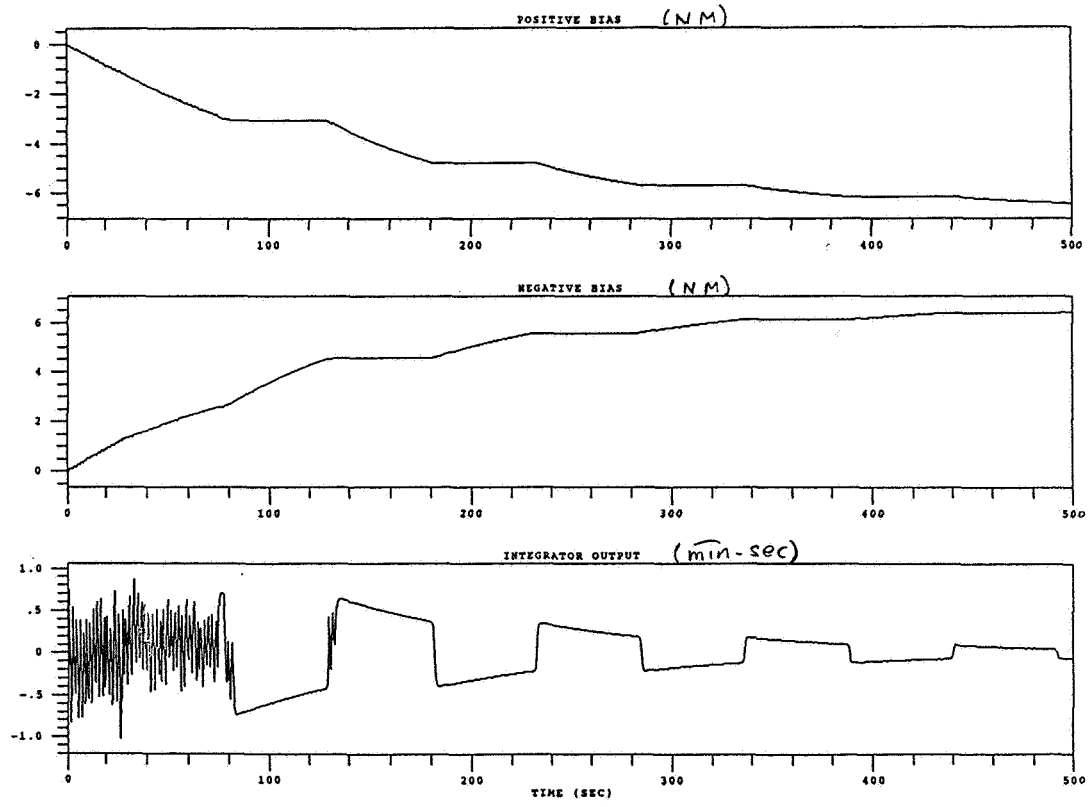
WITH ADAPTIVE BIAS

POINTING PERFORMANCE

FIGURES 5 & 6



POINTING PERFORMANCE



ADAPTIVE BIAS CONVERGENCE

EFFECTS OF INITIAL LARGE BIAS

FIGURES 7 & 8

A METHOD OF DETERMINING ATTITUDE FROM MAGNETOMETER DATA ONLY*

G. A. Natanson, S. F. McLaughlin, R. C. Nicklas
Computer Sciences Corporation

ABSTRACT

The paper presents a new algorithm to determine the attitude using only magnetometer data under the following conditions: (1) internal torques are known and (2) external torques are negligible. Torque-free rotation of a spacecraft in thruster firing acquisition phase and its magnetic despin in the B-dot mode give typical examples of such situations. A simple analytical formula has been derived in the limiting case of a spacecraft rotating with constant angular velocity. The formula has been tested using low-frequency telemetry data for the Earth Radiation Budget Satellite (ERBS) under normal conditions. Observed small oscillations of body-fixed components of the angular velocity vector near their mean values result in relatively minor errors of approximately 5 degrees. More significant errors come from processing digital magnetometer data. Higher resolution of digitized magnetometer measurements would significantly improve the accuracy of this deterministic scheme. Tests of the general version of the developed algorithm for a free-rotating spacecraft and for the B-dot mode are in progress.

*This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, under Contract NAS 5-31500.

1. INTRODUCTION

The idea of developing an attitude determination system using only three-axis magnetometer measurements has been attracting attention for many years, despite its relatively low accuracy. The light weight and low cost of such a system are usually considered its main advantages. For a spacecraft in low-attitude Earth orbit, Kalman filtering has been proven to be an effective tool to derive the attitude from magnetometer measurements with a 2-degree (deg) accuracy (see References 1 and 2).

This paper is intended to develop an attitude determination algorithm using only magnetometer measurements under contingency conditions such as loss of attitude control of spacecraft. Due to high-speed rotation of a spacecraft, all other sensors, such as Sun sensors or star trackers, would become unreliable. Our research was inspired by studies of the attitude motion of the Earth Radiation Budget Satellite (ERBS) during the July 2, 1987, control anomaly. An analysis of the playback data (see Reference 3), revealed that the stimulation of the Sun sensor by bright Earth during one of the real-time passes led to an initially incorrect conclusion about the spacecraft orientation in the post G-Rate mode.

Although the attitude control system does not utilize gyro measurements under normal conditions, our analysis showed that these measurements can be effectively coupled with the magnetometer data to determine the attitude when angular rates are lower than the saturation limits on gyro output. Nevertheless, to give a worst case, we also assume a gyro failure either because of exceeding the telemetry limit or like that recently experienced by the Cosmic Background Explorer (COBE).

Therefore, the problem is to determine the attitude using only magnetometer data with no *a priori* knowledge of the spacecraft orientation. The latter requirement makes this research essentially different from the previous studies of attitude determination from magnetometer-only data via the Kalman filtering (see References 1 and 2). This is because the dynamical equations must first be linearized near their approximate solution. The solution was assumed known in References 1 and 2, which discussed a spacecraft under normal conditions, whereas this paper is focused on development of a deterministic algorithm for making the first guess in a situation when the attitude of the spacecraft deviates substantially from the expectations. After an approximate solution is found through a deterministic algorithm, it could be improved using the filtering technique (see References 1 and 2).

We have identified the two most typical attitude acquisition phases likely to be encountered under the contingency conditions:

(1) No thruster firing acquisition phase (angular rates <0.2 degree/second (deg/sec))

(2) Thruster firing acquisition phase (spinning rates within 10-50 deg/sec)

Due to relatively small angular rates in phase 1, the control system can significantly affect the spacecraft tumbling and, as a result, several situations should be studied. The following operational modes have been identified as the most representative choices:

(1a) Magnetic despin of a spacecraft (the B-dot mode) (see References 4 and 5)

(1b) Control system turned-off

(1c) A "blind" control system randomly rocking the spacecraft

(1d) Stabilization of the spacecraft by means of nutation damping

For Phase 2, the control system is expected to play a relatively minor role, and, consequently, spacecraft tumbling is expected to be predominantly governed by the torque-free Euler equations.

The paper presents a new deterministic algorithm, which works under the conditions that (1) internal torques are known and (2) external torques are negligible. Environmental torques are expected to be negligible either because of large angular momentum of the spacecraft or when compared with internal torques. Thruster firing acquisition phase and the B-dot mode give typical examples of such situations. Also, the algorithm can be used (at least in principle) to determine the attitude of a spacecraft governed by a "blind" control system (operational mode (1c)), when momentum wheel and scanwheel speeds and electromagnetic dipole moments are available from the telemetry data.

2. ANGULAR RATE UNCERTAINTY CIRCLE (ARUC)

Let $\vec{B}^A$ and $\vec{B}^R$ be the vectors of geomagnetic field measured in the body-fixed and reference frames, respectively:

$$\underset{\approx}{A}\,\vec{B}^R = \vec{B}^A \tag{2-1a}$$

The time derivatives $\dot{\vec{B}}^A$ and $\dot{\vec{B}}^R$ of two vectors are connected by the relation

$$\underset{\approx}{A}\,\dot{\vec{B}}^R = \dot{\vec{B}}^A + \vec{\omega}^A \times \vec{B}^A \tag{2-1b}$$

where $\vec{\omega}^A$ is the angular velocity vector referred to the body-fixed frame and the attitude matrix $\underset{\approx}{A}$ represents the orientation of one frame with respect to another. The vector $\dot{\vec{B}}^A$ can be computed from two sequential magnetometer measurements $\vec{B}_1^A$ and $\vec{B}_2^A$ by using the finite-difference approximation. The vector $\dot{\vec{B}}^R$, like the vector $\vec{B}^R$ itself, is found from the geomagnetic field model, assuming that the position of the spacecraft in space is known.

If the angular velocity vector $\vec{\omega}^A$ can be extracted from gyro measurements, Equations (2-1a) and (2-1b) can be directly used to determine the attitude via the TRIAD

algorithm (see Reference 6), which implements the so-called "algebraic method" of three-axis attitude determination (see Reference 7).

If only magnetometer measurements are used, the set of Equations (2-1a), (2-1b) is incomplete. In particular, the projection of the angular velocity vector $\vec{\omega}^A$ on geomagnetic field can be arbitrarily changed without violating Equation (2-1b). It is shown below that the projection of $\vec{\omega}^A$ on the plane perpendicular to the vector $\vec{B}^A$ is restricted by Equations (2-1a), (2-1b) to a circle, referred to below as the Angular Rate Uncertainty Circle (ARUC). To determine the attitude, it is necessary to know the position of the latter projection on the ARUC (i.e., the angle Φ in Figure 1, explained below). This requires the third sequential magnetometer measurement, which makes it possible to compute the second derivative of the vector $\vec{B}^A$ with respect to time. The algorithm that allows one to unambiguously determine both the attitude matrix $\underset{\approx}{A}$ and the angular velocity error $\vec{\omega}^A$ is outlined in Section 3.

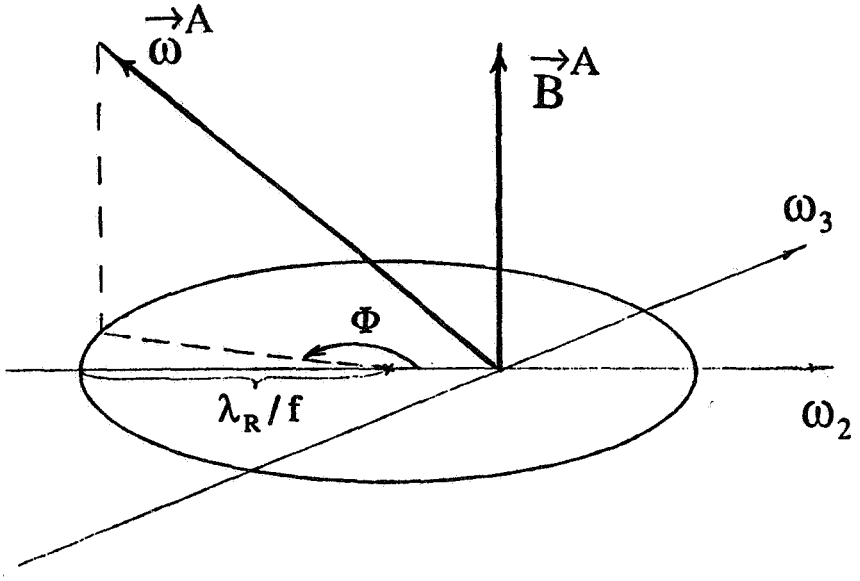


Figure 1. Angular Rate Uncertainty Circle (ARUC)

This section is focused on the information that can be extracted only from two sequential magnetometer measurements, giving rise to the particular ARUC. Calculating the square of magnitude of the vectors in the left- and right-hand sides of Equation (2-1) to exclude the attitude matrix, we come to the equation

$$|\dot{\vec{B}}^R|^2 - |\dot{\vec{B}}^A|^2 = |\vec{\omega}^A_\perp \times \vec{B}^A|^2 + 2\vec{\omega}^A_\perp \bullet (\dot{\vec{B}}^A \times \vec{B}^A) \qquad (2\text{-}2)$$

which contains only the projection $\vec{\omega}^A_\perp$ of the angular velocity vector $\vec{\omega}^A$ on the plane perpendicular to the geomagnetic field. (The vector $\vec{\omega}^A_\perp$ is referred to below as the transverse angular velocity.) Denoting the projections of $\vec{\omega}^A$ on the mutually perpendicular vectors,

$$\vec{D}_1 \equiv \vec{B}^A,\ \vec{D}_2 \equiv \vec{D}_1 \times \dot{\vec{B}}^A,\ \vec{D}_3 \equiv \vec{D}_1 \times \vec{D}_2 \tag{2-3}$$

by ω_1, ω_2, ω_3, one can easily see that the projections ω_2 and ω_3 lie on the circle:

$$(f\omega_2 + \lambda_A)^2 + f^2\,\omega_3^2 = \lambda_R^2 \tag{2-4}$$

(See Figure 1). The parameters f, λ_A, and λ_R are defined as follows:

$$f \equiv |\vec{B}^R|/|\dot{\vec{B}}^R|,\ \lambda_A \equiv a\,\sin\psi_A,\ \lambda_R \equiv \sin\psi_R, \tag{2-5}$$

where

$$a \equiv |\dot{\vec{B}}^A|/|\dot{\vec{B}}^R| \tag{2-6}$$

and ψ_K (K = A, R) is the angle between the vectors $\vec{B}^K$ and $\dot{\vec{B}}^K$ (K = A, R). The center of the ARUC always lies in the left semiplane of the $\omega_2\,\omega_3$ plane. Depending on the value of the parameter a, the ARUC either lies completely in this semiplane (if $a > 1$) or crosses the ordinate at two points (if $a < 1$). For $a = 1$, the ARUC is tangent to the ordinate at the origin, and this is the only case when zero angular velocity is among the allowed solutions; otherwise, the spacecraft must rotate. The projection of angular velocity along the vector $\vec{B}^A$ remains completely unrestricted unless the second derivatives of the geomagnetic field with respect to time are taken into account.

By analogy with the TRIAD algorithm (see Reference 6), we introduce three normalized reference vectors:

$$\hat{U}_1^R = \hat{B}^R,\ \hat{U}_2^R = \hat{U}_1^R \times \dot{\vec{B}}^R/(|\dot{\vec{B}}^R|\sin\psi_R),\ \hat{U}_3^R = \hat{U}_1^R \times \hat{U}_2^R \tag{2-7}$$

The crucial difference, however, comes from the fact that they can be transformed into their counterparts, $\hat{D}_1$, $\hat{D}_2$, $\hat{D}_3$ by the rotation $\underset{\approx}{A}$ only when the angular velocity vector is directed along the geomagnetic field.

It follows from Equation (2-7) that the unit vectors $\underset{\approx}{A}\,\hat{U}_2^R$ and $\underset{\approx}{A}\,\hat{U}_3^R$ are both orthogonal to the vector $\vec{B}^A$. As the same is true for the unit vectors $\hat{D}_2$ and $\hat{D}_3$ by definition, these two pairs of the mutually orthogonal vectors are related to each other as follows:

$$\underset{\approx}{A}\,\hat{U}_2^R = \cos\Phi\,\hat{D}_2 + \sin\Phi\,\hat{D}_3 \tag{2-8a}$$

$$\underset{\approx}{A}\,\hat{U}_3^R = -\sin\Phi\,\hat{D}_2 + \cos\Phi\,\hat{D}_3 \tag{2-8b}$$

where the angle Φ ranges between 0 and 2π. Introducing the 3-by-3 orthogonal matrix,

$$\underset{\approx}{T}_1(\Phi) \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\Phi & -\sin\Phi \\ 0 & \sin\Phi & \cos\Phi \end{bmatrix} \tag{2-9}$$

Equations (2-8a) and (2-8b) can be represented in the matrix form

$$\underset{\approx}{A}\,\underset{\approx}{U} = \underset{\approx}{D}\,\underset{\approx}{T}_1(\Phi) \tag{2-10}$$

where $\underset{\approx}{D}$ and $\underset{\approx}{U}$ are 3-by-3 orthogonal matrixes having the vectors $\hat{D}_j$ and $\hat{U}_j$ (j = 1, 2, 3), respectively, as their columns,

$$\underset{\approx}{D} \equiv \left[\hat{D}_1 : \hat{D}_2 : \hat{D}_3\right], \quad \underset{\approx}{U} \equiv \left[\hat{U}_1^R : \hat{U}_2^R : \hat{U}_3^R\right] \tag{2-11}$$

Therefore,

$$\underset{\approx}{A} = \underset{\approx}{D}\,\underset{\approx}{T}_1(\Phi)\,\underset{\approx}{U}^{-1} \tag{2-12}$$

The angle Φ has a simple physical meaning; namely, it determines the position of the transverse angular velocity $\vec{\omega}_\perp^A$ on the ARUC. To prove this assertion, the vector $\vec{B}^A$ is written in terms of $\hat{D}_1$ and $\hat{D}_3$ using the relation

$$\hat{D}_3 = (\kappa_R \hat{D}_1 - \vec{B}^A/|\vec{B}^R|)/\lambda_A \tag{2-13}$$

which directly follows from the definition of the vector $\hat{D}_3$. This leads to the expression

$$\dot{\vec{B}}^A = |\dot{\vec{B}}^R| (\kappa_R \hat{D}_1 - \lambda_A \hat{D}_3) \tag{2-14}$$

Substituting Equation (2-14) in the right-hand side of Equation (2-1b) and representing $\vec{\omega}^A$ as $\omega_1 \hat{D}_1 + \omega_2 \hat{D}_2 + \omega_3 \hat{D}_3$ one finds

$$\underset{\approx}{A}\, \dot{\vec{B}}^R = |\dot{\vec{B}}^R| \left[\kappa_R \hat{D}_1 - \lambda_A \hat{D}_3 + f(\omega_3 \hat{D}_2 - \omega_2 \hat{D}_3)\right] \tag{2-15}$$

The vector $\dot{\vec{B}}^R$ in the left side of Equation (2-15) is expressed in terms of $\hat{U}_1^R$, $\hat{U}_3^R$ by analogy with Equation (2-14):

$$\dot{\vec{B}}^R = |\dot{\vec{B}}^R| \left[\kappa_R \hat{U}_1^R - \lambda_R \hat{U}_3^R\right] \tag{2-16}$$

Using Equation (2-8b) and comparing the coefficients of the vectors $\hat{D}_2$, $\hat{D}_3$ in both sides of the resulting equation, we get the relation:

$$f\omega_2 + \lambda_A = \lambda_R \cos\Phi, \quad f\omega_3 = \lambda_R \sin\Phi \tag{2-17}$$

that uniquely determines the transverse angular velocity $\vec{\omega}_\perp$ after the angle Φ is found. Coupled with Equation (2-12) for the attitude, this relation completes the information that can be extracted simply from Equations (2-1a), (2-1b), exploiting only two sequential magnetometer measurements.

3. USE OF THE SECOND DERIVATIVE OF GEOMAGNETIC FIELD WITH RESPECT TO TIME

In this section, we show how the position of the transverse angular velocity on the ARUC can be determined by using the second derivative of the geomagnetic field with respect to time in the case when body-fixed projections of the total torque acting on the spacecraft are known. To do it we differentiate Equation (2-1b) with respect to time and represent the resulting relation between second derivatives of the geomagnetic field measured in body-fixed and reference frames as

$$\underset{\approx}{A}\, \ddot{\vec{B}}^R = \ddot{\vec{B}}^A + \dot{\vec{\omega}}^A \times \vec{B}^A + 2\,\vec{\omega}_\perp^A \times \dot{\vec{B}}^A - \omega_\perp^2\, \vec{B}^A + \omega_1 \vec{C}_2 \tag{3-1}$$

where $\omega_\perp \equiv |\vec{\omega}^A_\perp|$ and

$$\vec{C}_2 \equiv 2\hat{D}_1 \times \dot{\vec{B}}^A + \hat{D}_1 \times (\vec{\omega}^A_\perp \times \vec{B}^A) \tag{3-2}$$

To calculate the second derivatives of the geomagnetic field, at least three measurements are needed: $\vec{B}^A_1$, $\vec{B}^A_2$, $\vec{B}^A_3$. To close the set of equations, it is also necessary to have an equation for $\dot{\vec{\omega}}^A$. As explained below, this equation can be easily included in the case of negligibly small external torques. Otherwise, it explicitly contains the unknown attitude matrix. The external torques can thus be taken into account only through an iterative procedure, which is vulnerable to measurement accuracy and may diverge.

For the particular case of constant angular velocity, $(\dot{\vec{\omega}}^A = \vec{0})$ projecting vector Equation (3-1) on the plane perpendicular to $\vec{C}_2$ makes it possible to exclude ω_1. It is convenient to use the same computation for the general case of nonzero $\dot{\vec{\omega}}^A$. The final equations are thus obtained by projecting vector Equation (3-1) on two mutually orthogonal unit vectors

$$\hat{C}_2 = [(\lambda_A + \lambda_R \cos \Phi)\, \hat{D}_2 + \lambda_R \sin \Phi \hat{D}_3]/c(\Phi) \tag{3-3}$$

and

$$\hat{C}_3 \equiv \hat{B}^A \times \hat{C}_2 = \left[(\lambda_A + \lambda_R \cos \Phi)\, \hat{D}_3 - \lambda_R \sin \Phi\, \hat{D}_2\right] /c(\Phi) \tag{3-4}$$

with

$$c(\Phi) = \sqrt{(\lambda_A + \lambda_R \cos \Phi)^2 + \lambda_R^2 \sin^2 \Phi} \tag{3-5}$$

(To derive Equation (3-3) from Equation (3-2) we used Equations (2-14) and (2-17) together with the definition of the vectors $\vec{D}_1$, $\vec{D}_2$, $\vec{D}_3$ (see Equation (2-3)). Note that there is no need to consider the equation obtained by projecting Equation (3-1) on the direction $\hat{B}^A$ of the magnetic field. In fact, Equation (2-2) shows that the projected equation can be represented as

$$\vec{B}^R \bullet \ddot{\vec{B}}^R + |\dot{\vec{B}}^R|^2 = \vec{B}^A \bullet \ddot{\vec{B}}^A + |\dot{\vec{B}}^A|^2 \tag{3-6}$$

Hence, it is equivalent to the first derivative of the equality $\vec{B}^R \bullet \dot{\vec{B}}^R = \vec{B}^A \bullet \dot{\vec{B}}^A$ with respect to time. Therefore, this projection simply describes the change in the parameters of the ARUC with time.

To compute the projections of the left-hand side, we first expressed the vectors $\vec{D}_2$, $\vec{D}_3$ in (3-3) and (3-4) in terms of $\underset{\approx}{A}\hat{U}_2^R$, $\underset{\approx}{A}\hat{U}_3^R$ from Equations (2-8a) and (2-8b). The final equations have the form

$$\delta_2(\Phi) - s_2(\Phi) = |\vec{B}|c(\Phi)\dot{\vec{\omega}}^A \cdot \hat{C}_3 + \omega_1(\Phi)|\dot{\vec{B}}^R|c^2(\Phi) \tag{3-7a}$$

$$\delta_3(\Phi) - s_3 = -|\vec{B}|c(\Phi)\dot{\vec{\omega}}^A \cdot \hat{C}_2 \tag{3-7b}$$

where

$$s_2(\Phi) \equiv -4\omega_3(\Phi)\lambda_A |\dot{\vec{B}}^R|\kappa_R \tag{3-8a}$$

$$s_3 \equiv -2\kappa_R|\dot{\vec{B}}^R|(1-\alpha^2)/f \tag{3-8b}$$

and

$$\delta_2(\Phi) = \lambda_R\Delta_2^R - \lambda_A\Delta_2^A + (\lambda_A\Delta_2^R - \lambda_R\Delta_2^A)\cos\Phi - (\lambda_A\Delta_3^R + \lambda_R\Delta_3^A)\sin\Phi \tag{3-9a}$$

$$\delta_3(\Phi) = \lambda_R\Delta_3^R - \lambda_A\Delta_3^A + (\lambda_A\Delta_3^R - \lambda_R\Delta_3^A)\cos\Phi + (\lambda_A\Delta_2^R + \lambda_R\Delta_2^A)\sin\Phi \tag{3-9b}$$

with

$$\Delta_j^A \equiv \hat{D}_j \cdot \ddot{\vec{B}}^A, \qquad j = 2, 3 \tag{3-10a}$$

$$\Delta_j^R \equiv \hat{U}_j^R \cdot \ddot{\vec{B}}^R, \qquad j = 2, 3 \tag{3-10b}$$

The most important feature of Equations (3-7a), (3-7b) is that they do not contain the attitude matrix. The derived equations must be solved together with the dynamic equations of motion which make it possible to express $\vec{\omega}^A$ and $\dot{\vec{\omega}}^A$ in terms of torques. The full set of equations is closed provided that the torques are known.

3.1 CONSTANT ANGULAR VELOCITY

For constant angular velocity, the right-hand side of Equation (3-7b) vanishes and the resulting equation is transformed to a quadratic equation:

$$a_0 + 2a_1x + a_2x^2 = 0 \tag{3-11}$$

by the substitution $x = \tan(\Phi/2)$. The coefficients a_y ($y = 0,1,2$) in Equation (3-11) are defined as follows:

$$a_0 \equiv (\lambda_A + \lambda_R)(\Delta_3^R - \Delta_3^A) - s_3 \tag{3-12a}$$

$$a_1 \equiv \lambda_A\Delta_2^R + \lambda_R\Delta_2^A \tag{3-12b}$$

$$a_2 \equiv (\lambda_R - \lambda_A)(\Delta_3^R + \Delta_3^A) - s_3 \tag{3-12c}$$

After calculating two roots x_1 and x_2 of quadratic Equation (3-11) and substituting the appropriate values $\Phi_1 = 2 \arctan x_1$, $\Phi_2 = 2 \arctan x_2$, of the angle Φ in Equation (2-12), two possible solutions $\underset{\approx}{A}(\Phi_1)$ and $\underset{\approx}{A}(\Phi_2)$ for the attitude matrix are found. To select the correct solution it is necessary to calculate the angular velocity vector $\vec{\omega}^A(\Phi)$ for $\Phi = \Phi_k$, ($k = 1, 2$), using Equation (2-17) for $\omega_2(\Phi_k)$, $\omega_3(\Phi_k)$ and Equation (3-7a) for $\omega_1(\Phi_k)$:

$$\omega_1(\Phi_k) = [\delta_2(\Phi_k) - s_2(\Phi_k)]/[|\dot{\vec{B}}^R|c^2(\Phi_k)] \tag{3-13}$$

Taking into account that $\hat{B}^A = \underset{\approx}{A}(\Phi)\hat{B}^R$ for any point Φ on the ARUC (regardless of any error in data), the loss function is written as

$$L_{1,3}(\Phi_k) = \left[|\vec{B}_1^A - \underset{\approx}{A}^-(\Phi_k)\vec{B}_1^R| + |\vec{B}_3^A - \underset{\approx}{A}^+(\Phi_k)\vec{B}_3^R|\right]/(2\,dt) \tag{3-14}$$

where the matrices $\underset{\approx}{A}^-(\Phi_k)$ and $\underset{\approx}{A}^+(\Phi_k)$ are obtained by analytical propagation (see Equation (12-7b) in Reference 8) of the attitude backward ($t = -dt$) and forward ($t = dt$) in time t with constant angular velocity $\vec{\omega}^A(\Phi_k)$, starting from the matrix $\underset{\approx}{A}(\Phi_k)$ and assuming an equal time step dt between each sequential measurement. The correct root of Equation (3-11) is expected to give a smaller value for function (3-14), if all the time derivatives used in the algorithm are calculated accurately enough.

3.2 KNOWN INTERNAL TORQUES

Assuming that external torques are negligible, dynamic equations of motion are written as

$$\vec{N} = \underset{\approx}{I}\,\dot{\vec{\omega}}^{A} + \vec{\omega}^{A} \times \underset{\approx}{I}\,\vec{\omega}^{A} \tag{3-15}$$

where $\underset{\approx}{I}$ is the moment of inertia tensor and the internal torque $\vec{N}$ is a known function either of time or of the geomagnetic field. Two most important examples are torque-free rotation ($\vec{N} \equiv \vec{O}$) and the B-dot mode (see References 4 and 5). Expressing the components of the vector $\dot{\vec{\omega}}^{A}$ as quadratic polynomials of $\omega_1(\Phi), \omega_2(\Phi), \omega_3(\Phi)$ from Equation (3-15) and substituting the resulting expressions in Equation (3-7a) gives the quadratic equation for ω_1 with coefficients dependent on Φ. Each of two roots $\omega'_1(\Phi)$ and $\omega''_1(\Phi)$ of this quadratic equation is then substituted in Equation (3-7b), giving rise to two transcendental equations. After all possible solutions Φ_k of both transcendental equations are found, together with the appropriate vectors $\vec{\omega}^{A}(\Phi_k)$ and $\dot{\vec{\omega}}^{A}(\Phi_k)$, they are tested using loss function (3-14), where the matrices $\underset{\approx}{A}^{-}(\Phi_k)$ and $\underset{\approx}{A}^{+}(\Phi_k)$ are obtained by propagating numerically both the attitude and the angular velocity vector backward and forward in time, starting from the matrix $\underset{\approx}{A}(\Phi_k)$ and assuming the vector $\dot{\vec{\omega}}^{A}(\Phi_k)$ to be constant. Again the solution sought is expected to give the smallest value for loss function (3-14).

4. TESTS OF THE ALGORITHM

Both the algorithm and its software implementation have been tested for the ERBS in the arbitrarily selected time interval from 890115.000025 to 890115.005937. Geocentric inertial coordinates (GCI) were used as the reference frame. The observed attitude matrices $\underset{\approx}{A}$ were constructed with the same time step of 8 sec as that used in the processed engineering data (low-frequency format) containing both the magnetometer measurements $\vec{B}^{A}$ and the model geomagnetic field $\vec{B}^{R}$ in the GCI. The angular velocity $\vec{\omega}^{A}$ was calculated by numerically differentiating the matrix function $\underset{\approx}{A}$ (t) with respect to time t.

As the first step, oscillations of the body-fixed components of the angular velocity vector near its average value of [–0.018, 0.049, –0.034] deg/sec were neglected and propagation of the attitude matrix was performed analytically, assuming constant angular velocity. The body-fixed projections of the geomagnetic field were computed by means of Equation (2-1a), using the analytically calculated attitude matrix and the model geomagnetic field read with a time step of either 8 or 16 sec.

In Figure 2, we present two solutions of Equation (3-11) as functions of time t. The small plateau in the upper curve represents the region where discriminant becomes negative due to numerical errors in the vectors $\vec{B}^{A}$ and $\vec{B}^{R}$ evaluated using the finite-difference approximation. At these points, the program simply sets the discriminant equal to zero (see Figure 3) and picks up both solutions from the previous time step. The small spike in the

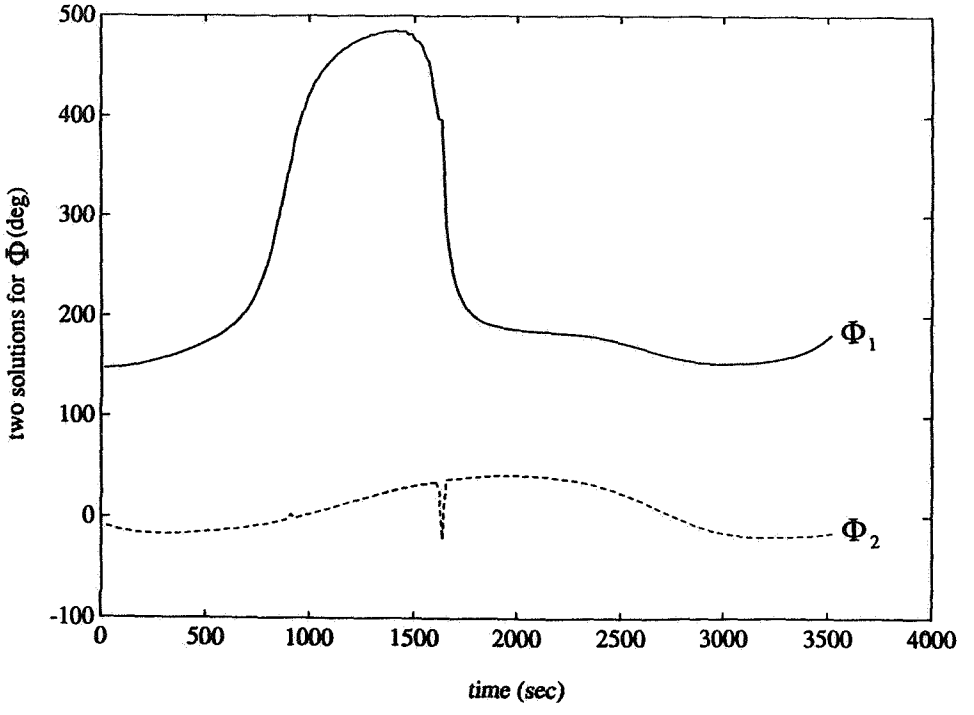


Figure 2. Test for Constant Angular Velocity (Step = 16 sec)

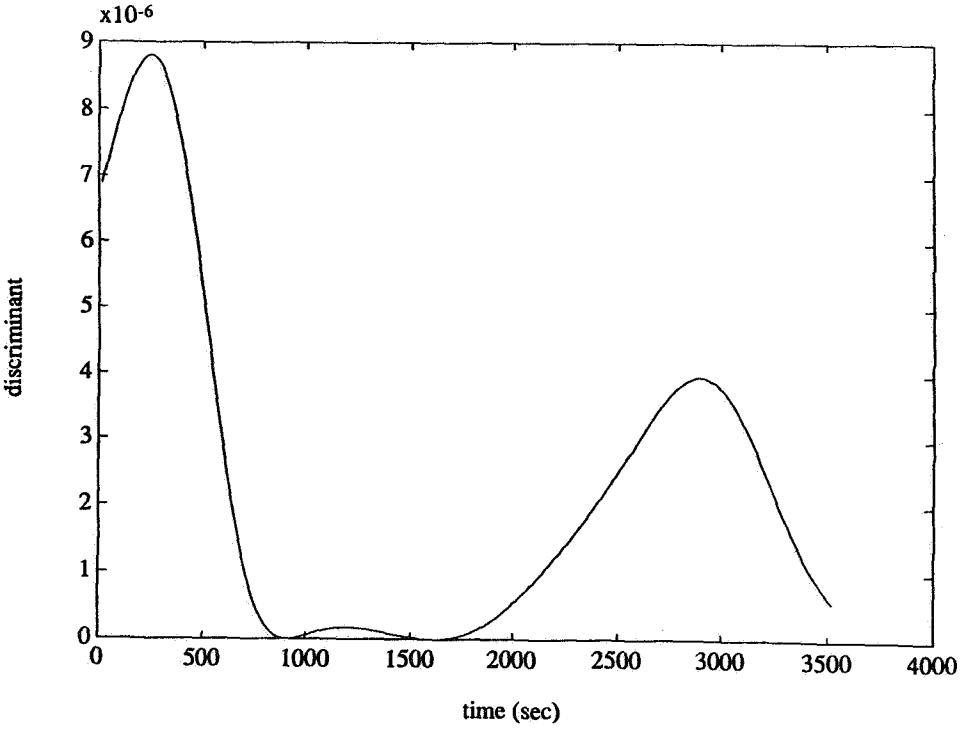


Figure 3. Test for Constant Angular Velocity (Step = 16 sec)

curve $\Phi_2(t)$ in Figure 2 at approximately 900 sec takes place where the discriminant illustrated in Figure 3 first touches the abscissa. The values of loss function (3-14) for each solution are presented in Figure 4. Due to errors in the time derivatives, two curves cross each other, and as a result, loss function (3-14) can be used to select the correct solution only in the region where the discriminant of quadratic Equation (3-11) is large.

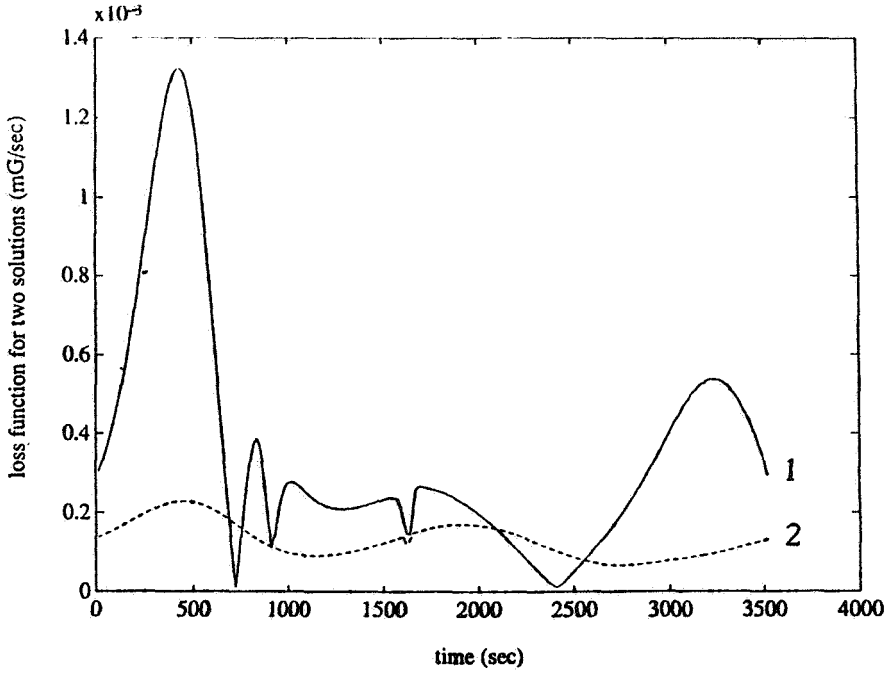
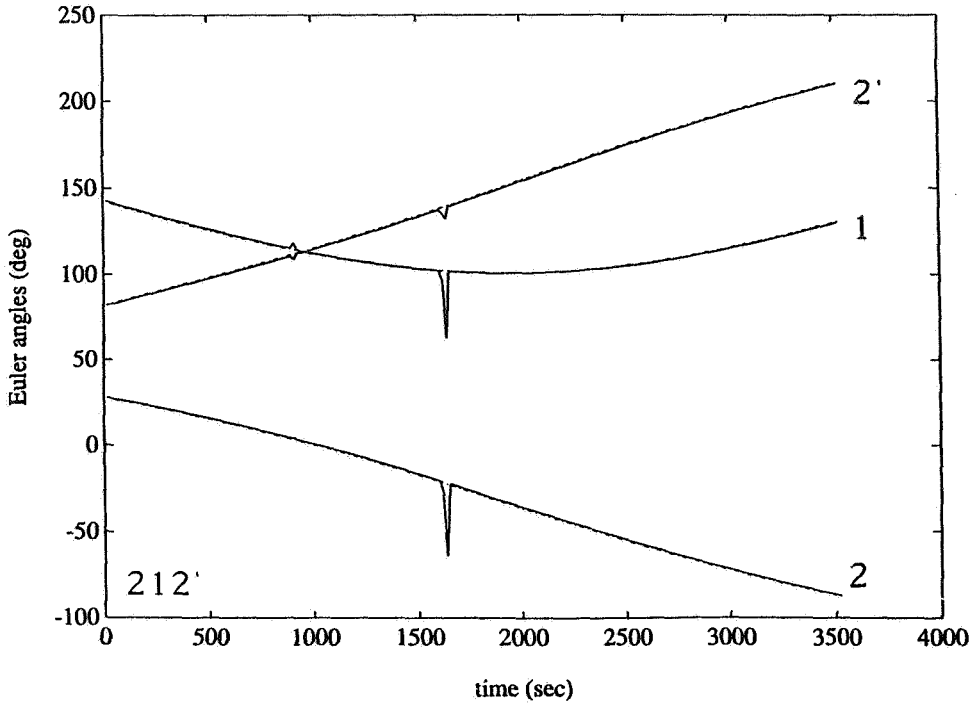


Figure 4. Test With Constant Angular Velocity (Step = 16 sec)

The attitude matrix $\underset{\approx}{A}$ is described here by a (212′) sequence of Euler rotations, using analytical formulae similar to Equations (12-21a) through (12-21c) in Reference 8. The values of the Euler angles determined by means of the developed algorithm are represented in Figure 5 by solid lines. The dot-dashed lines in Figure 5 represent the expected values in the limit of an infinitely small time step (the Euler angles were obtained from the analytically calculated attitude matrix). The agreement is reasonably good, except for the spikes in the region of significantly negative discriminant. It is worth mentioning that the small spikes observed in two upper curves in Figure 5 at approximately 900 sec completely disappeared when the smaller step of 8 sec was used to calculate the time derivatives of the geomagnetic field. This observation is in agreement with our statement that the observed errors are caused by a relatively large time step used for evaluating these derivatives.

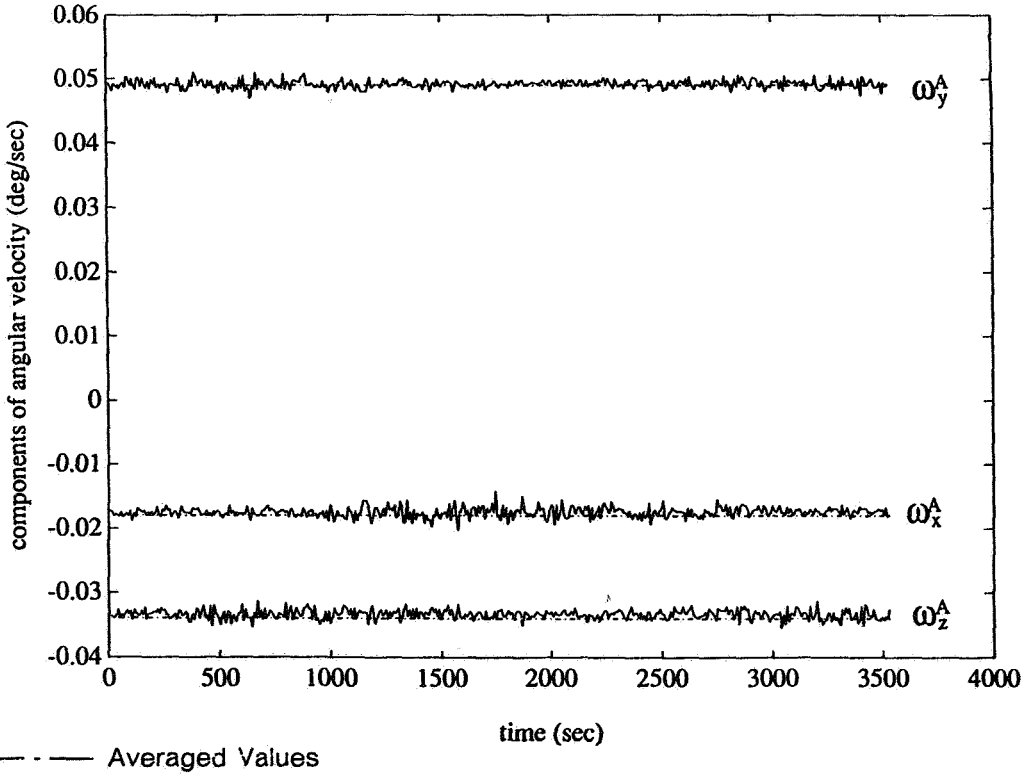
The solid lines in Figure 6 present the components of the angular velocity vector obtained by numerically differentiating the attitude matrix derived from the low-frequency telemetry data. The dot-dashed straight lines show the average values that were used for propagation of the attitude matrix in the tests discussed above. Despite the fact that high-frequency oscillations are relatively small, they essentially affect the attitude, as clearly seen from Figure 7, where the solid lines are the observed values of the Euler angles, and the dot-dashed curves are from Figure 5. The physical significance of the



— · — · — Propagated attitude

——————— Attitude determined from Equation (3-11)

Figure 5. Test for Constant Angular Velocity (Step = 16 sec)



— · — · — Averaged Values

——————— Extracted from the telemetry data

Figure 6. Measured/Averaged Angular Velocity

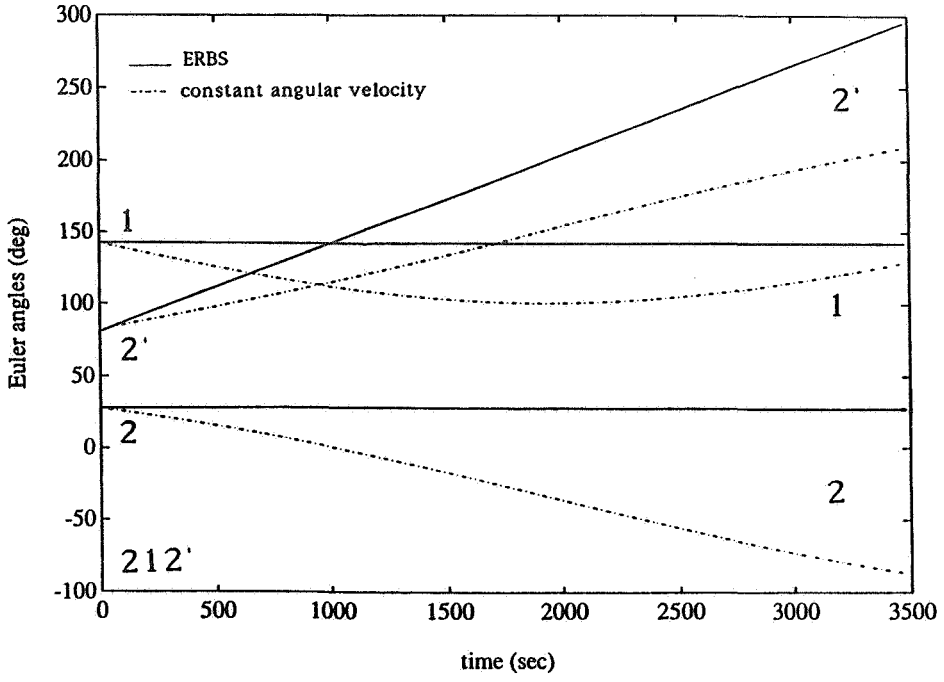
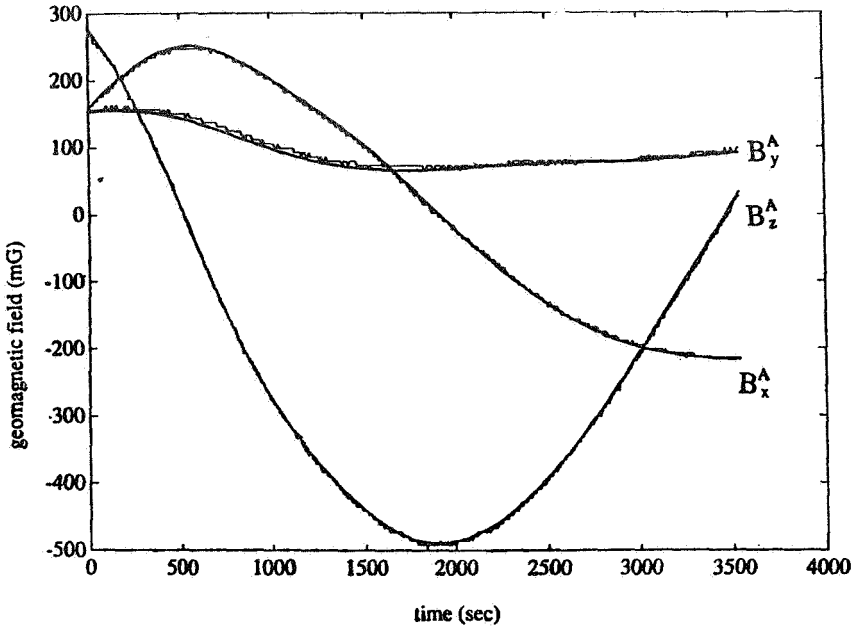


Figure 7. Effect of Averaging Angular Velocity

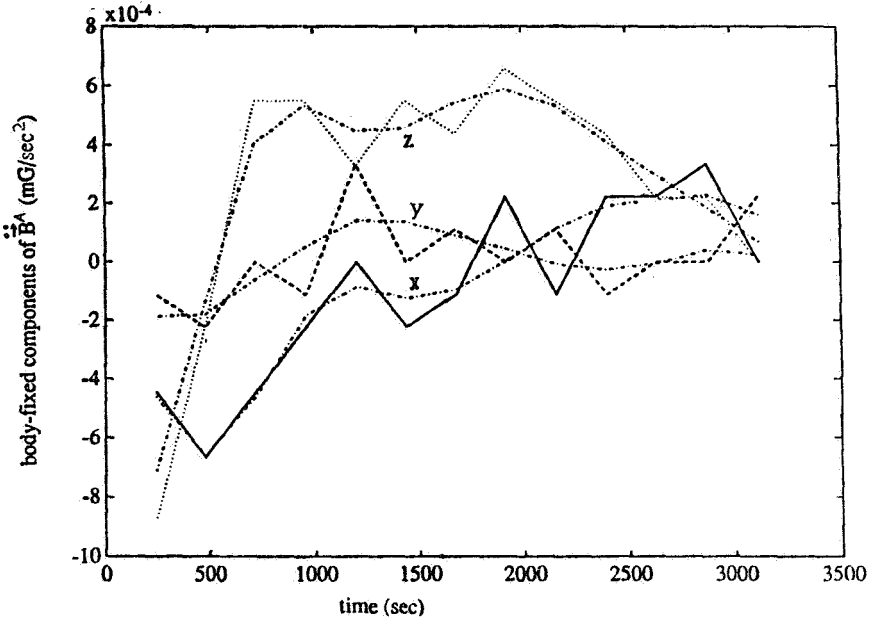
oscillations can be understood by analyzing behavior of the Euler angles 1 and 2, which determine the direction of the pitch body-fixed axis in the GCI (cf. Equation (12-20) in Reference 8). The oscillations simply force this direction to remain unchanged. It is remarkable that no oscillations are seen in the solid curves in Figure 7, despite the fact that the oscillations in angular velocity significantly affect the attitude.

In Figure 8, the magnetometer measurements taken from the low-frequency telemetry data are plotted versus the calculated body-fixed components of the geomagnetic field. The latter were obtained by rotating the geomagnetic field from the GCI frame to the body-fixed axes by means of the observed attitude matrix derived from telemetry data. The agreement looks reasonably good, except for the stepwise behavior of the measured data due to their analog-to-digital conversion with the increment of ~6.44 milligauss (mG). The coarse digitization of the magnetometer measurements creates an obstacle in calculating the second derivative $\ddot{\vec{B}}^A$ of the geomagnetic field. This is illustrated by Figure 9, where the zigzag lines were obtained by processing the magnetometer measurements and the dash-dotted lines represent the second derivative of the calculated geomagnetic field with the same finite-difference scheme and the same time step of 240 sec used in both cases. The digitization results in relatively large errors of $\pm$20-deg in attitude determination. In Figure 10, we plot the determined Euler angles (solid lines) versus their observed values (dot-dashed lines) selected at a time step of 240 sec. In Figure 11, for comparison, we give a similar plot for the Euler angles which were determined by utilizing the attitude information in the telemetry data to model a field measurement in the body-fixed frame and then using this in the algorithm to show the upper limit on accuracy. In addition to the curves exploiting the time step of 240 sec (solid lines) to calculate the



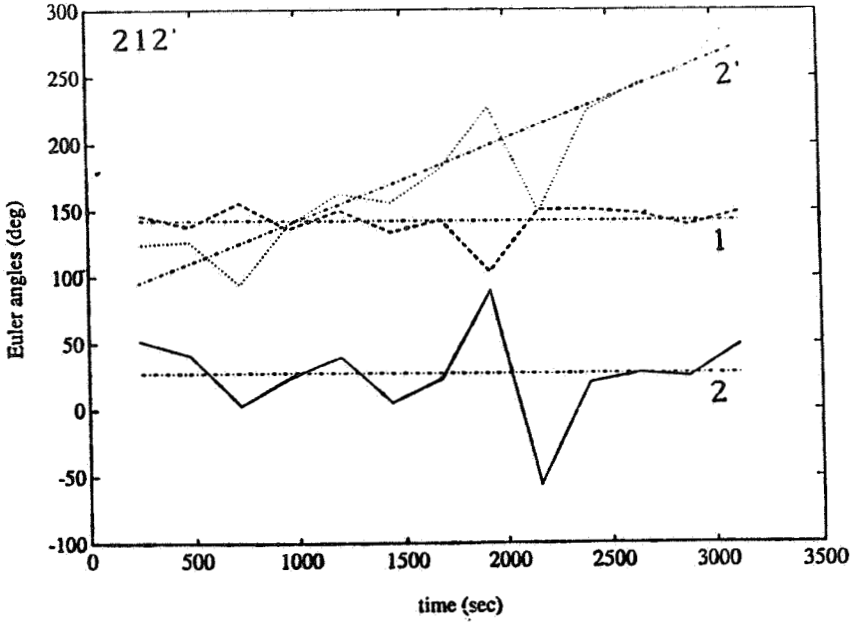
Stepwise lines show the magnetometer data; smooth curves were calculated by using Equation (2-1a).

Figure 8. Measured Versus Calculated Geomagnetic Field



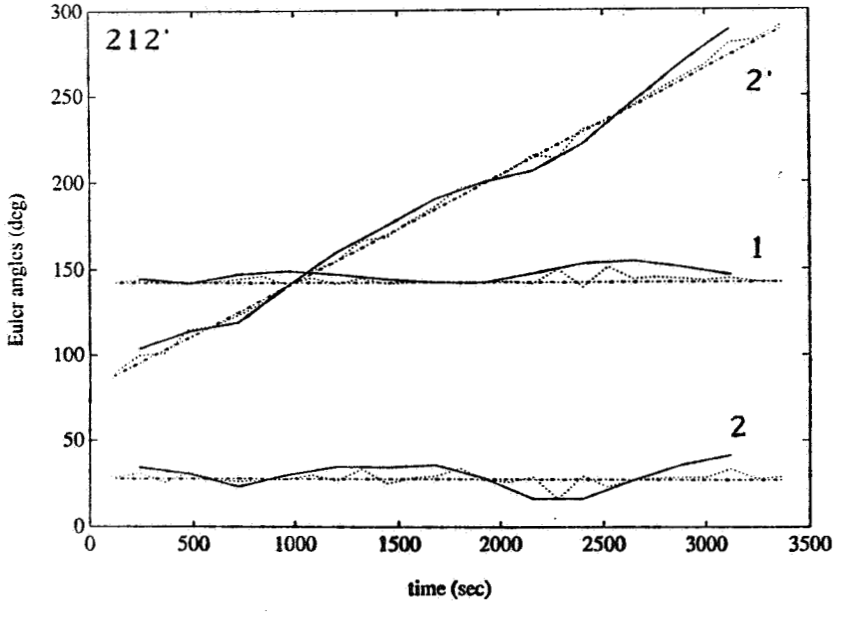
Zigzag lines obtained by processing the magnetometer data; smooth curves represent the second derivative of the calculated geomagnetic field with respect to time.

Figure 9. Second Derivatives of Measured Versus Calculated Geomagnetic Field With Respect to Time (Step = 240 sec)



Zigzag lines show the Euler angles determined from the magnetometer data; dot-dashed lines show the Euler angles determined using the calculated body-fixed components at the geomagnetic field.

Figure 10. Use of Measured Geomagnetic Field (Step = 240 sec)



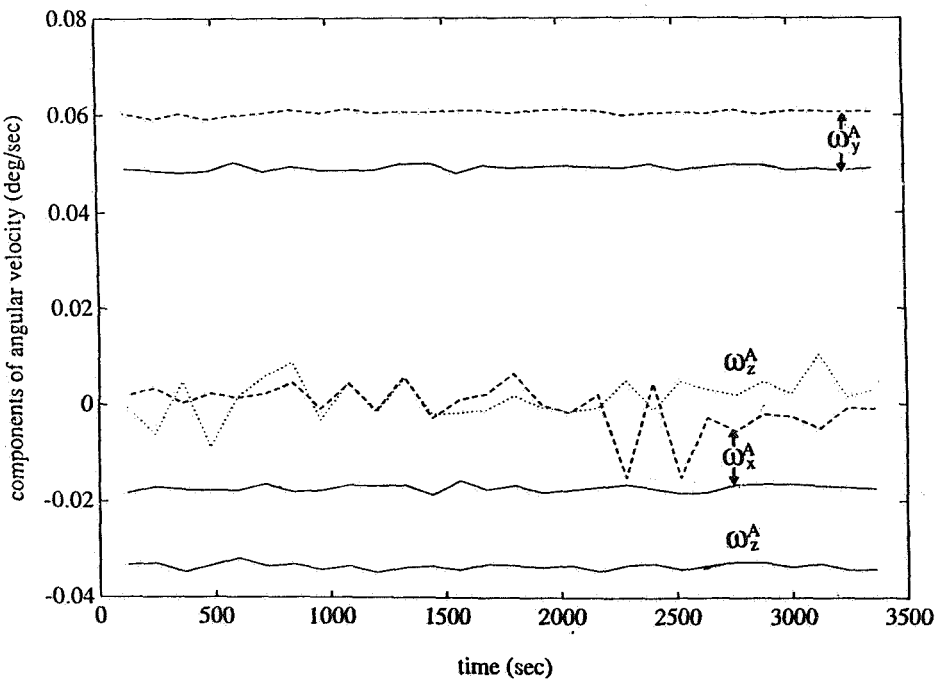
- - - - - Determined; with a step of 120 sec; —— · —— · —— observed

——— Determined; with a step of 240 sec

Figure 11. Use of Calculated Geomagnetic Field (Steps = 120, 240 sec)

necessary time derivatives, Figure 11 also reproduces the Euler angles determined using the time step of 120 sec (dotted lines). In the case of the measured geomagnetic field, use of the large step was necessary to smooth the data. However, it results in some systematic errors clearly seen in Figure 11. A further decrease in a time step used to compute the time derivatives of the calculated geomagnetic field results in accumulation of errors caused by oscillations of angular velocity, which are disregarded in the algorithm. Therefore, the time step of 120 sec turns out to be an optimum compromise, providing accuracy of ~5 deg for each angle.

The observed oscillations of the angular velocity vector significantly affected the ability of the algorithm to determine its body-fixed components. In Figure 12, the dashed and dotted lines show the values of these components determined using a time step of 120 sec, and the solid lines show the observed values selected at the same time step. The total angular rate of 0.062 deg/sec is reasonably well reproduced by the dominant y-component of the determined angular velocity vector, whereas the two remaining components are too small to contribute and are thus in obvious disagreement with the observations.



Zigzag lines show the values determined using the developed algorithm; solid lines show the observed values.

Figure 12. Use of Calculated Geomagnetic Field (Step = 120 sec)

5. CONCLUSIONS AND FURTHER DEVELOPMENTS

The reported preliminary analysis demonstrates that the deterministic approach to coarse attitude determination, using only magnetometer data, is feasible. A successful

implementation could benefit significantly from more accurate representation of magnetometer measurements in telemetry records than is provided for the ERBS.

Our study of the applicability of the algorithm to attitude determination under normal conditions is mostly methodological and illustrative. As mentioned in the introduction, the main objective is to develop an attitude determination system for application under contingency conditions when only magnetometer data are available. In particular, the analytical formula derived here for the limiting case of constant angular velocity could be applied to a spacecraft rotating around its major principal axis after it was stabilized using nutation damping. At this time, we are studying applicability of the developed algorithm to a spacecraft in the B-dot mode and to a spacecraft freely rotating with high angular speeds caused by thruster firing. The errors from neglecting environmental effects in both cases are now being investigated.

ACKNOWLEDGMENTS

The authors thank M. Phenneger for his guidance and extremely valuable critical remarks on the manuscript. One of the authors (G. A. Natanson) is also indebted to D. Chu for interesting, encouraging discussions, to F. L. Markley for his comments on the literature on magnetic despin of a spacecraft, and to B. Rashkin for his helpful suggestions on the presentation of the results.

REFERENCES

1. G. A. Heyler, "Attitude Determination by Enhanced Kalman Filtering Using Euler Parameter Dynamics and Rotational Update Equations," AIAA Paper No. A81-45832, AAS/AIAA Astrodynamics Specialist Conference, Lake Tahoe, Nevada, Aug. 3-5, 1981

2. F. Martel, P. K. Pal, and M. L. Psiaki, "Three-Axis Attitude Determination via Kalman Filtering of Magnetometer Data," Paper No. 17 for the Flight Mechanics/ Estimation Theory Symposium, NASA/Goddard Space Flight Center, Greenbelt, Maryland, May 10 and 11, 1988

3. J. Kronenwetter and M. Phenneger, *Attitude Analysis of the Earth Radiation Budget Satellite (ERBS) Control Anomaly*, CSC/TM-88/6154

4. A. C. Stickler and K. T. Alfriend, *Elementary Magnetic Attitude Control System*, Journal of Spacecraft, Volume 13, pp. 282-287, 1976

5. F. L. Markley "Attitude Control Algorithms for the Solar Maximum Mission," AIAA Paper No. 78-1247 for the 1978 AIAA Guidance and Control Conference, Palo Alto, CA 1978

6. M. D. Shuster and S. D. Oh, *Three-Axis Attitude Determination from Vector Observations*, J. Guidance and Control, 4, pp. 70-77, 1981

7. G. M. Lerner, "Three-Axis Attitude Determination," in *Spacecraft Attitude Determination and Control*, J. R. Wertz, ed. Dordrecht, Holland: D. Reidel Publishing Co., 1978, pp. 420-428

8. G. M. Lerner, "Parameterization of the Attitude," in *Spacecraft Attitude Determination and Control*, J. R. Wertz, ed. Dordrecht, Holland: D. Reidel Publishing Co., 1978, pp. 412-420

On-board Attitude Determination and Control Algorithms for SAMPEX

Thomas W. Flatley
Josephine K. Forden
Debra A. Henretty
E. Glenn Lightsey
F. Landis Markley

Guidance and Control Branch
NASA/Goddard Space Flight Center

Algorithms for onboard attitude determination and control of the Solar, Anomalous, and Magnetospheric Particle Explorer (SAMPEX) have been developed. The algorithms include: spacecraft ephemeris and geomagnetic field models, attitude determination with 2 degree accuracy, control of pitch axis pointing to the sun and yaw axis pointing away from the earth to achieve control of pitch axis within 5 degrees of sunline, momentum unloading, and nutation damping. The closed loop simulations were performed on a VAX 8830 using a prototype version of the on-board software.

I. INTRODUCTION

The Solar, Anomalous, and Magnetospheric Particle Explorer (SAMPEX), scheduled for launch in June 1992, is the first in the Small Explorer (SMEX) series of spacecraft. The SAMPEX software algorithm design philosophy is to develop the common aspects of the attitude determination and control (ADC), such as sensor data processing, attitude determination, ephemeris propagation, and command and telemetry processing, in as generic a form as possible.

SAMPEX, a 169 kg payload, will be launched into a 580 km circular orbit with 82 degrees inclination by a Scout launch vehicle. The spacecraft attitude must simultaneously satisfy a variety of on-orbit pointing requirements. The scientific instruments are located on the +z side of the spacecraft (Figure 1), and detect the impingement of magnetically aligned solar and cosmic particles. The spacecraft z-axis (instrument boresight) should therefore point within 15 degrees of local vertical (zenith) in the polar regions. Due to damaging space dust and orbital debris, the experiment boresight is required to point at least 45 degrees away from the spacecraft ram velocity vector. The spacecraft has fixed solar arrays with all cells on the +y side. Pointing of the +y axis to within 5 degrees of the sun line is desired during the entire mission. Attitude knowledge to 2 degrees or better is also required. [1]

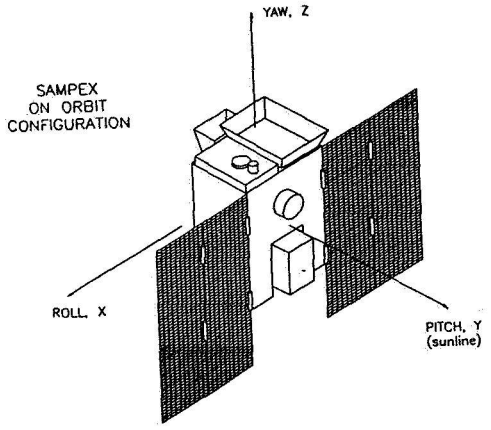


Figure 1. Spacecraft Mechanical Configuration

The ADC sensors include a two-axis fine digital sun sensor, five coarse sun sensors , and a triaxial search coil magnetometer. All sensor data are acquired and processed at the ADC sampling frequency, 2 Hz. Three magnetic torquer bars and a pitch axis momentum wheel are used as control actuators. The sensor data processing algorithms are presented in Section II.

There are two science pointing modes - vertical pointing and orbit rate rotation. The computation of the pitch error angle is different for the two modes. The on-board attitude determination algorithm is based on the sun and magnetic field data only. The rate information will be generated without gyros and the nadir

vector will be obtained without earth sensors. Therefore accurate spacecraft ephemeris and magnetic field vector knowledge are essential. The detailed description of the attitude determination is given in Section III.

The vertical pointing mode minimizes the angle between the spacecraft z-axis and the zenith vector within the sun pointing constraint. However, this mode has the undesirable property of pointing the experiment boresight directly into the ram vector twice per orbit when the orbit plane is parallel with the sun vector. The orbit rate rotation mode does not maintain the instrument boresight as close to the zenith vector as the vertical pointing mode, but it satisfies the spacecraft pointing requirement as well as the velocity vector avoidance criterion. The control scheme for each of these modes is described in Section IV. The digital sun pointing only mode, which does not perform pitch axis control, is used for reacquisition and is also described in this section.

This paper also discusses the attitude dynamic simulator and results used for analysis in Section V, and presents a summary in Section VI.

II. SENSOR DATA PROCESSING

Sun Sensor Data Processing

The fine sun sensor (FSS) is used for on-board attitude determination and sun pointing. It is a two-axis sun sensor with a field of view of ± 64 degrees. It outputs 8 bits of data in gray code for each axis with a resolution of 0.5 degrees. After the gray coded data is converted to binary, the two binary counts (N_a and N_b) are converted into coordinates of the sun's image within the sensor, i.e.

$$x = 0.002754\, N_a - 0.350625$$
$$z = 0.002754\, N_b - 0.350625$$

A sun vector (s) in the spacecraft body frame is computed using the following equations [2]:

$$s = \frac{1}{\sqrt{x^2 + z^2 + t^2}} \begin{bmatrix} -n\,x \\ \sqrt{t^2 - (n^2 - 1)\,(x^2 + z^2)} \\ -n\,z \end{bmatrix}$$

Where:
- n - the refraction index of the FSS glass
- t - the glass thickness (cm)
- x, z - measurements in the sensor frame (cm)

The above equations are derived based on the assumptions : 1) the sensor frame x, y,

z is aligned with the body frame x, y, z axes respectively, 2) the misalignment between sensor frame and the body frame is within tolerance, 3) pitch axis is normal to FSS. However, the FSS algorithm has means to correct 90 degree and 180 degree misalignment of the x, z axes.

The coarse sun sensor (CSS) is used for initial acquisition, reacquisition, and on-board attitude determination when the sun is out of the field of view of the FSS. The CSS system consists of five eyes: one pitch eye mounted on the center of the spacecraft facing the negative pitch(y) axis, two roll(x) eyes and two yaw(z) eyes. Each pair is mounted on the solar array panels facing 180 degrees apart along the axis. The roll and yaw eye outputs are differenced. All outputs are converted to 12 bit words. When divided by the maximum expected output these signals represent the x, y, and z components of the sun vector in body frame. When the negative pitch eye is illuminated, the normalized coarse sun vector (s_c) in body frame is given by the following equation:

$$s_c = \frac{1}{\sqrt{x^2 + y^2 + z^2}} \begin{bmatrix} x \\ -y \\ z \end{bmatrix}$$

and when the negative pitch eye is not illuminated, s_c is given by the following equation:

$$s_c = \begin{bmatrix} x \\ \sqrt{1 - x^2 - z^2} \\ z \end{bmatrix}$$

when $x^2+z^2 >= 1$, the above equation will become:

$$s_c = \frac{1}{\sqrt{x^2 + z^2}} \begin{bmatrix} x \\ 0 \\ z \end{bmatrix}$$

Magnetometer Data Processing

The triaxial search coil magnetometer is used for on-board attitude determination and momentum management (sun pointing and momentum unloading). It outputs 12 bit words for each axis. Its resolution is 0.3125×10^{-7} tesla and its range is $\pm 640 \times 10^{-7}$ tesla. The earth's magnetic field in the body frame is computed using the following equation [2]:

$$\mathbf{B} = [\text{Scale}]\mathbf{V} + \mathbf{Bias} - [\text{C}]\mathbf{dm}$$

where: **B** - magnetic field vector (tesla)
[Scale] - scale factor matrix with all off-diagonal elements being

zero (tesla/volts)

V - magnetometer reading (volts)

Bias - constant magnetic field bias (tesla)

[C] - torquer bar & magnetometer coupling matrix (tesla/A-m^2)

dm - dipole moment of torque rods computed from the torque magnetic assembly current feedback (A-m^2).

The SAMPEX algorithm contains a procedure to calibrate the torque-rod/ magnetometer coupling matrix ([C]). This procedure is excited periodically via ground command. It turns off all torquer bars first and obtains the uncontaminated magnetometer readings ($\mathbf{B_0}$). Then it sequentially turns on each torquer bar with a 10 A-m^2 excitation and obtains the contaminated magnetometer reading ($\mathbf{B_c}$). Finally [C] is computed using the following equation:

$$\mathbf{B_c} = \mathbf{B_0} + [\mathbf{C}]\mathbf{dm}.$$

This technique enables attitude determination to be performed in the presence of significant contamination of the magnetometer signal during magnetic actuation.

Momentum Wheel Speed Processing

The momentum wheel is used for pitch error angle control. The momentum wheel speed is used to generate the system momentum vector for momentum management. The processing of momentum wheel speed simply multiplies the raw tachometer signal by a scale factor and corrects it with a constant bias if necessary.

III. ATTITUDE DETERMINATION

The objective of on-board attitude determination is to produce an estimate of the inertial-to-body transformation matrix for the system, a representation of its three-axis attitude. The attitude determination process includes: generation of the spacecraft position vector, sun vector and earth's magnetic field vector in the inertial frame, construction of the transformation matrix, and computation of the pitch error angle, error rate and body angular momentum for control.

The attitude determination algorithm is disabled when the sun vector aligns with the earth's magnetic field vector.

During eclipse, when the sun information is not available, the body pitch axis is assumed to remain inertially fixed at the value given by the second row of the last inertial-to-body transformation matrix computed before entering eclipse. This vector J and j, a y-axis unit vector in the body frame, are substituted for the inertial

and body sun vectors in the construction of the transformation matrix.

Figure 2 illustrates the scheme of construction of the inertial-to-body transformation matrix.

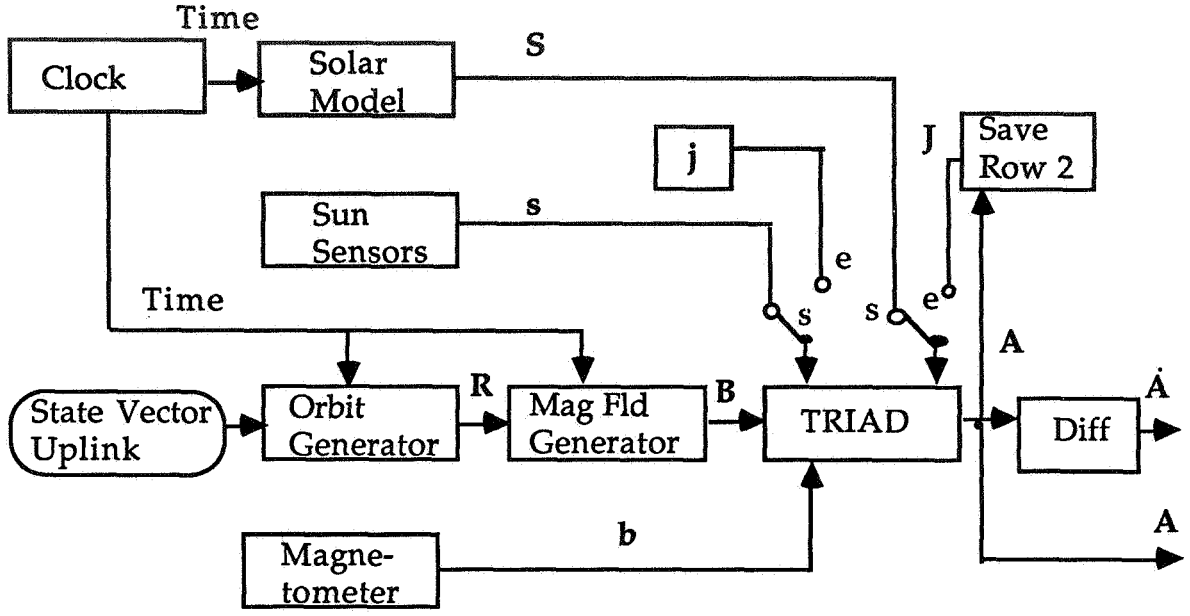


Figure 2. Construction of Inertial-To-Body Transformation Matrix

Reference Vector Generation

Although pointing requirements for SAMPEX are modest, accurate on-board spacecraft ephemeris and magnetic field vectors are required to meet the design goal of 2 degrees accuracy for on-board attitude determination.

The spacecraft ephemeris is propagated employing the fourth order Runge-Kutta integration method with integration step size equal to the ADC sampling frequency. The dynamic equation for generating the spacecraft ephemeris includes the perturbations due to the zonal gravity harmonics and the atmospheric drag [3]:

$$\ddot{\mathbf{R}} = -\frac{\mu}{|\mathbf{R}|^3}\mathbf{R} + \mathbf{a}_g + \mathbf{a}_d$$

where:
- $\mathbf{R}$ - spacecraft position vector (km)
- μ - earth's gravitational constant, 3.9860064×10^5 (km^3/sec^2)
- $|\mathbf{R}|$ - magnitude of $\mathbf{R}$ (km)
- $\mathbf{a}_g$ - acceleration due to zonal gravity harmonics (km/sec^2)
- $\mathbf{a}_d$ - acceleration due to atmospheric drag (km/sec^2)

The nonspherical contribution to the gravity acceleration is computed with the spherical harmonic model using zonal harmonic coefficients up to J_4.

$$a_g = \frac{\mu R_e^2}{2.0\,|R|^5}\begin{bmatrix} R(1)\,(f_{xy2}+f_{xy3}+f_{xy4}) \\ R(2)\,(f_{xy2}+f_{xy3}+f_{xy4}) \\ R(3)(f_{z2}+f_{z4})+f_{z3} \end{bmatrix}$$

where
$$f_{xy2} = 3.0\,j_2\,(5.0\,Z_r^2 - 1.0)$$
$$f_{xy3} = 5.0\,j_3\,(R_e/|R|)\,(7.0\,Z_r^3 - 3.0Z_r)$$
$$f_{xy4} = 3.75\,j_4\,(R_e/|R|)^2(21.0\,Z_r^4 - 14.0\,Z_r^2 + 1.0)$$
$$f_{z2} = 3.0\,j_2\,(5.0\,Z_r^2 - 3.0)$$
$$f_{z3} = j_3\,(R_e/|R|)\,(Z_r\,(35.0\,Z_r^3 - 30.0\,Z_r) + 3.0|R|)$$
$$f_{z4} = 1.25\,j_4\,(R_e/|R|)^2\,(63.0\,Z_r^4 - 70.0\,Z_r^2 + 15.0)$$
$$j_2 = 1.0826271 \times 10^{-3}$$
$$j_3 = -2.5358868 \times 10^{-6}$$
$$j_4 = -1.6246180 \times 10^{-6}$$
R_e = 6.37814×10^3, earth radius at equator (km)
$$Z_r = \frac{R(3)}{|R|}$$

The drag model is

$$a_d = [-\rho d(1+f)A_C/2m)\;|\dot{R}_r|\;\dot{R}_r$$

where:
- ρ - atmospheric density at the spacecraft's altitude computed using a Jacchia-Roberts density model (kg/km^3), [4]
- d - drag coefficient (nominally 2.2)
- f - drag scale factor
- A_C - cross sectional area of spacecraft perpendicular to the direction of motion (km^2)
- m - mass of the spacecraft (kg)
- $|\dot{R}_r|$ - magnitude of $\dot{R}_r$ (km/sec)
- $\dot{R}_r$ - velocity of the spacecraft relative to the rotating atmosphere (km/sec)

$$\dot{R}_r = \begin{bmatrix} \dot{x} + \dot{\theta}\,y \\ \dot{y} - \dot{\theta}\,x \\ \dot{z} \end{bmatrix}$$

where $[\,\dot{x}\;\;\dot{y}\;\;\dot{z}\,]^T$ is the inertial velocity of the spacecraft and $\dot{\theta}$ is the rate of rotation

of the earth ($7.2921151467 \times 10^{-5}$ rad/sec).

The state vector, drag scale factor and parameters to specify the Jacchia-Roberts model are reinitialized every 24 hours via ground command. They are generated by Flight Dynamic Support System.

The earth's magnetic field is generated using the spherical harmonic model with order eight [2]. The Gaussian coefficients are updated to epoch 1985. The position of the sun is computed using a rapid analytical technique [2]. The parameters are updated to epoch 1991. Both the earth's magnetic field vector and sun vector are used to construct the transformation matrix.

Transformation Matrix Construction

The transformation matrix [A] is constructed by employing the standard TRIAD method [5]. This method only requires two body frame vectors and two reference frame vectors. The sun vector is more heavily weighted than the magnetic field vector except when the two vectors are perpendicular. The TRIAD method is given by:

$$[A] = [P]\,[Q]^T$$

where [P] and [Q] are orthogonal matrices. [P] is a matrix constructed using the sun (s) and earth's magnetic field (b) vectors in the body frame, and [Q] is a matrix constructed using the sun (S) and earth's magnetic field (B) vectors in the inertial frame:

$$\mathbf{P} = \left[\, \mathbf{s}\,,\ \frac{\mathbf{s}\times\mathbf{b}}{|\mathbf{s}\times\mathbf{b}|}\,,\ \frac{\mathbf{s}\times(\mathbf{s}\times\mathbf{b})}{|\mathbf{s}\times\mathbf{b}|}\,\right]$$

$$\mathbf{Q} = \left[\, \mathbf{S}\,,\ \frac{\mathbf{S}\times\mathbf{B}}{|\mathbf{S}\times\mathbf{B}|}\,,\ \frac{\mathbf{S}\times(\mathbf{S}\times\mathbf{B})}{|\mathbf{S}\times\mathbf{B}|}\,\right]$$

Computation of Angular Velocity

The angular velocity vector along with the momentum wheel speed is used to compute the system angular momentum for momentum management.

The angular velocity matrix is determined from the matrix identity:

$$\begin{bmatrix} 0 & \omega_z & -\omega_y \\ -\omega_z & 0 & \omega_x \\ \omega_y & -\omega_x & 0 \end{bmatrix} = [\dot{A}]\,[A]^T$$

where $[\dot{A}]$ is the derivative of the inertial-to-body transformation matrix [A], and is estimated by simple differencing of sequential values of [A] as decribed in the

following equation :

$$[\dot{A}] = \frac{\Delta[A]}{\Delta t} = \frac{A_i - A_{i-1}}{\Delta t}$$

Where Δt is the difference between the current time and the time when the previous sample is taken. The angular velocity vector is formed using the average of the two off-diagonal elements for each of the rates ω_x, ω_y, and ω_z.

Computation of Pitch Error Angle, and Error Rate

Both the pitch error angle and error rate are used in the computation of the momentum wheel control torque.

In vertical pointing control mode, the pitch error angle is computed using the following equation:

$$e = \tan^{-1}[-r(1)/r(3)]$$

where r(1) and r(3) are x and z components of the spacecraft position vector in the body frame, respectively. This vector is obtained by transforming the spacecraft position vector in the inertial frame to the body frame by using the transformation matrix [A].

In the orbit rate rotation mode, the spacecraft y-axis is desired to point at the sun while the z-axis rotates at one revolution per orbit in a plane perpendicular to the sun vector. At the same time, the z-axis is desired to point as close to north as possible at the northernmost point in the orbit, south at the southernmost point, and parallel to the equator at the equatorial crossings.

Let θ be the spacecraft orbit angle (measured from the northernmost point of the orbit), Figures 3 and 4 show the desired pointing direction for the z-axis for the two cases $\mathbf{S} \bullet \mathbf{N} > 0$ and $\mathbf{S} \bullet \mathbf{N} < 0$ respectively, where N is the orbit normal in the direction of $\mathbf{R} \times \dot{\mathbf{R}}$ and S is the inertial sun vector.

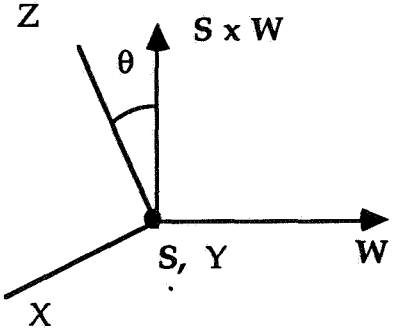


Figure 3. Desired Spacecraft Orientation for $\mathbf{S} \bullet \mathbf{N} > 0$

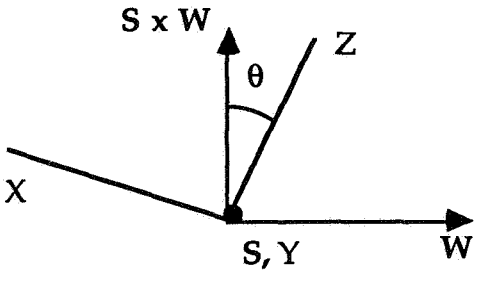


Figure 4. Desired Spacecraft Orientation for $\mathbf{S} \bullet \mathbf{N} < 0$

In these figures, W is a unit vector in the direction of $\mathbf{NP} \times \mathbf{S}$, where NP is a

north pole unit vector, $[0, 0, 1]^T$. An inertial target (U) for z-axis pointing will be given by :

$$\mathbf{U} = \cos\theta\, \mathbf{S} \times \mathbf{W} + \sin\theta\ \mathbf{W} \qquad \text{for } \mathbf{S} \bullet \mathbf{N} < 0$$

$$\mathbf{U} = \cos\theta\, \mathbf{S} \times \mathbf{W} - \sin\theta\ \mathbf{W} \qquad \text{for } \mathbf{S} \bullet \mathbf{N} > 0$$

where $$\mathbf{W} = \frac{\mathbf{NP} \times \mathbf{S}}{|\mathbf{NP} \times \mathbf{S}|}$$

The sine and cosine of the orbit angle are determined from the spacecraft position (**R**) and velocity ($\dot{\mathbf{R}}$) vectors. If **AN** is a unit vector in the direction of **NP** × **N** and locates the orbit ascending node, and **NMP** is a unit vector in the direction of **N** × **AN** and locates the northernmost point, then sinθ and cosθ are defined by:

$$\sin\theta = -\frac{\mathbf{R} \bullet \mathbf{AN}}{|\mathbf{R}|}$$

$$\cos\theta = \frac{\mathbf{R} \bullet \mathbf{NMP}}{|\mathbf{R}|}$$

The test for the sign of the dot product of the sun and the orbit normal vectors will automatically change the spin direction when the sun passes through the orbit plane. The pitch error angle for control can be computed using:

$$e = \tan^{-1}(-u(1)/u(3))$$

where **u** is obtained by transforming the target vector **U** in the inertial frame to the body frame by using the transformation matrix [**A**] .

The pitch error rate is estimated by differencing of sequential values of pitch error angle.

IV. CONTROL LAWS

In the two science control modes, z-axis pointing is accomplished by controlling the momentum wheel speed, and sun pointing is accomplished by magnetic torquing. In addition to pointing control, the magnetic torquer bars provide nutation damping.

When the sun vector is parallel to the earth's magnetic field vector (at singularity), the momentum wheel speed is held constant at the speed before singularity occurs, and all torquer bars are turned off.

During eclipse, all magnetic torques are turned off, and momentum wheel is controlled using the normal control law except in the singular case described above.

In the sun pointing only control mode, the momentum wheel speed is held at a commandable rate to provide a momentum bias, and the magnetic torquer bars are activated for sun pointing as well as nutation damping.

Momentum Wheel Control Law for Science Pointing

The goal of controlling the momentum wheel speed is to control the motion of the spacecraft about pitch axis (y-axis), i.e. z-axis pointing control. The torque to the momentum wheel is driven by rate and position error:

$$T_y = I_y(\omega_n^2 e + 2\xi\omega_n \dot{e})$$

where I_y is inertia (N-m-sec^2), ω_n is the control frequency and is set to 0.01 (rad/sec), ξ is a damping constant and is set to 0.707, e is the pitch error angle (rad), and $\dot{e}$ is the pitch error rate (rad/sec). Both e and $\dot{e}$ are products from the attitude determination algorithm.

Magnetic Control for Science Pointing

SAMPEX attitude control involves a momentum bias along the spacecraft y-axis and the pointing of this axis at the sun. Let H_o be a desired level of angular momentum, with **j** a y-axis unit vector and **s** the sun unit vector in the body frame. Ideally, we would like to have the system angular momentum vector (**H**) equal to both H_o **j** and H_o **s**. This will be true only if the y-axis is pointed at the sun, momentum is at the desired level and there is no nutation.

We now consider two momentum error vectors, $\mathbf{H} - H_o\,\mathbf{j}$ and $\mathbf{H} - H_o\,\mathbf{s}$ and add them together to form an" excess" momentum vector to be "unloaded" by the magnetic torquing system. Let $\Delta\mathbf{H} = 2\mathbf{H} - H_o(\mathbf{j} + \mathbf{s})$. One common method of momentum unloading is to let $\mathbf{dm} = \Delta\mathbf{H} \times \mathbf{b}$ where **dm** is an applied magnetic moment, $\Delta\mathbf{H}$ is the undesired system momentum and **b** is the magnetic field vector in the body frame. In our case, we let

$$\mathbf{dm} = \text{Mag\_gain}\ (\Delta\mathbf{H} \times \mathbf{b})$$

where Mag_gain - an appropriate control gain (A-m^2/N-m-sec-tesla)
$\Delta\mathbf{H}$ - undesired system momentum (N-m-sec)
b - Measured earth's magnetic field (tesla)

In more detail, this becomes

$$dm = Mag\_gain \begin{vmatrix} i & j & k \\ 2H_x - H_o\, s_x & 2H_y - H_o\,(1+s_y) & 2H_z - H_o\, s_z \\ b_x & b_y & b_z \end{vmatrix}$$

Sun Point Only Mode (Software Safe Hold)

The sun pointing only mode is a non-science control mode. It is executed at computer power on, or when the spacecraft pitch axis drifts away from the sun line by more than 15 degrees.

The goal of this control mode is to remove any excess spacecraft body spin rates, and to precess the spacecraft pitch axis to within 15 degrees of the sun line. In this mode, there is no pitch axis control and the momentum wheel speed is held constant. The torque signal in N-m is given by:

$$T_y = G_{mw}(fixed\_rate - v)$$

where v is the momentum wheel speed (rad/sec), and G_{mw} is momentum wheel gain in N-m-sec/rad.

The roll and yaw torquer bars are used for spin rate control (spin control) and momentum unloading (Bdot control). In addition to these two goals, the pitch axis torquer bar is used to precess the pitch axis toward the sun line (y-axis precession control). The following equation describes these three controls:

$$\mathbf{dm} = -G_{Bdot}(\mathbf{b}-\mathbf{b_p})/\Delta t + SC_y G_s \begin{bmatrix} b(3) \\ 0 \\ b(1) \end{bmatrix} + SC_z G_s \begin{bmatrix} b(2) \\ b(1) \\ 0 \end{bmatrix} + SIGN(bs) G_p F_e F_p \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

where
- G_{Bdot} - Bdot control gain (A-m^2-sec/tesla)
- G_s - spin control gain (A-m^2/tesla)
- G_p - y-axis precession gain (A-m^2)
- bs - the second element of **b**×**s**, **s** is the body sun vector and **b** is the measured earth's magnetic field vector (tesla)
- $\mathbf{b_p}$ - previous measured earth's magnetic field vector (tesla)
- SC_y - y-axis spin control flag, values can be -1, 0, 1 corresponding to despin, off, spin
- SC_z - z-axis spin control flag, values can be -1, 0, 1 corresponding to despin, off, spin
- F_e - eclipse flag, 1 = no eclipse, 0 = eclipse
- F_p - y-axis precession flag, 1 = y-axis precession enable

0 = y-axis precession disable

Δt - the difference between the current time and the time when the previous sample is taken (sec)

The first term of the above equation describes the Bdot control for the momentum unloading. The dipole moment for each torquer bar is proportional to the rate of change of the measured earth's magnetic field along each axis. The second and third terms describe the spin control with the second term for y-axis spin and the third term for z-axis spin. These two spin controls will not be on simultaneously. The control options spin, despin, or off are selected by ground command. The last term describes y-axis precession control. The dipole moment for this control is driven by a gain with a polarity determined by the cross product of the measured sun and magnetic field vectors. The precession dipole moment is not computed when the spacecraft is in eclipse and can be disabled by ground command. All gains are modifiable via ground command.

V. SIMULATION

The dynamic simulator has two parts: 1) system state model (SSM) and 2) on-board attitude determination and control. The on-board attitude control and determination has been discussed in sections II, III, and IV. The SSM includes the kinematics (including the quaternion and direction cosine matrix), the equations of motion, and the sensor models. Figure 5 below describes the dynamic simulator.

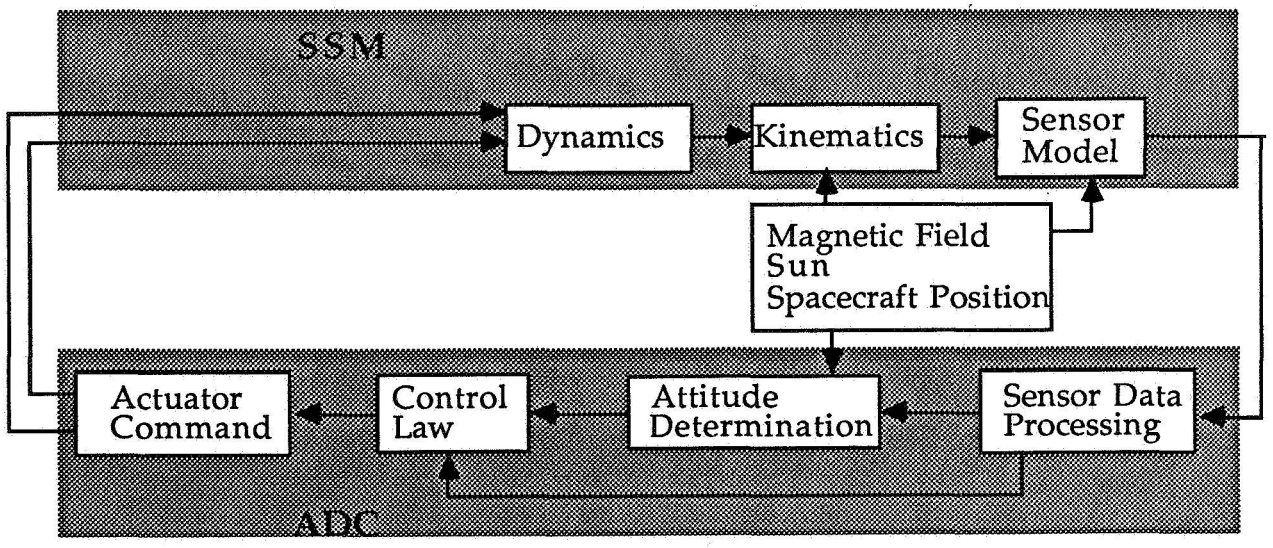


Figure 5. The Dynamic Simulator

Runge-Kutta integration on the derivatives of the system momentum and wheel momentum, the spacecraft position and velocity vectors, and the quaternion.

Momentum

The system momentum is computed by integrating Euler's equation [2]:

$$\dot{\mathbf{H}} = \mathbf{T}_{aero} + \mathbf{T}_{gg} + \mathbf{T}_{mu} - \boldsymbol{\omega} \times \mathbf{H}$$

where:
- $\mathbf{H}$ - system angular momentum (N-m-sec)
- $\mathbf{T}_{aero}$ - aerodynamic torque (N-m)
- $\mathbf{T}_{gg}$ - gravity gradient torque (N-m)
- $\mathbf{T}_{mu}$ - magnetic unloading torque (N-m)
- $\boldsymbol{\omega}$ - spacecraft rate (rad/sec)

The Pictorial Solar Pressure (PSP) program is used to compute the aerodynamic torques on SAMPEX. PSP takes into account the effects of shadowing on the spacecraft. The density is selected for a worst case solar cycle.

$$T_{aero} = conv[(1/2)\rho V^2 C_d T_{norm}]$$

where:
- conv - conversion factor $\left(16387.3 \frac{N-m-sec^2}{(gm/cm^3)km^2\ in^3}\right)$
- ρ - atmospheric density (gm/cm^3)
- V - orbit velocity (km/sec)
- C_d - coefficient of drag
- T_{norm} - normalized torque output from PSP (in^3)

The gravity gradient torque is computed by [2]:

$$\mathbf{T}_{gg} = 3(\mu/|\mathbf{R}|^3)[\mathbf{r}_b \times ([I] \cdot \mathbf{r}_b)]$$

where:
- μ - earth's gravitational constant ($3.98601 \times 10^5\ km^3/sec^2$)
- $\mathbf{r}_b$ - zenith vector in body frame
- $|\mathbf{R}|$ - magnitude of radius vector (km)
- I - inertia tensor ($N\text{-}m\text{-}sec^2$)

The magnetic unloading torque is determined from:

$$\mathbf{T}_{mu} = k_1 M_{lim} \times \mathbf{B}$$

where $\mathbb{B}$ is the earth's magnetic field strength in tesla computed using a tenth order spherical harmonic magnetic field model [2] and M_{lim} is the limited magnetic dipole moment with a torquer bar limit of 20 A-m^2. The magnetic dipole moment is determined from the sun pointing and nutation damping error vectors discussed earlier in Section IV. The constant k_1 is a conversion factor (1.0 N-m/A-m^2-tesla).

The $\boldsymbol{\omega} \times \mathbf{H}$ term is the gyroscopic reaction torque. It is the cross product of the body rates and the system angular momentum vector.

The system angular momentum is determined from the following integral

$$\mathbf{H} = \int \dot{\mathbf{H}}\, dt = \int (\mathbf{T}_{ext} - \boldsymbol{\omega} \times \mathbf{H}) dt$$

where T_{ext} is the sum of the external torques on the spacecraft.

Similarly, the reaction wheel angular momentum is the integral of the applied wheel torque

$$h_w = \int (T_w + T_{friction}) dt$$

where T_w is the commanded wheel torque and $T_{friction}$ is the torque due to friction.

Quaternion

The quaternion obeys the kinematic equation of motion [2]:

$$\dot{q} = (1/2)\Omega q$$

where $\Omega = \begin{bmatrix} 0 & \omega_z & -\omega_y & \omega_x \\ -\omega_z & 0 & \omega_x & \omega_y \\ \omega_y & -\omega_x & 0 & \omega_z \\ -\omega_x & -\omega_y & -\omega_z & 0 \end{bmatrix}$ is the skew-symmetric matrix

$$\omega = [I]^{-1}(H - h_w j)$$

h_w - wheel momentum (N-m-sec)

j - unit vector along the y-axis.

Using this quaternion, a direction cosine matrix can be generated:

$$\text{AMATRIX} = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 + q_3q_4) & 2(q_1q_3 - q_2q_4) \\ 2(q_1q_2 - q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 + q_1q_4) \\ 2(q_1q_3 + q_2q_4) & 2(q_2q_3 - q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix}$$

This matrix and its transpose are used to transform vectors from the inertial frame to the body frame and vice versa. This process is used in the sensor models in the simulation.

Sensors

The spacecraft ephemeris is generated by integrating the equations of motion including the perturbations due to nonspherical earth gravity effects and aerodynamic forces.

The sun vector in inertial frame is used to generate the sun vector in body frame and for attitude determination. It is computed using 1985 coefficients.

The magnetometer reading includes the earth's magnetic field, computed using a 10th order spherical harmonic magnetic field model [2], and the magnetic field produced by the torquer bars. The contamination due to the torquer bars is approximated by modelling the torquer bar as a dipole. For thin cylindrical bar magnets, the distance between the two poles of the dipole is approximately five-sixths the length of the magnet [6] (the torquer bar in our case). A contamination matrix is formed based on the contamination due to each bar in each of the axes of the three-axis magnetometer.

Results

The dynamic simulator has been used to simulate both the vertical pointing mode and the orbit rate rotation mode for various orbit configurations. These configurations include 6 PM, 9 PM, and midnight orbits, where the time is the local time at the ascending node. The simulator also examines how the relative motion of the sun and the earth throughout the year effects such things as the sun pointing error and pitch loop position error.

Plots of the sun pointing error and the pitch loop position error for 6 PM, 9 PM, and midnight orbits are shown in Figures 6, 7, and 8, respectively, for the vertical pointing mode and in Figures 9, 10, and 11 for the orbit rate rotation mode. These simulation runs are for the winter solstice time of year. The results of the simulation show that the five degree sun pointing requirement is met for all cases investigated for both control modes.

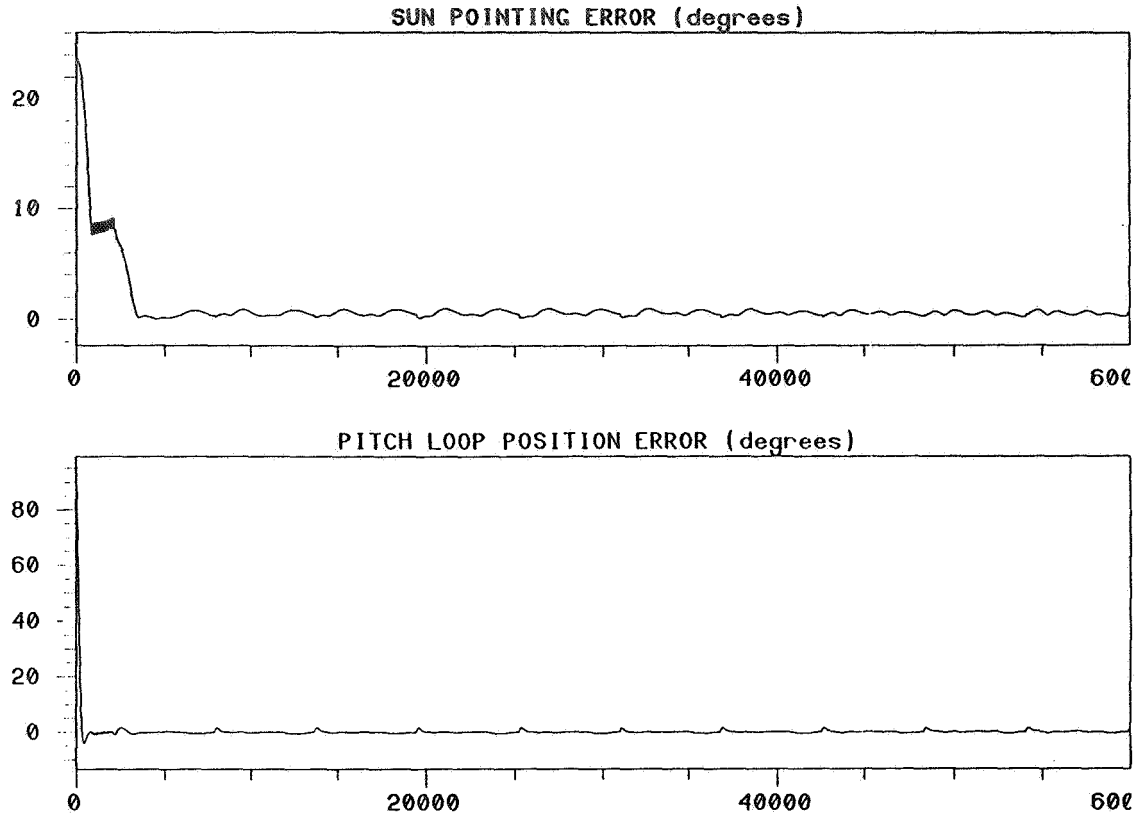


Figure 6. 6 PM Orbit - Vertical Pointing Mode

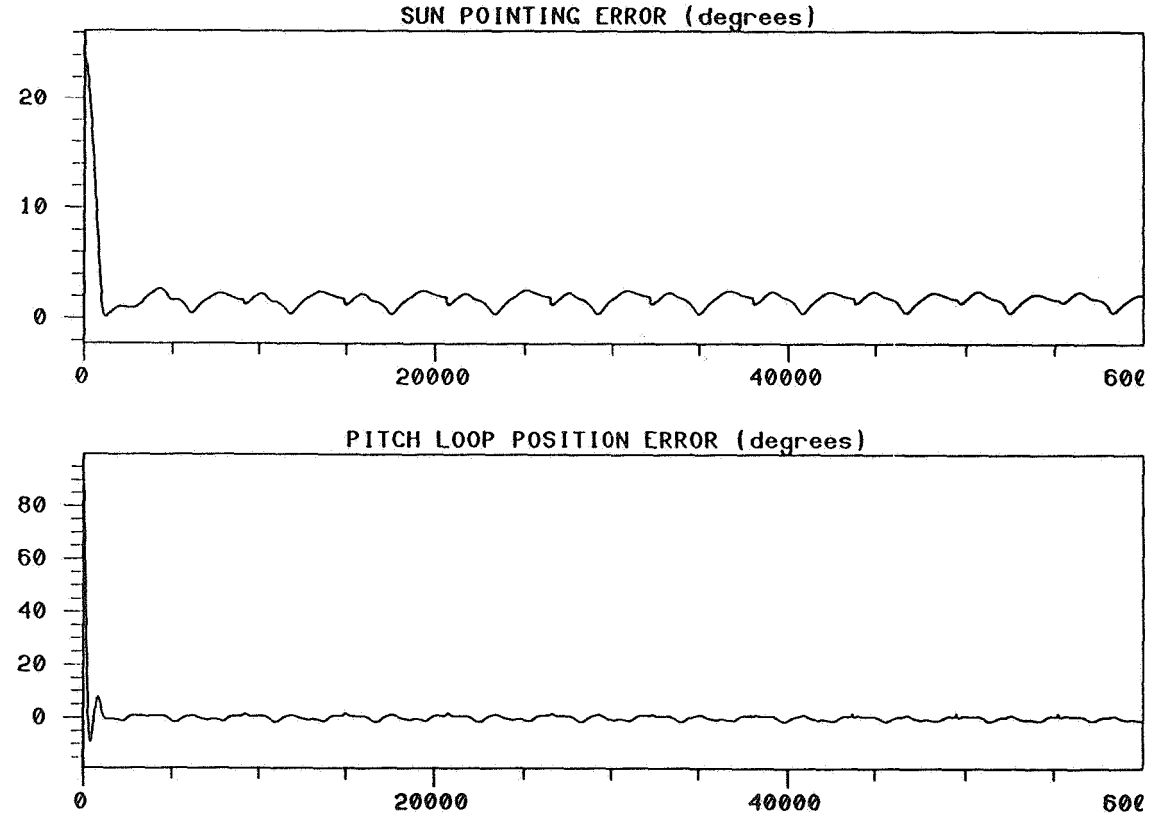


Figure 7. 9 PM Orbit - Vertical Pointing Mode

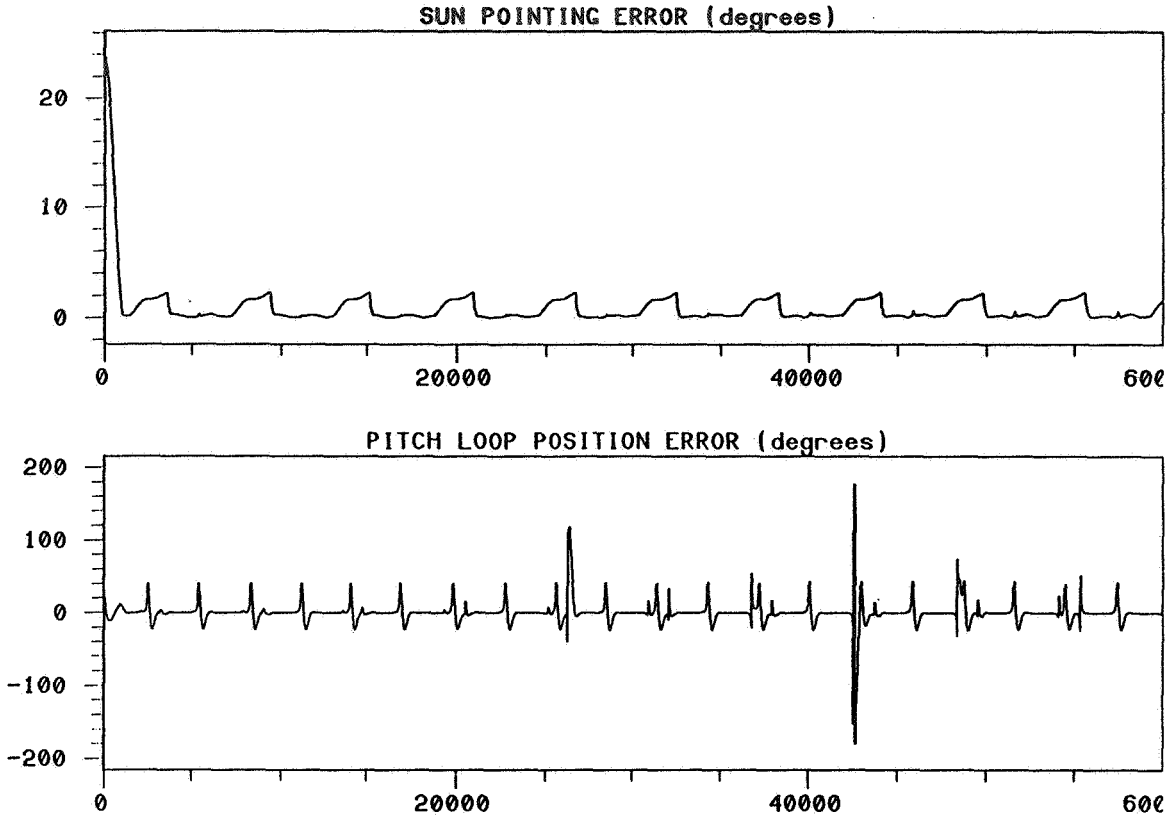


Figure 8. Midnight Orbit - Vertical Pointing Mode

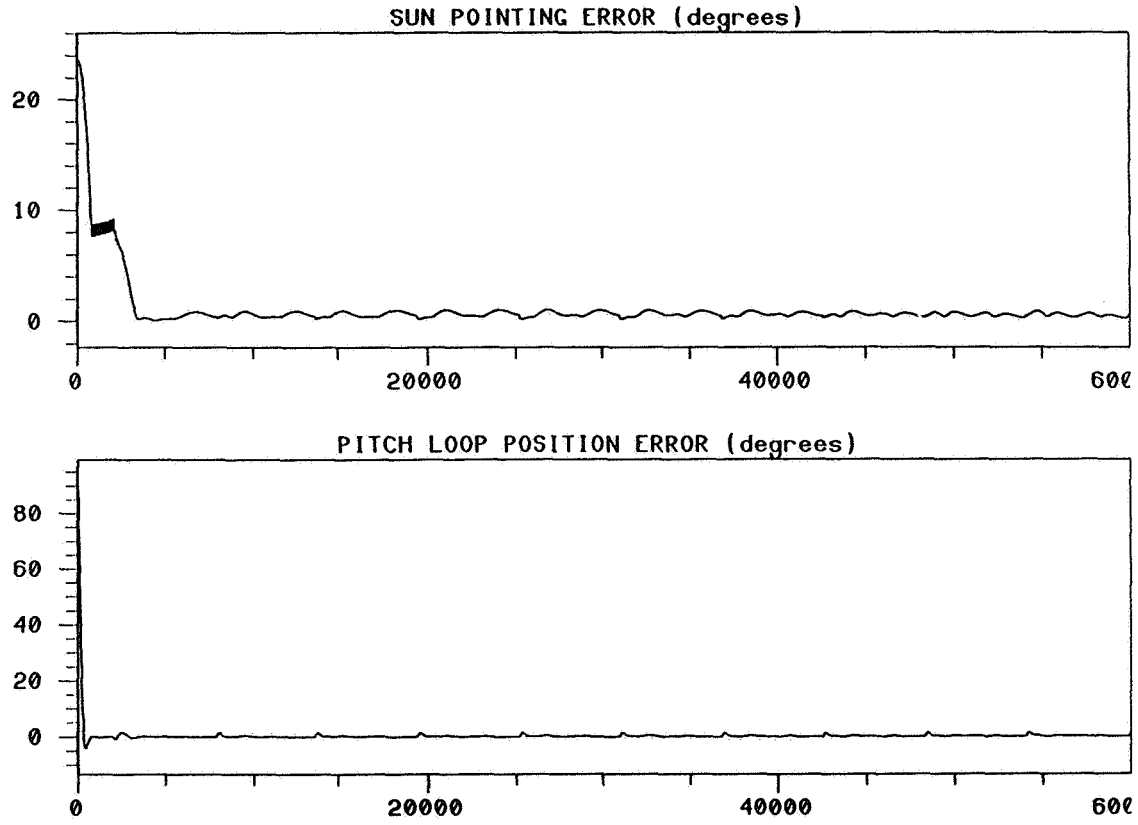


Figure 9. 6 PM Orbit - Orbit Rate Rotation Mode

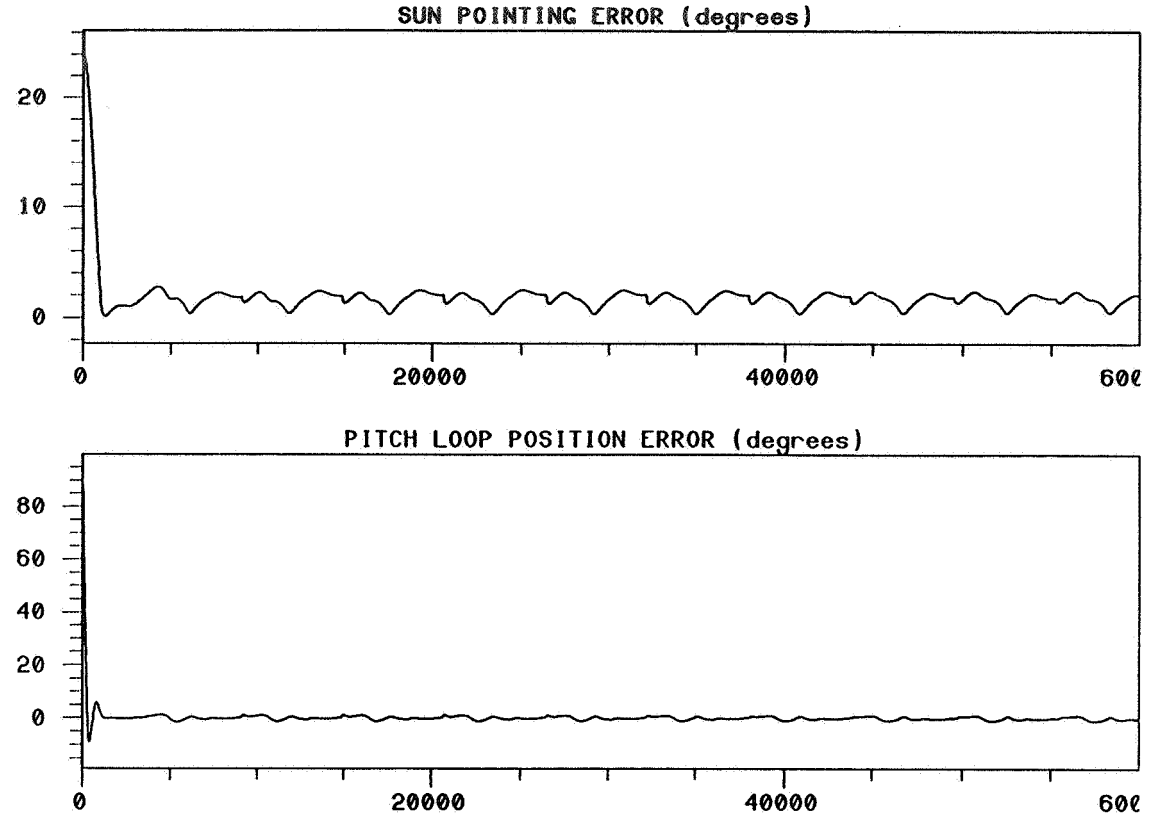


Figure 10. 9 PM Orbit - Orbit Rate Rotation Mode

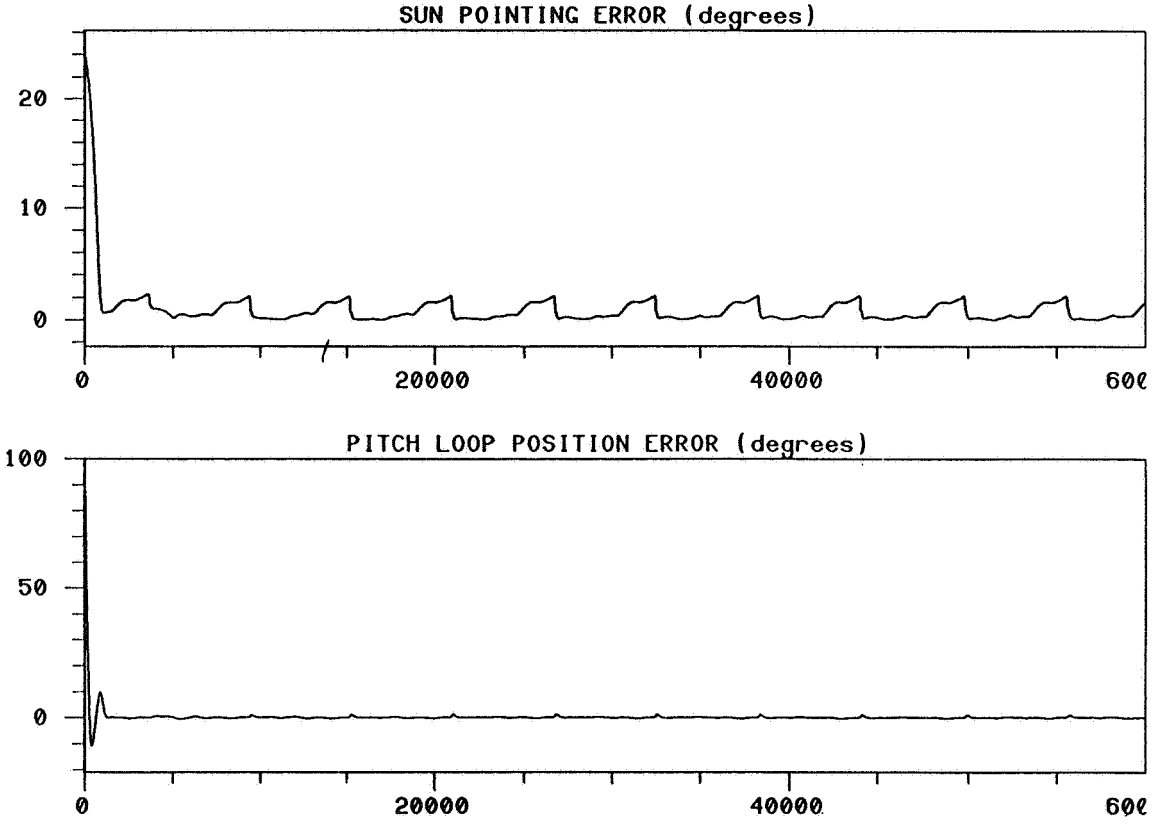


Figure 11. Midnight Orbit - Orbit Rate Rotation Mode

VI. SUMMARY

The simulation results show that the on-board attitude determination and control algorithm satisfies the attitude knowledge and control requirements. Since the orbit rate rotation mode satisfies the spacecraft pointing requirement as well as the velocity vector avoidance criterion, this mode is recommended as the base control mode for normal operation.

ACKNOWLEDGEMENT

We would like to acknowledge the assistance of John Azzolini, Jim Donohue, Henry Hoffman, the Guidance and Control Branch, and the SMEX project at Goddard Space Flight Center for their contributions to the SAMPEX project. Their input aided us in the development of the software algorithm and the design andanalysis of the dynamic simulator.

REFERENCES

[1] *Proposal for a Solar, Anomalous, and Magnetospheric Particle Explorer* (SAMPEX) , submitted to NASA in response to A. O. No. OSSA 2-88 for the Small-Class Explorer Mission, September 1988

[2] Wertz, James R. (editor), *Spacecraft Attitude Determination and Control,* 1986

[3] Report pending under NAS5-39412, Task 18

[4] GSFC, Flight Dynamics Division Document No. FDD/552-89/001

[5] Shuster, M. D. and Oh, S. D., *Three-axis Attitude Determination from Vector Observations,* Journal of Guidance and Control, 1981 Jan-Feb, P70-77

[6] Lemon, Harvey Brace and Ference, Michael, Jr, *Analytical Experimental Physics,* P244

Attitude-Dependent Launch Window Analysis for the Hubble Space Telescope Mission

James A. Lenio

Rockwell Space Operations Company

ABSTRACT

Launch window analysis for Space Shuttle missions determines the launch times which will ensure that all payload and Shuttle requirements for the mission are met. Attitude and pointing analysis determines Shuttle Orbiter attitudes that meet various communication, viewing, and thermal requirements for the Orbiter and its payloads. Historically, launch window analysis and attitude and pointing analysis for Shuttle missions have been done separately, without directly influencing each other. However, methods have been developed to consider simultaneously dependencies between launch window and attitude and pointing requirements if they arise. These methods were developed from the launch window analysis for STS-31, the Hubble Space Telescope (HST) deployment mission.

To release the HST, the Orbiter attitude had to remain inertially fixed while pointing the HST at the Sun. The Orbiter release attitude and the HST release time were determined from the position of the Sun and varied with launch time and launch date. The launch window analysis for STS-31 centered on how to determine the range of launch times for a given launch date that would allow the Shuttle to release the HST and simultaneously satisfy communication, attitude, and lighting requirements for the deployment operations.

This paper discusses how the HST deployment requirements determined the launch window and how the Orbiter release attitude affected the launch window.

INTRODUCTION

Launch Window and Attitude and Pointing Analysis for Shuttle Missions

The orbital trajectory of the Space Shuttle is made up of a series of events that must occur at precise times to meet the objectives of the mission. The timing of these events usually originates from requirements to operate or deploy Shuttle payloads at locations that ensure certain orbital conditions are met, such as communication, viewing, lighting, and thermal exposure. These orbital conditions usually have a certain geometrical relationship relative to Earth-fixed targets or celestial targets. The times that these mission events occur are also influenced by crew schedules, Shuttle performance limitations, and duration of the mission. For all planned events to occur in sequence and satisfy the objectives of the mission, the Shuttle can launch only at certain predetermined times.

A launch window for a particular Shuttle payload is the range of launch times, on a given launch date, that achieve the payload's orbital objectives and satisfy the crew, Shuttle, and mission requirements and constraints. The launch window for the Hubble Space Telescope (HST), for example, is constrained by several objectives, such as deploying the satellite on a particular orbit, ensuring orbital lighting at deployment, pointing the HST at the Sun at release, and maintaining communication with the Tracking and Data Relay Satellite System (TDRSS) during deployment.

Attitude and pointing analysis is another specialized area of Shuttle orbital flight analysis. Attitude and pointing analysis determines Shuttle attitudes that satisfy the communication, viewing, lighting, thermal, and microgravity requirements of the Orbiter and its payloads at certain times during a mission. Such analysis, for example, determines the attitude that is required at a certain time to point the Orbiter's star trackers and recalibrate the onboard navigation platform. This pointing analysis is necessary because of the Orbiter's own operational requirements. But payloads may also require specific Orbiter attitudes at certain times during a mission, and they may require their own special attitude and pointing analysis. The attitude and pointing analysis for the HST mission determined the Orbiter attitudes that were needed to make contact with TDRSS and still point HST at the Sun before its release from the Orbiter's robotic arm.

Orbiter attitudes are usually determined separately from the orbital times required for deploying or operating Shuttle payloads; therefore, attitude and pointing requirements are usually independent of launch time. However, attitude and pointing analysis for STS-31, the Shuttle mission to deploy the HST, showed that Orbiter attitudes affect the orbital conditions required at HST deployment, and that the availability of these conditions due to Orbiter attitudes varied with launch time. Special analysis was required to analyze this interdependency between Orbiter attitude requirements and launch time.

HST PAYLOAD REQUIREMENTS FOR DEPLOYMENT

Three deployment requirements dominated the launch window analysis for the HST mission: (1) there had to be communication with TDRSS throughout the deployment period, (2) the Orbiter had to be in a particular attitude to deploy HST and separate from it, and (3) HST release had to occur in orbital daylight.

Communication Constraints

For ground control to uplink commands and receive telemetry from HST during the deployment operations, the Orbiter had to relay Ku-band transmissions through one of the two TDRSS satellites, TDRS-West or TDRS-East. Since these geosynchronous satellites are fixed with respect to the Earth, the orbiting Shuttle continually moves into and out of contact with each satellite, creating certain times of TDRSS acquisition of signal (AOS) and loss of signal (LOS) due to Earth occultation. The release of HST had to occur at least 5 minutes after AOS and at least 25 minutes before LOS on a given TDRSS pass. The 5 minutes before release were needed for the Orbiter's directional Ku-band antenna to begin tracking the TDRSS satellite. The 25 minutes after release were needed to ensure the HST maneuvered to a stable attitude.

Attitude Constraints

While HST was in its release position on the Orbiter's robotic arm, the Orbiter was required to point the satellite's aft end, or -V1 axis, at the Sun. The aft end contains coarse Sun sensors that allow HST's attitude control system to determine and stabilize the satellite's orientation after release. The release orientation was further constrained by the need to protect the telescope during the Orbiter's separation burn that occurred 2 minutes after release. To prevent recontact with the HST or contaminating the HST with the Orbiter jet plumes during the burn, the Orbiter's Z-body axis had to remain in the orbit plane and point in the direction of orbital motion at the time of HST release.

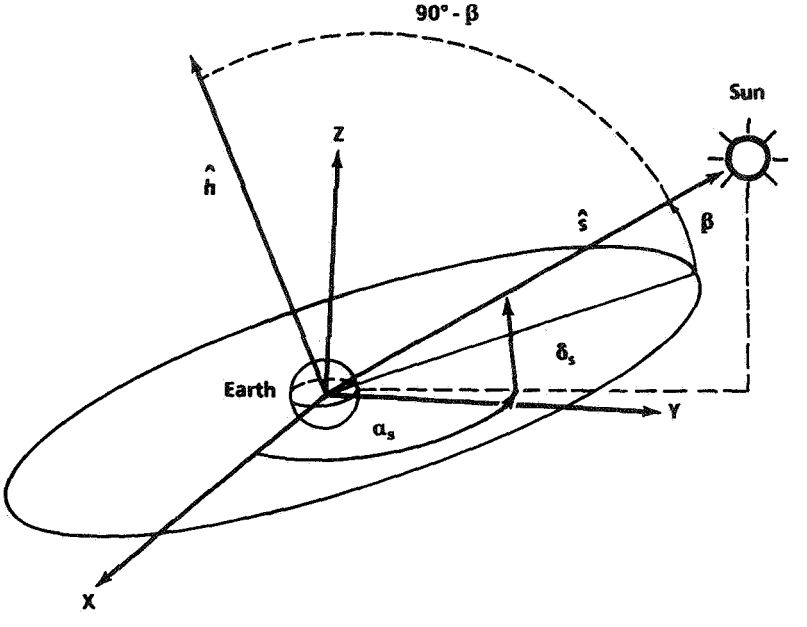
Lighting Constraints

To ensure that HST's solar arrays were deployed and to verify the Orbiter's safe separation from the HST, the Shuttle crew had to be able to observe the telescope in orbital daylight before and after its release. To meet this lighting requirement throughout the deployment sequence, the release was timed to coincide approximately with the location of orbital noon. Since orbital noon is approximately halfway into the daylight portion of an orbit, HST release near orbital noon would provide sufficient lighting during the deployment operations.

ANALYSIS AND DEVELOPMENT OF THE HST LAUNCH WINDOW

The primary objective of the STS-31 mission was to release HST on the earliest deployment opportunity after satisfying HST checkout requirements and crew activity constraints. This opportunity occurred on orbit 19. The actual time of release on orbit 19 was based on the combination of the communication, attitude, and lighting requirements that constrained release.

First, HST release had to occur while the Orbiter was in contact with one of the TDRSS satellites on orbit 19. Second, the release time also had to satisfy the specific attitude and lighting constraints during this period of contact. To determine the actual release time after orbital noon, a formula was derived in which the incidence angle between the Sun and the orbit plane was the only variable. This angle is generally called the solar Beta angle and is illustrated in the HST release time derivation in Figure 1.



α, δ_s = right ascension and declination to the Sun, respectively, measured in the inertial X-Y-Z plane

$\hat{h}$ = the direction of the orbital angular momentum vector, perpendicular to the orbit plane

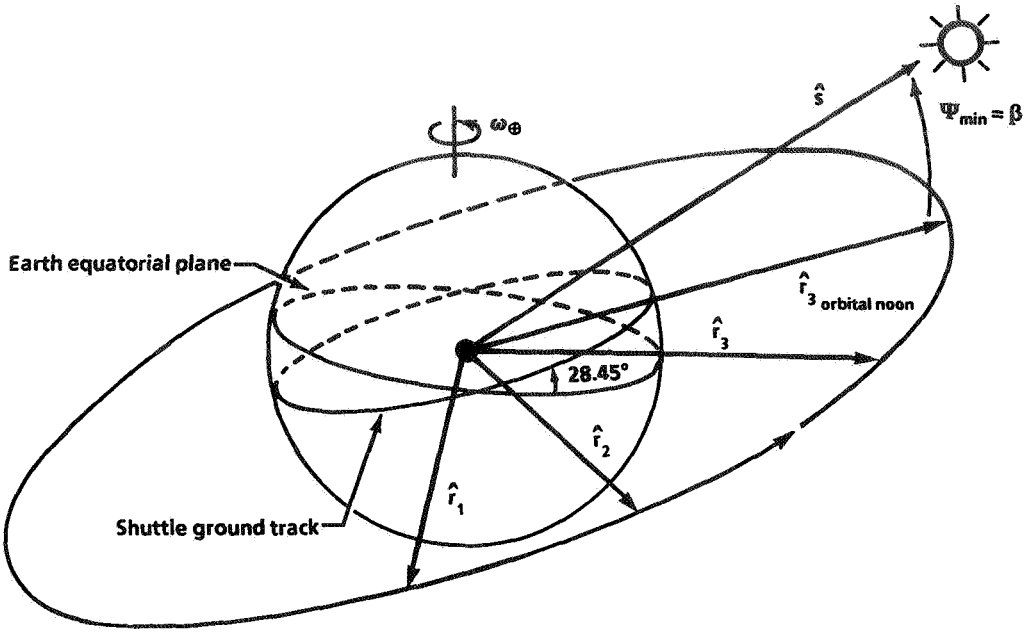
$\hat{s}$ = the direction of the Sun pointing vector from the center of the Earth

β = the solar Beta angle measured from the orbit plane to the Sun pointing vector

$$\beta = 90° - \cos^{-1}(\hat{h} \cdot \hat{s})$$

(a) Definition of solar Beta angle.

Figure 1.– Derivation of the HST release time.



$\hat{r}$ ≡ the direction of the position vector for the Shuttle at a particular time

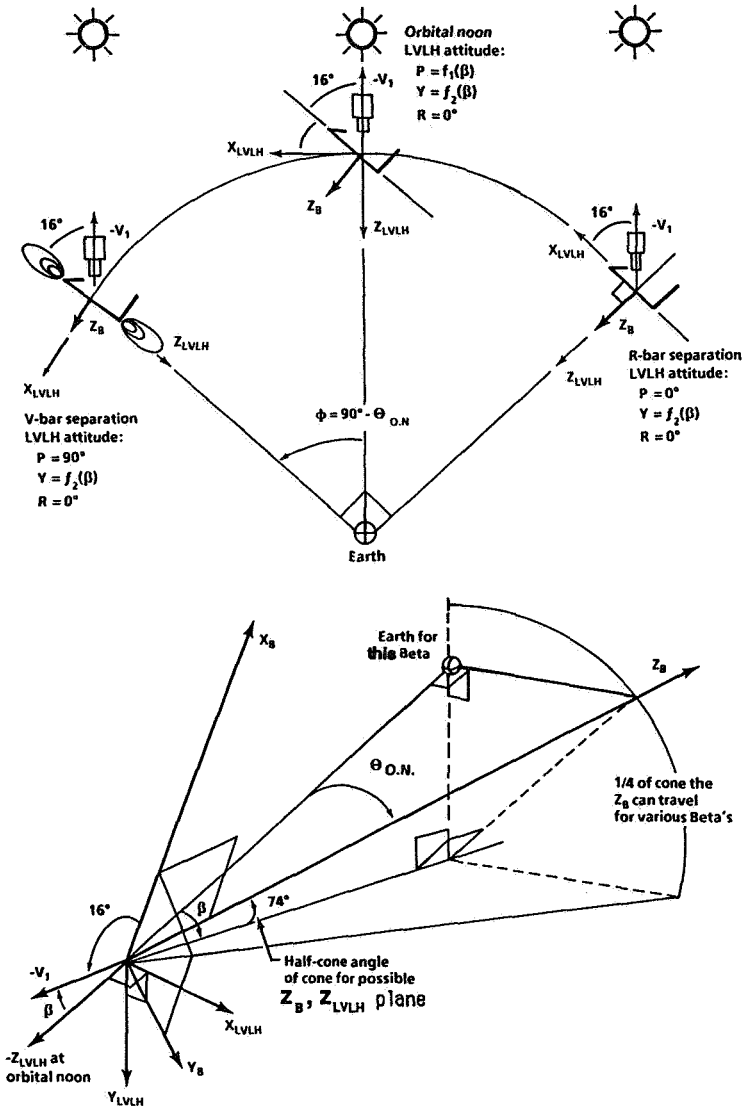
Ψ = Sun angle ≡ the geometric angle between the Shuttle's position vector, $\hat{r}$, and the Sun pointing vector, $\hat{s}$

Orbital noon occurs when the Sun angle is at a minimum:

$$\Psi_{min} = |\beta|$$

(b) Definition of orbital noon.
Figure 1.– Continued.

LVLH ≡ Local Vertical, Local Horizontal coordinate system whose origin is at the spacecraft's center of mass. The z-axis points towards the Earth along the spacecraft's position vector, the y-axis is in the opposite direction of the orbital angular momentum vector, and the x-axis completes the orthogonal coordinate system. The Euler rotation sequence for spacecraft attitudes referenced to the LVLH system is usually pitch, yaw, roll (P,Y,R)



LVLH pitch angle of Orbiter at orbital noon is:

$$\theta_{O.N.} = \cos^{-1}\left[\frac{\cos 74^\circ}{\cos \beta}\right] = f_1(\beta)$$

(c) Derivation of the Orbiter LVLH pitch angle at orbital noon.
Figure 1.- Continued.

Central angle from orbital noon to the V-bar separation is:

$$\phi = \theta_{sep} - \theta_{O.N.} = 90° - \cos^{-1}\left[\frac{\cos 74°}{\cos \beta}\right]$$

where β = solar Beta angle

θ_{sep} = Euler pitch angle of the Orbiter, in the LVLH coordinate system, at the V-bar separation; equal to 90°

$\theta_{O.N}$ = Euler pitch angle of the Orbiter, in the LVLH coordinate system, at orbital noon

ϕ = central angle from orbital noon to the V-bar separation

The time from orbital noon to the V-bar separation is:

$$\Delta t_{sep} = \frac{\phi}{360°} \cdot T$$

where ϕ = central angle from orbital noon to the V-bar separation

T = orbital period

Δt_{sep} = the time from orbital noon to the V-bar separation

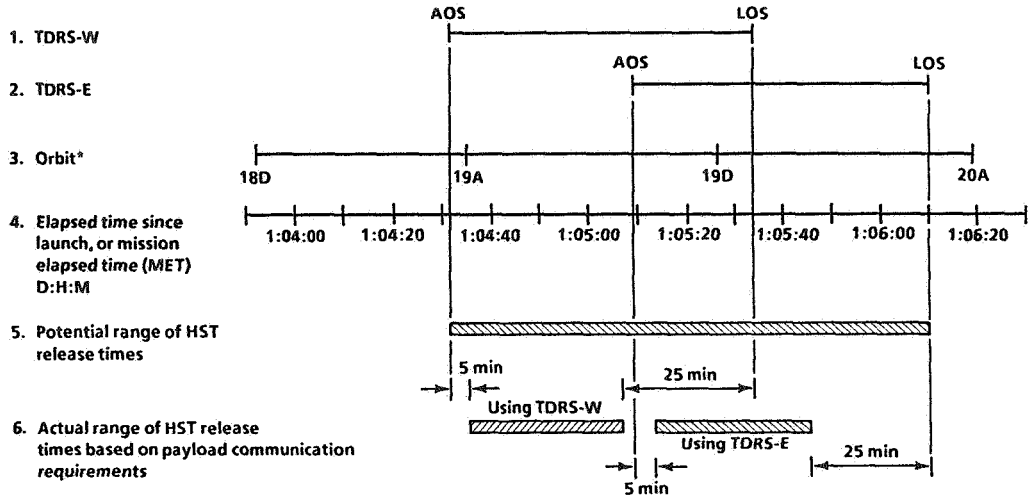
Finally, the time from orbital noon to HST release is:

$$\Delta t_{release} = \Delta t_{sep} - 2 \text{ min}$$

(d) Derivation of the central angle and time between orbital noon and HST release.

Figure 1.– Concluded.

The launch window was calculated from the range of release times on orbit 19 that simultaneously satisfied the communication, attitude, and lighting requirements for the HST deployment. Figure 2 shows how each of these requirements combine to produce the range of acceptable release times from which the launch window can be determined.



*18D corresponds to the descending equatorial crossing, or descending node, for orbit 18.
19A corresponds to the ascending node for orbit 19.

Figure 2.– Acceptable HST release times on orbit 19.

For an orbital trajectory with a given geometry and orientation, the variation in launch time over a range of possible launch days will depend on how the orbital conditions required for payload operations change with time. Because the TDRSS satellites are geosynchronous, their orbital period is equal to the Earth's sidereal day, and communication with TDRSS is considered contact with an Earth-fixed target. Since TDRSS is an Earth-fixed target, the mission elapsed time of TDRSS acquisition on a given orbit does not vary with launch time. However, the position of the Sun, and therefore orbital lighting, does vary with time, so the mission elapsed time of HST release on orbit 19 also varied with launch time and date. This variation of release time due to a change in launch time can be seen in Figure 3.

By simulating the trajectories resulting from various launch times over a 24-hour period, the launch window analyst can determine all possible launch times that produce the required orbital conditions at the desired time of payload operations. For the HST mission, the launch time was varied on a specified launch date until the release time, calculated from the orbital noon time and solar Beta angle, matched an acceptable time from the HST deployment timeline in Figure 2. This day-long launch window would ensure that the deployment conditions were met during the allowable period of release times on orbit 19.

The variation in launch time over a day or period of days is typically displayed as a graph showing the opening and closing of the launch window that meets the payload requirements. The day-long and year-long HST launch windows in Figure 4 show how the HST deployment requirements that open and close the launch window change with time.

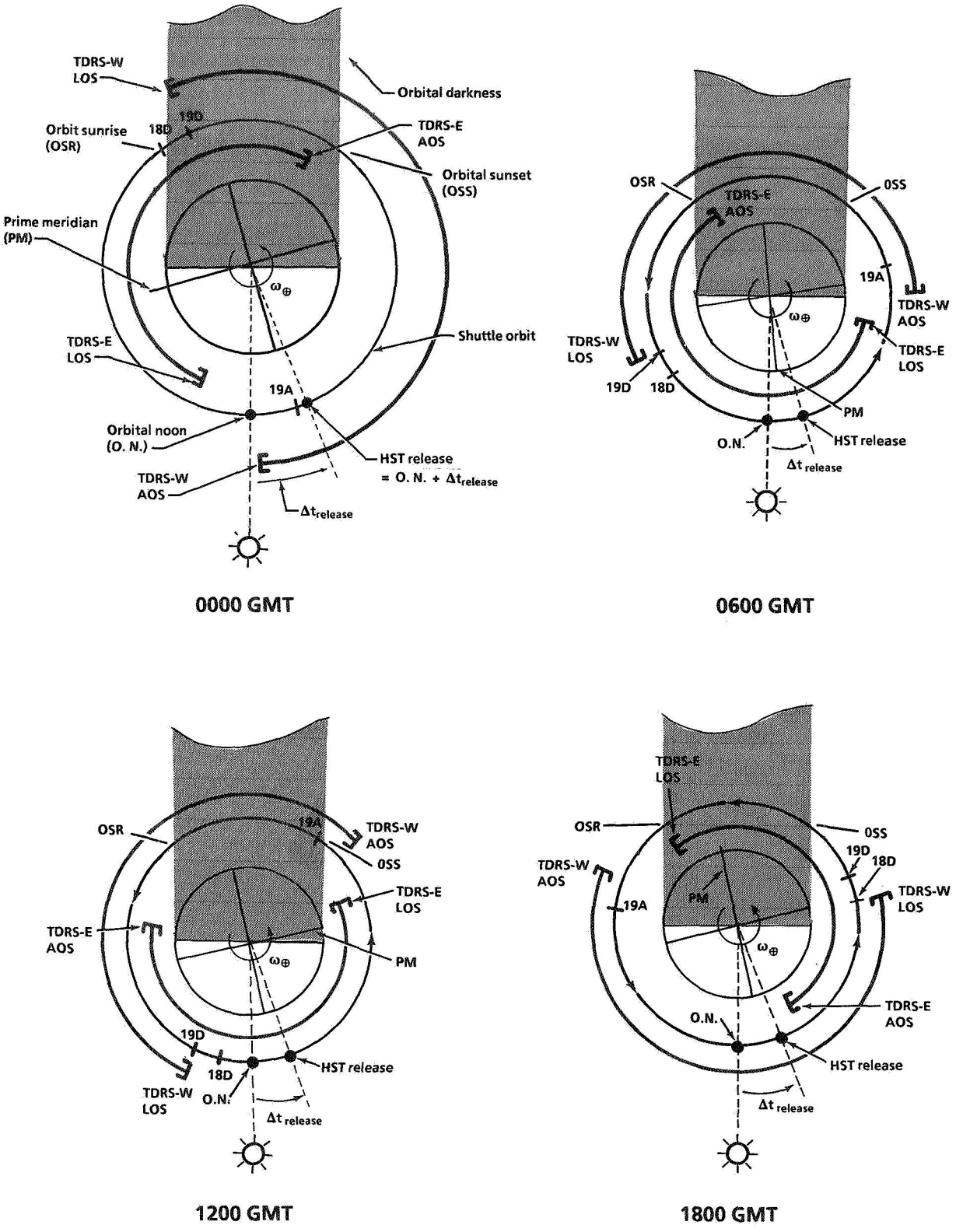
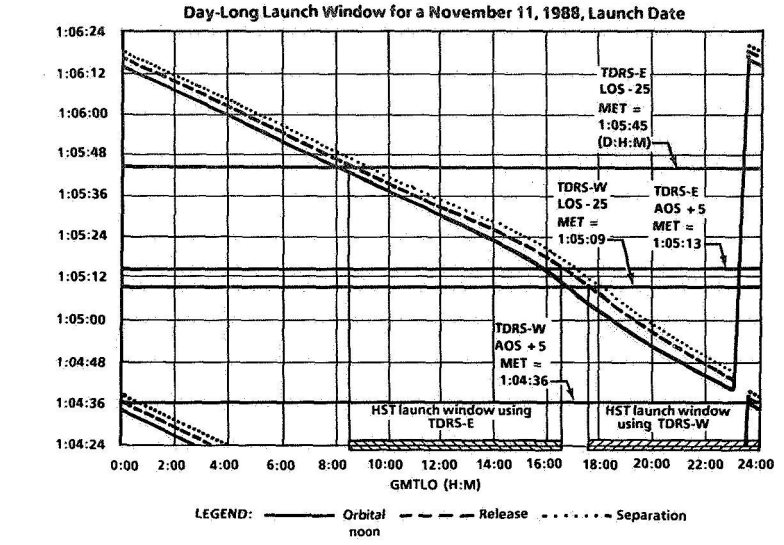
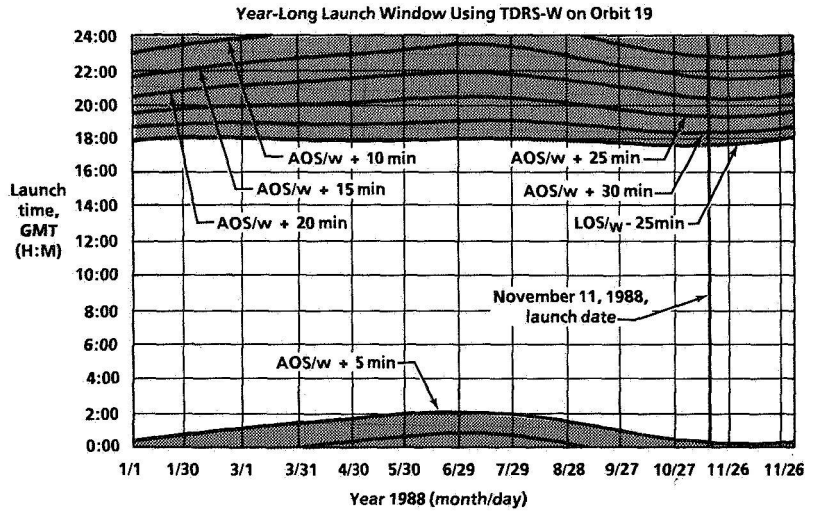


Figure 3.– Change in HST release time relative to TDRSS for various November 11, 1988, launch times.





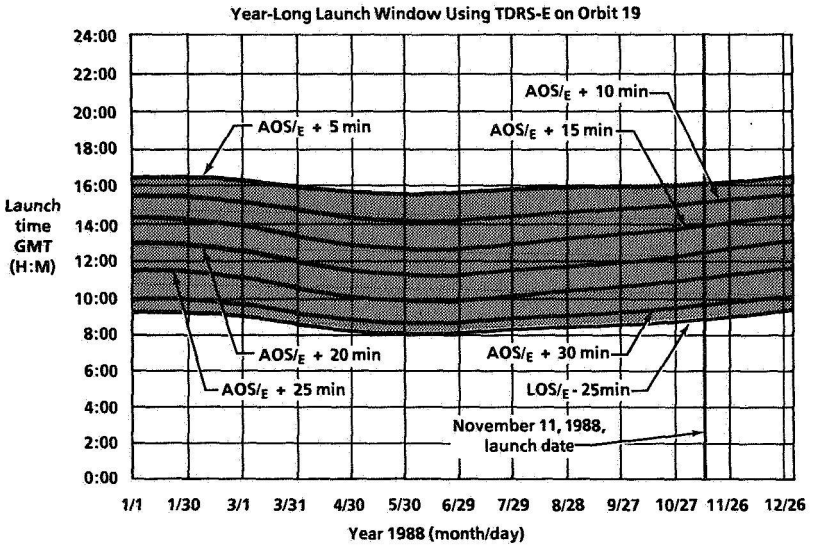


Figure 4.– Day-long and year-long HST launch windows for an orbit 19 deployment.

EFFECT OF ORBITER ATTITUDE REQUIREMENTS ON THE HST LAUNCH WINDOW

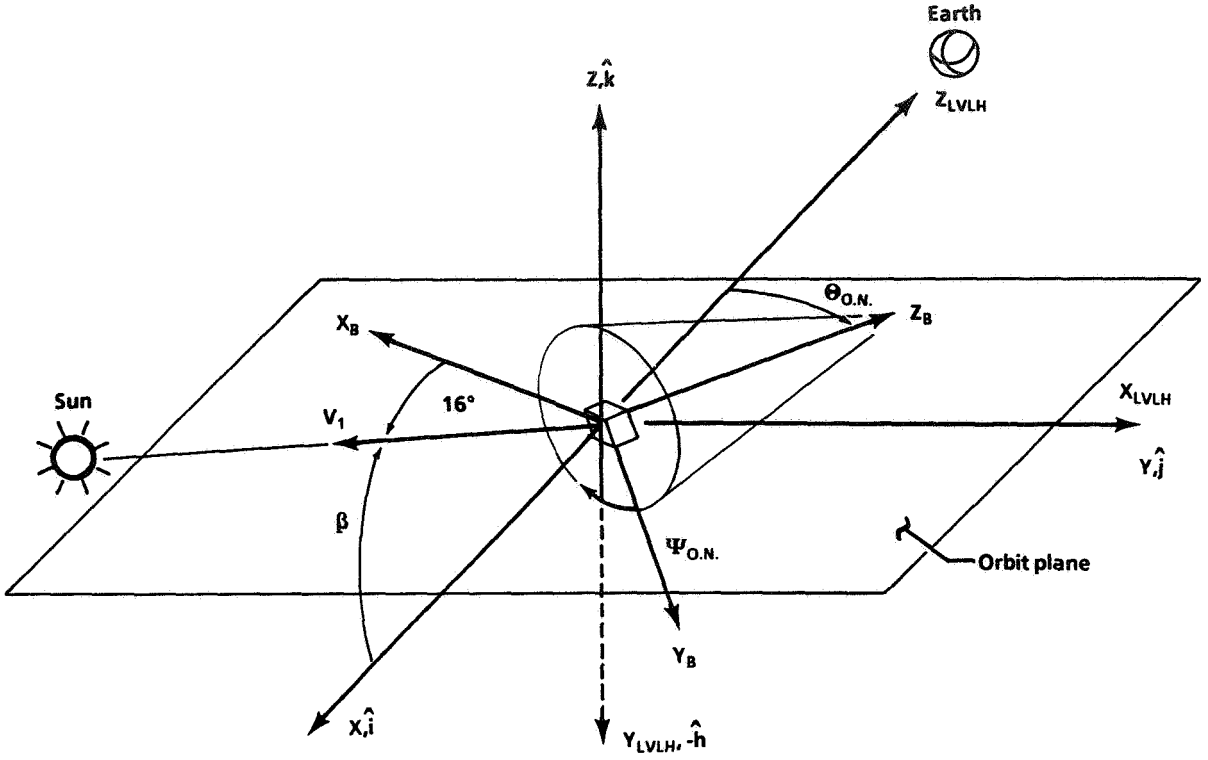
All launch times between the opening and closing of the HST launch window in Figure 4 ensured that communication, attitude, and lighting requirements could be met during the HST deployment on orbit 19. However, since launch time determined the Orbiter release attitude, certain launch times actually caused the Orbiter body to block the Ku-band antenna's line-of-sight communication with TDRSS. The launch times causing these unacceptable Orbiter release attitudes must be eliminated from the launch window.

To determine whether Orbiter attitudes required for payload operations affect a launch window, the analyst must first determine the attitudes that satisfy the payload requirements during the mission. Typically, more than one Orbiter attitude will meet the operational requirements of a free-flying or attached Shuttle payload. The analyst would limit this range of acceptable attitudes either by fixing the Orbiter orientation relative to the orbit plane or celestial sphere, or by requiring the Orbiter maintain an attitude that continually tracks a desired target.

Although the Orbiter attitude had to point the HST's -V1 axis at the Sun during the deployment, only one release attitude also aligned the Orbiter's Z-body axis with the orbital velocity vector at the separation burn. A formula was determined for this Sun-tracking, inertial attitude based on the known attitude and pointing constraints and the solar Beta angle at orbital noon. This geometrical derivation is illustrated in Figure 5. Like the HST release time, the Orbiter release attitude also varies with launch time and date because the solar Beta angle changes with launch time and date.

The Orbiter release attitudes could conceivably affect the lighting or communication required for HST deployment. However, the lighting requirement was met as long as HST was pointing at the Sun and its release was timed to coincide with orbital noon. But, determining whether the Orbiter release attitude interfered with TDRSS communication required special analysis, since the attitude varied with launch time and date. To assess the effect of attitude on communication, the launch window analyst had to construct a spherical look angle plot, or blockage pattern.

The blockage pattern is a graphical tool used in attitude and pointing analysis to determine the effect of Orbiter attitudes on the acquisition of targets relative to the Orbiter and payload. This plot represents a 360° spherical perspective from a pointing instrument, such as an antenna, sensor, or camera, that must make contact with a target. Plotting the time-varying and attitude-dependent positions of a target on the blockage pattern determines whether contact with the target can be made by the pointing instrument.



$\hat{h} \equiv$ orbital angular momentum vector

$-\hat{h} = -\hat{k}$

$Y_B \cdot -\hat{h} = \cos \Psi_{O.N}$

$$\Psi_{O.N.} = \cos^{-1}\left[\frac{\sin\theta_{O.N.}\cos\beta}{\sqrt{\sin^2\beta + \sin^2\theta_{O.N.}\cos^2\beta}}\right]$$

Therefore, the Orbiter LVLH attitude at orbital noon is:

$$P = \theta_{O.N.} = f_1(\beta)$$
$$Y = \Psi_{O.N.} = f_2(\beta)$$
$$R = 0^\circ$$

Figure 5.– The Orbiter attitude at orbital noon.

Knowing that Orbiter-to-TDRSS communication during the deployment would occur on Ku-band frequencies, a blockage pattern was created for the Orbiter's Ku-band antenna, as shown in Figure 6. The positions of the TDRSS satellites relative to the Orbiter on orbit 19 were plotted on the Ku-band antenna blockage pattern for the different launch times and the varying release attitudes that the Orbiter would be in over the duration of the launch window. The locations of the TDRSS satellites for a November 11, 1988, launch date can be seen in Figure 7. This plot shows that communication is affected by Orbiter and HST structures within the field-of-view of the Ku-band antenna, and by Earth occultation due to relative orbital motion between the Orbiter and a TDRSS satellite.

By plotting TDRSS satellite positions on the Orbiter Ku-band blockage pattern for launch window durations over different days of the year, the analyst determined the launch times that produced Orbiter release attitudes that would allow contact with TDRSS and simultaneously point HST at the Sun while the Orbiter was in the proper separation attitude. The effect of the Orbiter release attitude on the year-long HST launch window is shown in Figure 8.

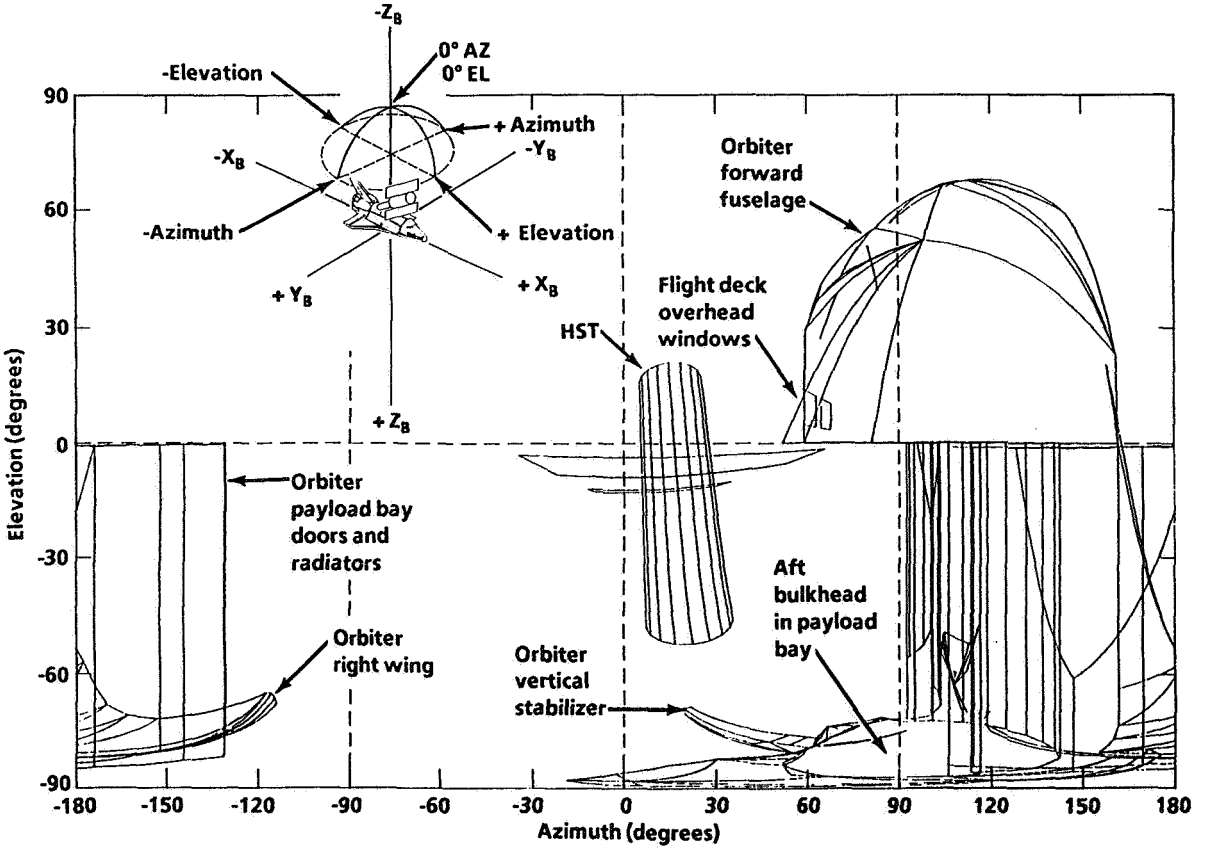
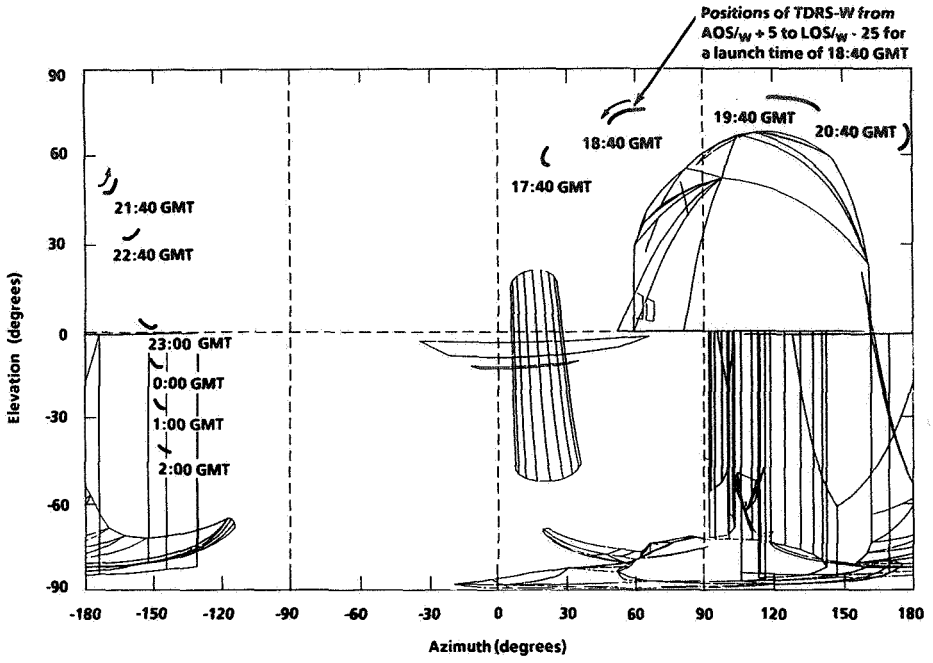
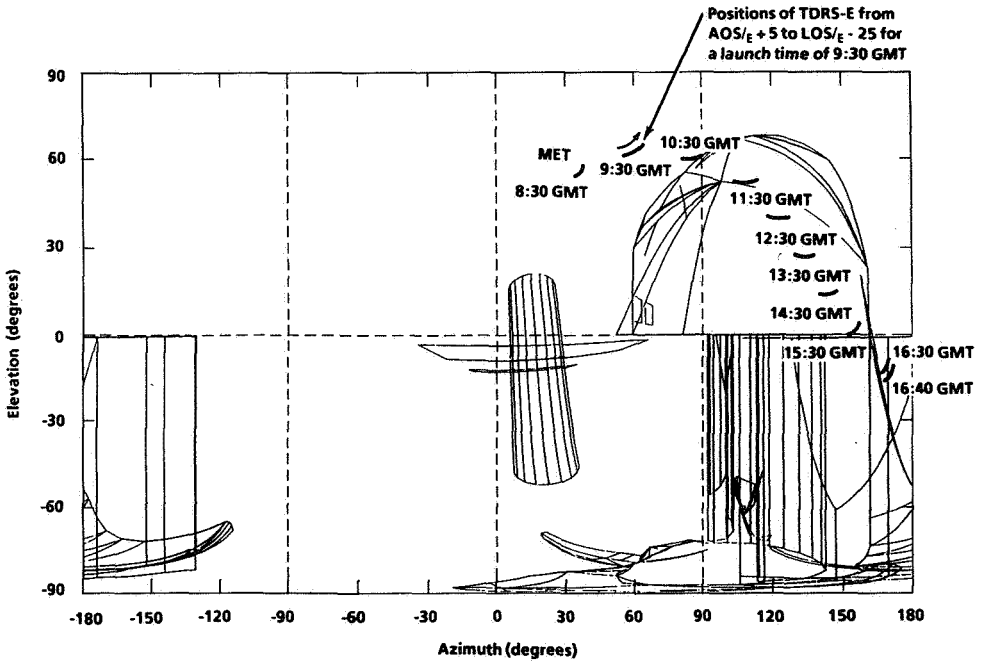


Figure 6.- Ku-band antenna blockage pattern.



(a) Positions of TDRS-W for the HST launch window using TDRS-W.



(b) Positions of TDRS-E for the HST launch window using TDRS-E.

Figure 7.– Positions of TDRSS from the Ku-band antenna for launch times on November 11, 1988.

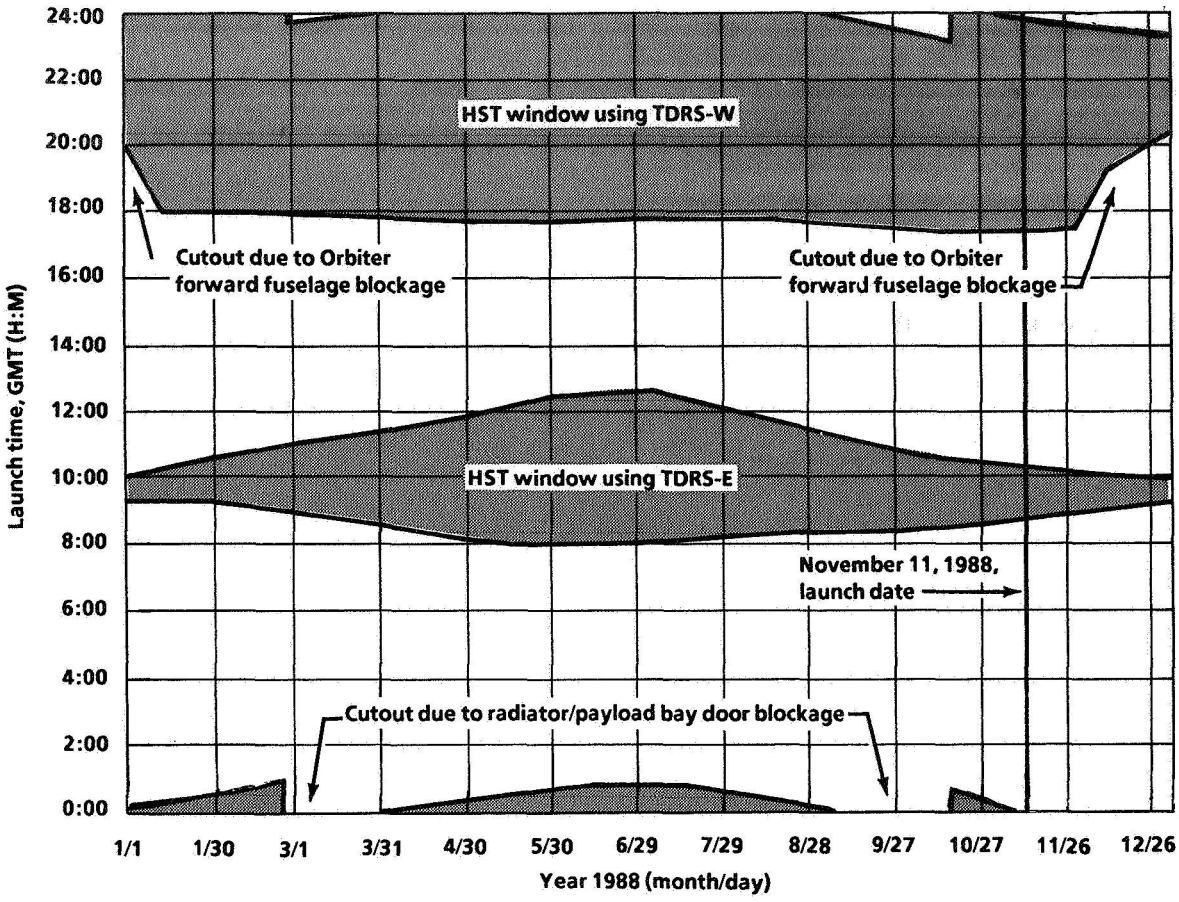


Figure 8.– HST launch window using Ku-band antenna for orbit 19 deployment.

SUMMARY

The launch window analysis described in this paper was used to determine the Shuttle launch times to deploy the HST. Shuttle attitudes that are required for payload operations are usually independent of launch time, but analysis for STS-31 showed that the Shuttle Orbiter attitude required to release the HST varied with launch time and date. The results of the analysis also showed that launch time determined the acceptable times to deploy the HST. The launch window that was eventually developed for STS-31 allowed the mission planner to determine the Orbiter release attitude and HST release time for a specific launch time, and to make prelaunch decisions if a launch delay occurred or if a TDRSS satellite was not available.

Analytical solutions were derived for the HST release time and the Orbiter release attitude based on the requirements for the deployment operations. A spherical look angle plot, or blockage pattern, was a visualization tool used to predict the effect of the Orbiter release attitude and the HST release time on the launch window. The methods developed from the HST launch window analysis can also be used in the launch window analysis for Shuttle missions with payloads with communication, pointing, and lighting requirements similar to HST, such as Spacelab missions or missions deploying low Earth-orbiting satellites.

For the HST deployment mission, a change in launch time rotated the Shuttle's trajectory in inertial space so the required orbital conditions for HST release occurred at a particular time. Changing the orientation of the orbital trajectory with launch time was the basis for the HST launch window analysis. However, the trajectory profile for the HST mission could have been altered by other means to obtain the orbital period, altitude, inclination, or Earth-latitude crossing that would ensure the required orbital conditions for the HST deployment. This could have been accomplished by a number of design options, such as ascent yaw-steering, nonstandard inclinations, alternative insertion altitudes, alternative days for payload operations, and orbit adjust maneuvers. Although some of these options were considered for the HST mission, discussion of these options is an entire subject by itself and was not covered in this paper.

ACKNOWLEDGMENTS

I would like to thank the Rockwell Space Operations Company and the NASA Johnson Space Center for giving me the opportunity to publish this paper. I would especially like to thank Dirk Deam for his supervision and technical support throughout the writing of this paper. I would also like to thank the editing, graphics, and word processing groups from the Omniplan Integrated Documentation Support Department. Their patience and diligence contributed immeasurably to the publication of this paper.

This paper is dedicated to the present and former members of the STS-31 Orbit Design Team: Scott Appelbaum, Vicky Whynott, Steve Stich, Jessica Landisman, Kevin Schaefer, and Alexie Benney. Their enduring efforts and enthusiasm over the last 4 years have made a significant contribution to the success of the mission.

REFERENCES

1. Hubble Space Telescope Payload Integration Plan, National Aeronautics and Space Administration, rev C, NSTS-14009, Aug. 1988.

2. Cargo Systems Manual: Hubble Space Telescope, All Flights, vol. I, rev A, JSC-2206, July 1988.

Contribution of Zonal Harmonics to Gravitational Moment

Carlos M. Roithmayr*
NASA Johnson Space Center, Houston, Texas

May 23, 1990

Abstract

A celestial body produces a gravitational moment about the mass center of a small orbiting body, which affects the orientation of the smaller body. Each zonal harmonic in the gravitational potential of a celestial body is shown to make a contribution to the gravitational moment which can be expressed in a recursive vector-dyadic form. A formal derivation is presented, followed by an example in which the result is employed in obtaining the contribution of the zonal harmonic of 2nd degree. The contribution of the zonal harmonic of 3rd degree is also reported.

Introduction

The gravitational moment about the mass center of a body in orbit about a celestial body has an important effect on the orientation of the orbiting body. The more misshapen the celestial body, and the less uniform its mass distribution, the more involved is the calculation of the gravitational moment (and

*Aerospace Engineer, Control and Guidance Systems Branch

force) it exerts. Situations in which it might be important to calculate accurately the gravitational moment include the design of spacecraft for expeditions to asteroids, comets, and the moons of Mars.

In Ref [1], a method for obtaining a vector-dyadic expression for the moment exerted about a small body's mass center by an oblate spheroid was set forth. The derivation of that expression made use of a gravitational potential written in terms of the zonal harmonic of 2nd degree. When gravitational potentials containing zonal harmonics of degree 2 or greater are considered, each zonal harmonic makes a contribution to the gravitational moment.

Recorded below is a vector-dyadic expression for the contribution of a zonal harmonic of degree n to the gravitational moment, produced by a body, about the mass center of a small body. As is the case with all vector-dyadic expressions, this result is basis independent— that is, the vectors and dyadics can be expressed in any convenient vector basis.

The equation given below is recursive: the contribution to the gravitational moment from the zonal harmonic of degree n is a function of the moment contributions from the zonal harmonics of degree $n-1$ and $n-2$. The equation contains Legendre polynomials and derivatives of Legendre polynomials of degree $n-1$ and $n-2$. The Legendre polynomials, as well as their derivatives, can, themselves, be generated by means of recursion formulae.

As an example, the contribution to the gravitational moment from the zonal harmonic of degree 2 is worked out. The contribution of the zonal harmonic of degree 3 is also given.

Assertion

Figure 1 shows a small body B in the presence of an axisymmetric body E. The distance between B^*, the mass center of B, and E^*, the mass center of

E, is assumed to exceed the greatest distance from B^* to any point of B. The system of gravitational forces exerted by E on B produces a moment M about B^*, and M can be written as

$$M = \frac{3\mu}{R^3}\hat{r} \times \underline{I} \bullet \hat{r} + \sum_{n=2}^{\infty} M_n \tag{1}$$

where M_n is the contribution of the zonal harmonic of degree n and can be obtained by using the recursion relation

$$\begin{aligned} M_n = \frac{\mu J_n {R_E}^n}{nR^{n+3}} \Bigg\{ & \Big[(2n-1)\left[(n+2)P_{n-1} + 3(\hat{r} \bullet \hat{n})P'_{n-1}\right] - (2n-2)P'_{n-2}\Big] \\ & (\hat{n} \times \underline{I} \bullet \hat{r} + \hat{r} \times \underline{I} \bullet \hat{n}) \\ + & \Big[(4n-4)\left[(n+1)P_{n-2} + (\hat{r} \bullet \hat{n})P'_{n-2}\right] \\ & -(2n-1)(\hat{r} \bullet \hat{n})\left[(4n+8)P_{n-1} + 4(\hat{r} \bullet \hat{n})P'_{n-1}\right]\Big]\hat{r} \times \underline{I} \bullet \hat{r} \\ - & \Big[2(2n-1)P'_{n-1}\Big]\hat{n} \times \underline{I} \bullet \hat{n}\Bigg\} \\ + & \frac{2n-1}{n}\frac{J_n}{J_{n-1}}\frac{R_E}{R}(\hat{r} \bullet \hat{n})M_{n-1} - \frac{n-1}{n}\frac{J_n}{J_{n-2}}\left(\frac{R_E}{R}\right)^2 M_{n-2} \end{aligned} \tag{2}$$

where; μ is the gravitational parameter of E, J_n is the zonal harmonic coefficient of degree n, R_E is the mean equatorial radius of E, R is the distance from E^* to B^*, P_n is the Legendre polynomial of degree n and argument S_λ, S_λ is the Sine of λ, the latitude of B^* ($S_\lambda = \hat{r} \bullet \hat{n}$), P'_n is the first derivative, with respect to its argument, of P_n, $\hat{r}$ is the unit position vector from E^* to B^*, $\hat{n}$ is the unit vector in the direction of the axis of symmetry of E, and $\underline{I}$ is the inertia dyadic of B relative to B^*.

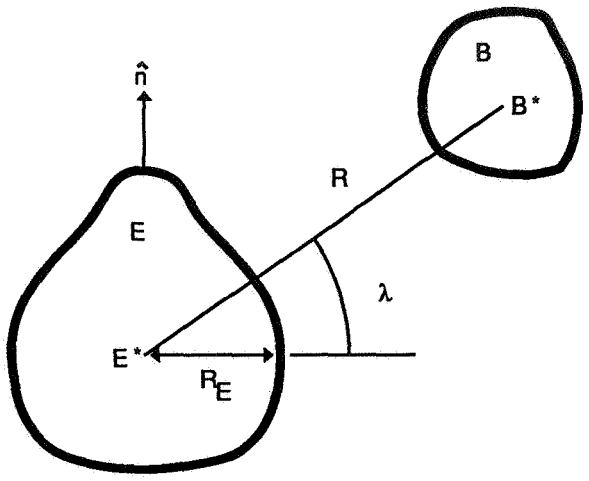


Figure 1: Body B in the Presence of Body E

Derivation

The gravitational forces exerted by a body E on a small body produce a moment $\boldsymbol{M}$ about the mass center B^* of B. $\boldsymbol{M}$ is given approximately by equation (2.18.1) of Ref [2],

$$\boldsymbol{M} = -\underline{I} \overset{\times}{\bullet} \nabla\nabla V(\boldsymbol{R}) \tag{3}$$

where $\boldsymbol{R}$ is the position vector from E^* to B^* and ∇ denotes differentiation with respect to the vector $\boldsymbol{R}$. Section 2.9 of Ref [2] contains a thorough explanation of how one differentiates with respect to a vector. The definition of the cross-dot product, $\overset{\times}{\bullet}$, appears on p. 156 of Ref [2]. The gravitational potential of E is symbolized by V.

Equation (2.13.14) in Ref [2] deals with the gravitational potential of an axisymmetric body and contains an infinite series of zonal harmonics. For a particle of unit mass coincident with B^*,

$$V = \frac{\mu}{R}\left[1 - \sum_{n=2}^{\infty}\left(\frac{R_E}{R}\right)^n J_n P_n(\mathrm{S}_\lambda)\right] \tag{4}$$

where S_λ, the argument of P_n, is equal to the sine of the geographic latitude of B^*. Eq. (4) can be simplified to

$$V = \frac{\mu}{R} + \sum_{n=2}^{\infty} V_n \tag{5}$$

when the contribution to the gravitational potential of the zonal harmonic of degree n, V_n, is defined as

$$V_n \triangleq -\frac{\mu {R_E}^n}{R^{n+1}} J_n P_n \tag{6}$$

Ref [1] shows that $-\underline{I} \overset{\times}{\bullet} \nabla\nabla(\mu/R) = (3\mu/R^3)\hat{r} \times \underline{I} \bullet \hat{r}$. Hence, eq. (3) can be rewritten as eq. (1),

$$M = \frac{3\mu}{R^3}\hat{r} \times \underline{I} \bullet \hat{r} + \sum_{n=2}^{\infty} M_n \tag{7}$$

so long as the contribution to the gravitational moment of the zonal harmonic of degree n, M_n, is defined as

$$M_n \triangleq -\underline{I} \overset{\times}{\bullet} \nabla\nabla V_n \tag{8}$$

The Legendre polynomial of degree n, $P_n(x)$, is expressed recursively in equation (8.71) of Ref [3] in terms of Legendre polynomials of degree $n-1$, $n-2$, and their argument, x, for $n \geq 2$, as follows:

$$P_n(x) = \frac{1}{n}[(2n-1)xP_{n-1}(x) - (n-1)P_{n-2}(x)] \tag{9}$$

Eq. (9) can also be produced with $m = 0$ in formula I of Table 1 in Ref [4]. Substituting from this recursion relation for the Legendre polynomials into eq. (6) leads to a recursive expression for V_n:

$$V_n = \frac{2n-1}{n}\frac{J_n}{J_{n-1}}\frac{R_E}{R}\mathrm{S}_\lambda V_{n-1} - \frac{n-1}{n}\frac{J_n}{J_{n-2}}\left(\frac{R_E}{R}\right)^2 V_{n-2} \tag{10}$$

Eq. (8) requires that a dyadic be formed by differentiating V_n twice with respect to $\boldsymbol{R}$. The first derivative of V_n with respect to $\boldsymbol{R}$ yields a recursion relation for the vector ∇V_n, the contribution of the zonal harmonic of degree n to the gravitational force exerted by E on a particle of unit mass coincident with B^*.

$$\begin{aligned}\nabla V_n = \frac{2n-1}{n}\frac{J_n}{J_{n-1}}R_E\left[\frac{V_{n-1}}{R}\nabla \mathrm{S}_\lambda + \frac{\mathrm{S}_\lambda}{R}\nabla V_{n-1} - \frac{\mathrm{S}_\lambda V_{n-1}}{R^3}\boldsymbol{R}\right] \\ - \frac{n-1}{n}\frac{J_n}{J_{n-2}}{R_E}^2\left[\frac{\nabla V_{n-2}}{R^2} - \frac{2V_{n-2}}{R^4}\boldsymbol{R}\right]\end{aligned} \tag{11}$$

The second derivative of V_n with respect to $\boldsymbol{R}$ yields a recursion relation for the symmetric dyadic $\nabla\nabla V_n$,

$$\begin{aligned}\nabla\nabla V_n &= \frac{2n-1}{n}\frac{J_n}{J_{n-1}}R_E\left\{\frac{V_{n-1}}{R}\nabla\nabla \mathrm{S}_\lambda + \frac{\nabla \mathrm{S}_\lambda \nabla V_{n-1}}{R} - \frac{V_{n-1}\nabla \mathrm{S}_\lambda \boldsymbol{R}}{R^3}\right. \\ &+ \frac{\nabla V_{n-1}\nabla \mathrm{S}_\lambda}{R} + \frac{\mathrm{S}_\lambda}{R}\nabla\nabla V_{n-1} - \frac{\mathrm{S}_\lambda \nabla V_{n-1}\boldsymbol{R}}{R^3} \\ &- \left.\frac{V_{n-1}\boldsymbol{R}\nabla \mathrm{S}_\lambda}{R^3} - \frac{\mathrm{S}_\lambda \boldsymbol{R}\nabla V_{n-1}}{R^3} + \frac{3\mathrm{S}_\lambda V_{n-1}\boldsymbol{R}\boldsymbol{R}}{R^5} - \frac{\mathrm{S}_\lambda V_{n-1}}{R^3}\underline{U}\right\} \\ &- \frac{n-1}{n}\frac{J_n}{J_{n-2}}{R_E}^2\left\{\frac{\nabla\nabla V_{n-2}}{R^2} - \frac{2\nabla V_{n-2}\boldsymbol{R}}{R^4}\right. \\ &- \left.\frac{2\boldsymbol{R}\nabla V_{n-2}}{R^4} + \frac{8V_{n-2}\boldsymbol{R}\boldsymbol{R}}{R^6} - \frac{2V_{n-2}}{R^4}\underline{U}\right\}\end{aligned} \tag{12}$$

where $\underline{U}$ is the unit dyadic.

Cross-Dot Products

We now perform the cross-dot product with dyadics $\underline{I}$ and $\nabla\nabla V_n$, making use of the right hand side of eq. (12) and the cross-dot identity $\underline{I} \overset{\times}{\bullet} \underline{U} = \mathbf{0}$, which is set forth in eq. (19) of Ref [1], obtaining

$$\begin{aligned}
\underline{I} \overset{\times}{\bullet} \nabla\nabla V_n \;=\; & \frac{2n-1}{n}\frac{J_n}{J_{n-1}}R_E\Big\{\frac{V_{n-1}}{R}\underline{I} \overset{\times}{\bullet} \nabla\nabla \mathrm{S}_\lambda - \frac{V_{n-1}}{R^3}\underline{I} \overset{\times}{\bullet} (\nabla \mathrm{S}_\lambda \boldsymbol{R} + \boldsymbol{R}\nabla \mathrm{S}_\lambda) \\
& + \frac{\mathrm{S}_\lambda}{R}\underline{I} \overset{\times}{\bullet} \nabla\nabla V_{n-1} + \frac{1}{R}\underline{I} \overset{\times}{\bullet} (\nabla \mathrm{S}_\lambda \nabla V_{n-1} + \nabla V_{n-1}\nabla \mathrm{S}_\lambda) \\
& - \frac{\mathrm{S}_\lambda}{R^3}\underline{I} \overset{\times}{\bullet} (\nabla V_{n-1}\boldsymbol{R} + \boldsymbol{R}\nabla V_{n-1}) + \frac{3\mathrm{S}_\lambda V_{n-1}}{R^5}\underline{I} \overset{\times}{\bullet} \boldsymbol{RR} - \mathbf{0}\Big\} \\
& - \frac{n-1}{n}\frac{J_n}{J_{n-2}}{R_E}^2\Big\{\frac{1}{R^2}\underline{I} \overset{\times}{\bullet} \nabla\nabla V_{n-2} + \frac{8V_{n-2}}{R^6}\underline{I} \overset{\times}{\bullet} \boldsymbol{RR} - \mathbf{0} \\
& - \frac{2}{R^4}\underline{I} \overset{\times}{\bullet} (\nabla V_{n-2}\boldsymbol{R} + \boldsymbol{R}\nabla V_{n-2})\Big\} \qquad (13)
\end{aligned}$$

In order to carry out the cross-dot products with $\underline{I}$ and the other dyadics on the right side of eq. (13), we will express these dyadics in terms of $\boldsymbol{R}$ and $\hat{\boldsymbol{n}}$, a unit vector parallel to the axis of symmetry of E.

The sine of λ can be expressed as $\sin\lambda = (\boldsymbol{R}\bullet\hat{\boldsymbol{n}})/R = \hat{\boldsymbol{r}}\bullet\hat{\boldsymbol{n}}$ so that the first derivative of S_λ with respect to $\boldsymbol{R}$ is

$$\nabla \mathrm{S}_\lambda = \nabla\frac{\boldsymbol{R}\bullet\hat{\boldsymbol{n}}}{R} = \frac{\hat{\boldsymbol{n}}}{R} - \frac{(\boldsymbol{R}\bullet\hat{\boldsymbol{n}})\boldsymbol{R}}{R^3} \qquad (14)$$

and the second derivative of S_λ with respect to $\boldsymbol{R}$ is

$$\nabla\nabla \mathrm{S}_\lambda = \frac{3(\boldsymbol{R}\bullet\hat{\boldsymbol{n}})\boldsymbol{RR}}{R^5} - \frac{1}{R^3}[\hat{\boldsymbol{n}}\boldsymbol{R} + \boldsymbol{R}\hat{\boldsymbol{n}} + (\boldsymbol{R}\bullet\hat{\boldsymbol{n}})\underline{U}] \qquad (15)$$

Eq. (20) of Ref [1] is a derivation of a cross-dot identity which will be used repeatedly throughout the sequel: For any dyad $\boldsymbol{uv}$ composed of vectors $\boldsymbol{u}$ and $\boldsymbol{v}$, it is shown that $\underline{I} \overset{\times}{\bullet} \boldsymbol{uv} = -\boldsymbol{u}\times\underline{I}\bullet\boldsymbol{v}$. By making use of this identity, and eq. (15) above, one can write

$$\underline{I} \overset{\times}{\bullet} \nabla\nabla \mathrm{S}_\lambda = \frac{1}{R^2}(\hat{\boldsymbol{n}}\times\underline{I}\bullet\hat{\boldsymbol{r}} + \hat{\boldsymbol{r}}\times\underline{I}\bullet\hat{\boldsymbol{n}}) + \mathbf{0} - \frac{3(\boldsymbol{R}\bullet\hat{\boldsymbol{n}})}{R^3}\hat{\boldsymbol{r}}\times\underline{I}\bullet\hat{\boldsymbol{r}} \qquad (16)$$

where $\hat{r}$ is a unit vector in the direction of $\boldsymbol{R}$. By recalling the definition in eq. (6), one can evaluate the first cross-dot product on the right side of eq. (13).

$$\frac{V_{n-1}}{R}\underline{I} \overset{\times}{\bullet} \nabla\nabla \mathrm{S}_\lambda = -\mu J_{n-1} {R_E}^{n-1} \frac{P_{n-1}}{R^{n+3}} [\hat{n} \times \underline{I} \bullet \hat{r} + \hat{r} \times \underline{I} \bullet \hat{n} - 3(\hat{r} \bullet \hat{n})\hat{r} \times \underline{I} \bullet \hat{r}] \quad (17)$$

The dyadic required for the second cross-dot product on the right side of eq. (13) is easily constructed by using eq. (14):

$$\nabla \mathrm{S}_\lambda \boldsymbol{R} + \boldsymbol{R} \nabla \mathrm{S}_\lambda = \hat{n}\hat{r} + \hat{r}\hat{n} - 2(\hat{r} \bullet \hat{n})\hat{r}\hat{r} \quad (18)$$

Thus,

$$\frac{V_{n-1}}{R^3}\underline{I} \overset{\times}{\bullet} (\nabla \mathrm{S}_\lambda \boldsymbol{R} + \boldsymbol{R} \nabla \mathrm{S}_\lambda) = -\mu J_{n-1} {R_E}^{n-1} \frac{P_{n-1}}{R^{n+3}} [2(\hat{r} \bullet \hat{n})\hat{r} \times \underline{I} \bullet \hat{r} - \hat{n} \times \underline{I} \bullet \hat{r} - \hat{r} \times \underline{I} \bullet \hat{n}] \quad (19)$$

Replacing n with $n-1$ in eq. (8) allows the third cross-dot product on the right side of eq. (13) to be immediately expressed in terms of $\boldsymbol{M}_{n-1}$.

$$\frac{\mathrm{S}_\lambda}{R}\underline{I} \overset{\times}{\bullet} \nabla\nabla V_{n-1} = -\frac{\hat{r} \bullet \hat{n}}{R} \boldsymbol{M}_{n-1} \quad (20)$$

The fourth cross-dot product on the right side of eq. (13) contains the quantity ∇V_{n-1}. Temporarily allow x to be the argument of P_{n-1} [See eq. (6)] and write the derivative of V_{n-1} with respect to $\boldsymbol{R}$ as

$$\nabla V_{n-1}(x) = -\mu J_{n-1} {R_E}^{n-1} \left[\frac{1}{R^n} \frac{\mathrm{d}}{\mathrm{dx}} P_{n-1}(x) \nabla x - \frac{n P_{n-1}(x)}{R^{n+2}} \boldsymbol{R} \right] \quad (21)$$

The first derivative of a Legendre polynomial P_n with respect to its argument is often denoted by P_n'. For $n \geq 2$, a useful recurrence formula for P_n' can be found in problem 8-9 of Ref [3], p. 393, or in formula I of Table 1 in Ref [4] (with $m = 1$):

$$P_n'(x) = \frac{1}{n-1} [(2n-1)x P_{n-1}'(x) - n P_{n-2}'(x)] \quad (22)$$

Since the argument of P_{n-1} is known to be S_λ, we can make use of eq. (14) to rewrite eq. (21) as

$$\nabla V_{n-1} = -\mu J_{n-1} {R_E}^{n-1} \left\{ \frac{P'_{n-1}}{R^n} \left[\frac{\hat{n}}{R} - \frac{(\boldsymbol{R} \bullet \hat{n})\boldsymbol{R}}{R^3} \right] - \frac{nP_{n-1}}{R^{n+2}} \boldsymbol{R} \right\} \tag{23}$$

The sum of two dyadics, formed by juxtaposing the vectors ∇S_λ and ∇V_{n-1} in opposite order, yields the symmetric dyadic

$$\nabla S_\lambda \nabla V_{n-1} + \nabla V_{n-1} \nabla S_\lambda = -\mu J_{n-1} {R_E}^{n-1} \left\{ \frac{2P'_{n-1}}{R^n} \left[\frac{\hat{n}\hat{n}}{R^2} + \frac{(\boldsymbol{R} \bullet \hat{n})^2}{R^6} \boldsymbol{R}\boldsymbol{R} - \frac{(\boldsymbol{R} \bullet \hat{n})}{R^4} (\boldsymbol{R}\hat{n} + \hat{n}\boldsymbol{R}) \right] - \frac{nP_{n-1}}{R^{n+2}} \left[\frac{1}{R} (\boldsymbol{R}\hat{n} + \hat{n}\boldsymbol{R}) - \frac{2(\boldsymbol{R} \bullet \hat{n})}{R^3} \boldsymbol{R}\boldsymbol{R} \right] \right\} \tag{24}$$

Consequently, the fourth cross-dot product on the right side of eq. (13) can be expressed as

$$\begin{aligned} \frac{1}{R} \underline{I} \overset{\times}{\bullet} (\nabla S_\lambda \nabla V_{n-1} + \nabla V_{n-1} \nabla S_\lambda) &= -\mu J_{n-1} {R_E}^{n-1} \\ &\left\{ \frac{2P'_{n-1}}{R^{n+3}} \left[(\hat{r} \bullet \hat{n})(\hat{r} \times \underline{I} \bullet \hat{n} + \hat{n} \times \underline{I} \bullet \hat{r}) - \hat{n} \times \underline{I} \bullet \hat{n} - (\hat{r} \bullet \hat{n})^2 \hat{r} \times \underline{I} \bullet \hat{r} \right] \right. \\ &- \left. \frac{nP_{n-1}}{R^{n+3}} \left[2(\hat{r} \bullet \hat{n}) \hat{r} \times \underline{I} \bullet \hat{r} - \hat{n} \times \underline{I} \bullet \hat{r} - \hat{r} \times \underline{I} \bullet \hat{n} \right] \right\} \end{aligned} \tag{25}$$

The dyadic required for the fifth cross-dot product on the right side of eq. (13) can be constructed rather easily by employing eq. (23), which yields

$$\begin{aligned} &\nabla V_{n-1} \boldsymbol{R} + \boldsymbol{R} \nabla V_{n-1} = \\ &\quad -\mu J_{n-1} {R_E}^{n-1} \left\{ \frac{P'_{n-1}}{R^n} \left[\hat{n}\hat{r} + \hat{r}\hat{n} - 2(\hat{r} \bullet \hat{n}) \hat{r}\hat{r} \right] - \frac{2nP_{n-1}}{R^n} \hat{r}\hat{r} \right\} \end{aligned} \tag{26}$$

so that

$$\begin{aligned} &\frac{S_\lambda}{R^3} \underline{I} \overset{\times}{\bullet} (\nabla V_{n-1} \boldsymbol{R} + \boldsymbol{R} \nabla V_{n-1}) = \\ &\quad -\mu J_{n-1} {R_E}^{n-1} \left\{ \frac{(\hat{r} \bullet \hat{n}) P'_{n-1}}{R^{n+3}} \left[2(\hat{r} \bullet \hat{n}) \hat{r} \times \underline{I} \bullet \hat{r} - \hat{n} \times \underline{I} \bullet \hat{r} - \hat{r} \times \underline{I} \bullet \hat{n} \right] \right. \\ &+ \left. \frac{2n(\hat{r} \bullet \hat{n}) P_{n-1}}{R^{n+3}} \hat{r} \times \underline{I} \bullet \hat{r} \right\} \end{aligned} \tag{27}$$

The sixth cross-dot multiplication which must be performed in order to obtain $\underline{I} \overset{\times}{\bullet} \nabla\nabla V_n$ is one of the easiest to carry out. That is,

$$\frac{3S_\lambda V_{n-1}}{R^5}\underline{I} \overset{\times}{\bullet} \boldsymbol{RR} = 3\mu J_{n-1} {R_E}^{n-1}(\hat{\boldsymbol{r}} \bullet \hat{\boldsymbol{n}})\frac{P_{n-1}}{R^{n+3}}\hat{\boldsymbol{r}} \times \underline{I} \bullet \hat{\boldsymbol{r}} \tag{28}$$

The seventh cross-dot product on the right side of eq. (13) can be expressed in terms of M_{n-2}. Replacing n with $n-2$ in eq. (8), we get

$$\frac{1}{R^2}\underline{I} \overset{\times}{\bullet} \nabla\nabla V_{n-2} = -\frac{1}{R^2}\boldsymbol{M}_{n-2} \tag{29}$$

The eighth cross-dot product to be evaluated is simply

$$\frac{8V_{n-2}}{R^6}\underline{I} \overset{\times}{\bullet} \boldsymbol{RR} = 8\mu J_{n-2} {R_E}^{n-2}\frac{P_{n-2}}{R^{n+3}}\hat{\boldsymbol{r}} \times \underline{I} \bullet \hat{\boldsymbol{r}} \tag{30}$$

The dyadic required for the final cross-dot product is similar to that needed for the fifth cross-dot product,

$$\nabla V_{n-2}\boldsymbol{R} + \boldsymbol{R}\nabla V_{n-2} =$$
$$-\mu J_{n-2} {R_E}^{n-2}\left\{\frac{P'_{n-2}}{R^{n-1}}\left[\hat{\boldsymbol{n}}\hat{\boldsymbol{r}} + \hat{\boldsymbol{r}}\hat{\boldsymbol{n}} - 2(\hat{\boldsymbol{r}} \bullet \hat{\boldsymbol{n}})\hat{\boldsymbol{r}}\hat{\boldsymbol{r}}\right] - \frac{2(n-1)P_{n-2}}{R^{n-1}}\hat{\boldsymbol{r}}\hat{\boldsymbol{r}}\right\} \tag{31}$$

so that

$$\frac{2}{R^4}\underline{I} \overset{\times}{\bullet} (\nabla V_{n-2}\boldsymbol{R} + \boldsymbol{R}\nabla V_{n-2}) =$$
$$-\mu J_{n-2} {R_E}^{n-2}\left\{\frac{2P'_{n-2}}{R^{n+3}}\left[2(\hat{\boldsymbol{r}} \bullet \hat{\boldsymbol{n}})\hat{\boldsymbol{r}} \times \underline{I} \bullet \hat{\boldsymbol{r}} - \hat{\boldsymbol{n}} \times \underline{I} \bullet \hat{\boldsymbol{r}} - \hat{\boldsymbol{r}} \times \underline{I} \bullet \hat{\boldsymbol{n}}\right]\right.$$
$$\left. + \quad \frac{4(n-1)P_{n-2}}{R^{n+3}}\hat{\boldsymbol{r}} \times \underline{I} \bullet \hat{\boldsymbol{r}}\right\} \tag{32}$$

Substituting from eqs. (17), (19), (20), (25), (27) - (30), and (32) into (13) and then into (8) leads to eq. (2), which is a recursive vector-dyadic expression for the contribution of the nth zonal harmonic to the gravitational moment.

Examples

In order to demonstrate the use of eq. (2), we will use it to obtain M_2, the contribution to the gravitational moment from the zonal harmonic of degree 2.

The two required Legendre polynomials are $P_1(S_\lambda) = S_\lambda = \hat{r} \bullet \hat{n}$ and $P_0(S_\lambda) = 1$. Legendre polynomials of degree greater than or equal to 2 can be obtained recursively by using eq. (9). The two derivatives of Legendre polynomials which will be needed are $P_1'(S_\lambda) = 1$ and $P_0'(S_\lambda) = 0$. Derivatives with respect to the argument of Legendre polynomials can be generated with the recursion formula (22) for $n \geq 2$.

Eq. (2) also requires knowledge of $\boldsymbol{M}_0$ and $\boldsymbol{M}_1$ in order to produce $\boldsymbol{M}_2$. Eq. (8) is helpful in developing expressions for $\boldsymbol{M}_0$ and $\boldsymbol{M}_1$.

The Legendre polynomial of degree zero is equal to 1, regardless of its argument, and the scalar V_0 [See eq. (6)] is

$$V_0 = -\mu J_0 P_0 / R = -\mu J_0 / R \tag{33}$$

The dyadic formed with V_0 is then

$$\nabla\nabla V_0 = \nabla\nabla(-\mu J_0/R) = \nabla(\mu J_0 \boldsymbol{R}/R^3) = \frac{\mu J_0}{R^3}(\underline{U} - 3\hat{r}\hat{r}) \tag{34}$$

Eq. (8) tells us that

$$\boldsymbol{M}_0 = -\frac{3\mu J_0}{R^3}\hat{r} \times \underline{I} \bullet \hat{r} \tag{35}$$

J_0 is an undefined constant, but the coefficient of $\boldsymbol{M}_0$ in eq. (2) contains J_0 in the denominator. Hence, a numerical value of J_0 is not required for constructing $\boldsymbol{M}_2$.

A similar process leads to $\boldsymbol{M}_1$. The value of the Legendre polynomial of degree 1 is identical to the argument, so the scalar V_1 is

$$V_1 = -\mu J_1 R_E \frac{P_1}{R^2} = -\mu J_1 R_E \frac{(\boldsymbol{R} \bullet \hat{n})}{R^3} \tag{36}$$

The dyadic formed with V_1 is

$$\nabla\nabla V_1 = -\mu J_1 R_E \nabla\nabla \frac{(\boldsymbol{R} \bullet \hat{n})}{R^3} = -\mu J_1 R_E \nabla \left[\frac{\hat{n}}{R^3} - \frac{3(\boldsymbol{R} \bullet \hat{n})}{R^5}\boldsymbol{R}\right] =$$
$$-\mu J_1 R_E \left\{\frac{15(\boldsymbol{R} \bullet \hat{n})}{R^7}\boldsymbol{R}\boldsymbol{R} - \frac{3}{R^5}[\boldsymbol{R}\hat{n} + \hat{n}\boldsymbol{R} + (\boldsymbol{R} \bullet \hat{n})\underline{U}]\right\} \tag{37}$$

so that

$$M_1 = \frac{3\mu J_1 R_E}{R^4}\left[\hat{n}\times\underline{I}\bullet\hat{r}+\hat{r}\times\underline{I}\bullet\hat{n}-5(\hat{r}\bullet\hat{n})\hat{r}\times\underline{I}\bullet\hat{r}\right] \qquad (38)$$

Like J_0, the constant J_1 is undefined, and unneeded for the purpose of obtaining M_2. Note that M_0 and M_1 do not represent contributions to the gravitational moment, but are required to begin the process of recursion which will generate moment contributions beginning with M_2.

By substituting from eqs (35) and (38) into (2), we arrive at the following result with $n = 2$:

$$\begin{aligned} M_2 &= \frac{\mu J_2 {R_E}^2}{2R^5}\Big\{\left[30(\hat{r}\bullet\hat{n})\right](\hat{n}\times\underline{I}\bullet\hat{r}+\hat{r}\times\underline{I}\bullet\hat{n}) \\ &+ \left[15-105(\hat{r}\bullet\hat{n})^2\right]\hat{r}\times\underline{I}\bullet\hat{r}-6\hat{n}\times\underline{I}\bullet\hat{n}\Big\} \end{aligned} \qquad (39)$$

If eq. (1) of Ref [1] is expressed as $M = (3\mu/R^3)\hat{r}\times\underline{I}\bullet\hat{r}+M_2$, it can be seen that M_2 from Ref [1] is identical to eq. (39), above.

The contribution M_3 can be obtained in a similar manner, using the values of $P_2(\mathrm{S}_\lambda)$, $P_1(\mathrm{S}_\lambda)$, $P_2'(\mathrm{S}_\lambda)$, $P_1'(\mathrm{S}_\lambda)$, M_2, and M_1.

$$\begin{aligned} M_3 &= \frac{\mu J_3 {R_E}^3}{6R^6}\Big\{\left[315(\hat{r}\bullet\hat{n})^2-45\right](\hat{n}\times\underline{I}\bullet\hat{r}+\hat{r}\times\underline{I}\bullet\hat{n}) \\ &+ \left[315(\hat{r}\bullet\hat{n})-945(\hat{r}\bullet\hat{n})^3\right]\hat{r}\times\underline{I}\bullet\hat{r}-90(\hat{r}\bullet\hat{n})\hat{n}\times\underline{I}\bullet\hat{n}\Big\} \end{aligned} \qquad (40)$$

Conclusions

A recursive vector-dyadic expression for the contribution of a zonal harmonic of degree n to the gravitational moment about the mass center of a small body can be obtained by a procedure which involves differentiating a celestial body's gravitational potential twice with respect to a vector. The recursive property of the result is a consequence of taking advantage of a recursion relation for Legendre polynomials that appear in the gravitational potential. When a celestial body's gravitational potential includes zonal harmonics, the vector-dyadic

expression above is useful for calculating their contributions to the gravitational moment. The contribution of the zonal harmonic of degree 2 is consistent with the gravitational moment exerted by an oblate spheroid.

References

[1] Roithmayr, C.M., "Gravitational Moment Exerted on a Small Body by an Oblate Body", *Journal of Guidance, Control, and Dynamics*, Vol 12, May-June 1989, pp 441-444

[2] Kane, T.R., Likins, P.W., and Levinson, D.A., *Spacecraft Dynamics*, McGraw-Hill, New York, 1983, Chapter 2

[3] Battin, R.H., *An Introduction to The Mathematics and Methods of Astrodynamics*, AIAA, New York, 1987, p. 393

[4] Lundberg, J.B., and Schutz, B.E., "Recursion Formulas of Legendre Functions for Use with Nonsingular Geopotential Models", *Journal of Guidance, Control, and Dynamics*, Vol 11, Jan-Feb 1988, pp. 32 - 38

A Technique for Optimal Temperature Estimation for Modeling Sunrise/Sunset Thermal Snap Disturbance Torque

D. F. Zimbelman*
C. J. Dennehy†
R. V. Welch‡
Fairchild Space Company
Fairchild Space Systems Division
Germantown, Maryland
U.S.A. 20874-1181

G. H. Born§
Colorado Center for Astrodynamics Research
University of Colorado
Boulder, Colorado
U.S.A. 80309-0431

ABSTRACT

This paper describes a predictive temperature estimation technique which can be used to drive a model of the Sunrise/Sunset thermal "snap" disturbance torque experienced by low Earth orbiting spacecraft. The twice per orbit impulsive disturbance torque is attributed to vehicle passage in and out of the Earth's shadow cone (umbra), during which, large flexible appendages undergo rapidly changing thermal conditions. Flexible members, in particular solar arrays, experience rapid cooling during umbra entrance (Sunset) and rapid heating during exit (Sunrise). The thermal "snap" phenomena has been observed during normal on-orbit operations of both the LANDSAT-4 satellite and the Communications Technology Satellite (CTS). Thermal "snap" has also been predicted to be a dominant source of error for the TOPEX satellite.

The fundamental equations used to model the Sunrise/Sunset thermal "snap" disturbance torque for a typical solar array like structure will be described. For this derivation the array is assumed to be a thin, cantilevered beam. The time varying thermal gradient is shown to be the driving force behind predicting the thermal "snap" disturbance torque and therefore motivates the need for accurate estimates of temperature. This paper will highlight the development of a technique to optimally estimate appendage surface temperatures. The objective analysis method used is structured on the Gauss-Markov Theorem and provides an optimal temperature estimate at a prescribed location given data from a distributed thermal sensor network. The optimally estimated surface temperatures could then be used to compute the thermal gradient across the body. The estimation technique is demonstrated using a typical satellite solar array.

---

*Engineer, GN&C Department. Graduate Student, Colorado Center for Astrodynamics Research/University of Colorado.

†Section Manager, GN&C Department.

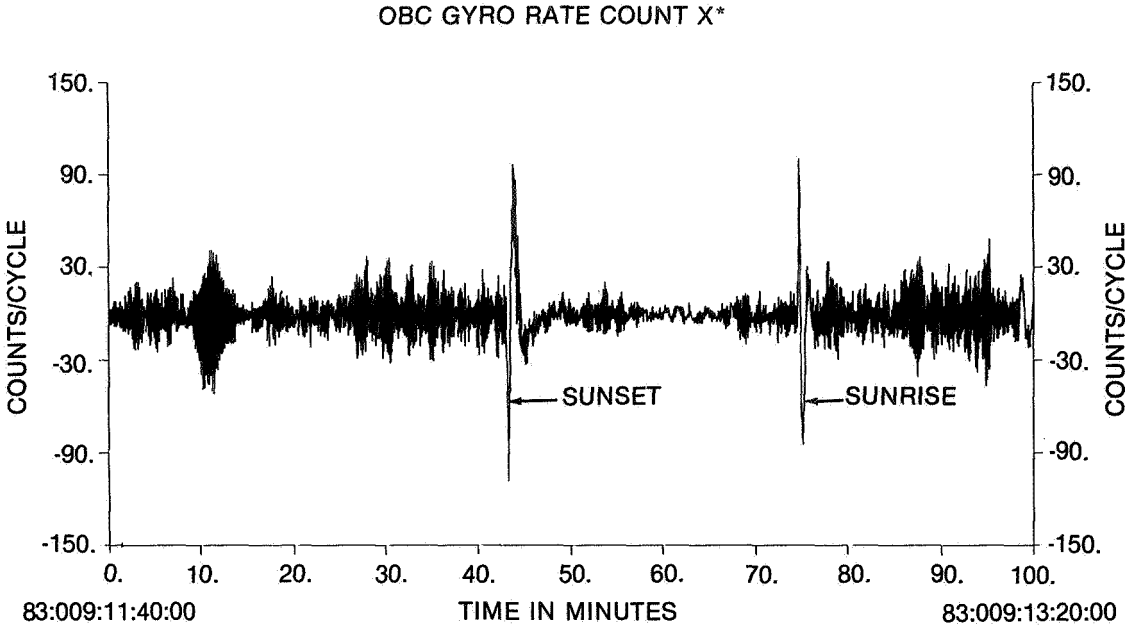
‡Senior Director, GN&C Department.

§Director, Colorado Center for Astrodynamics Research/University of Colorado. Professor, Aerospace Engineering Sciences.

INTRODUCTION

In the early 1980's, an unexpected perturbation was experienced by the LANDSAT-4 and 5 satellites. An anomously large, twice per orbit disturbance was observed in the flight data during normal on-orbit operations. An example of this disturbance is illustrated in Figures 1 and 2 which show the LANDSAT-4 roll and yaw axes derived rate telemetry data, respectively, over one orbital period. An inspection of Figures 1 and 2 reveals that the roll and yaw controlled response of the vehicle exhibits an impulsive velocity in one direction immediately followed by a similar motion in the opposite direction, and finally proceeded by a decaying step of the initial sign. There is minimal coupling present in the pitch axis data which indicates the disturbance is primarily distributed between the roll and yaw axes. The perturbation shown in Figures 1 and 2 has been correlated with the spacecraft's entrance (sunset) and exit (sunrise) to the eclipsed region of the orbit plane. During penumbral transitions, the vehicle undergoes rapidly changing thermal conditions which result in a thermally induced bending motion of the large single solar array. If the snaping or bending motion occurs at a non-constant rate a disturbance torque is generated about the hinged axis of the array, which is then transferred back onto the vehicle core-body. This concept is illustrated in Figure 3. This thermally induced disturbance, referred to as Thermal "snap" or Thermal Elastic Shock (TES), has also been observed during the three-axis stabilized operation of the Communications Technology Satellite (CTS), but to a much lesser degree. The difference in perturbation magnitude was a result of the differences between the two spacecraft designs. The TES disturbance is most pronounced for asymmetric satellite configurations, such as the single-wing LANDSAT vehicles. Satellites possesing a dual-wing array design, such as CTS, are significantly less affected by TES since the motion of both arrays tends to be self-compensating. However, a single array configuration is typically required to provide an unobstructed radiator view of cold space for proper thermal control of mission sensor payload instruments.

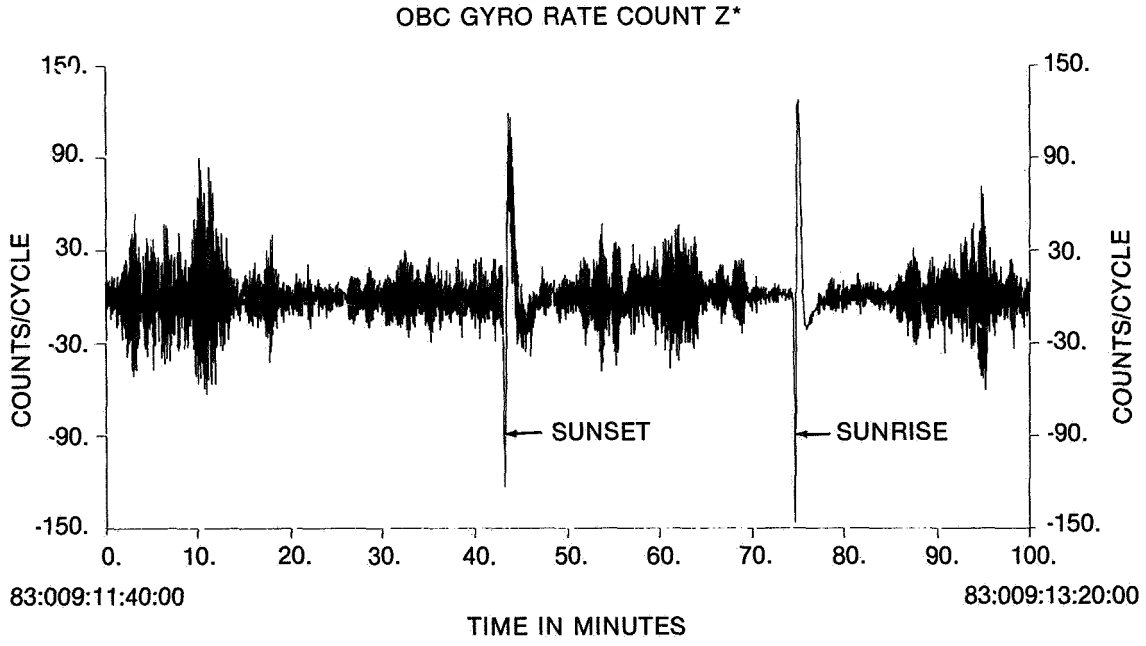
The significant attitude excursions experienced by the LANDSAT vehicles in response to the TES disturbance have aroused considerable concern for future satellite missions. This is especially true for the Upper Atmospheric Research Satellite (UARS) and the TOPography EXperiment (TOPEX) satellite since both vehicles utilize LANDSAT heritage. In particular, the TES disturbance has been predicted to be a dominant source of attitude perturbation for the TOPEX spacecraft [*Dennehy et al.*, 1988]. *Dennehy et al.* [1990] have analyzed the attitude pointing performance of the TOPEX spacecraft when subjected to the TES disturbance. Their analysis found that the TES disturbance was large enough to cause the TOPEX spacecraft to temporarily exceed its Normal Mission Mode (NMM) attitude pointing requirements. Consequently, some degraded performance of the primary scientific instrument will be experienced for a brief period of time. Thus, for TOPEX and other future satellites, a need is established to determine the on-orbit magnitude of the TES disturbance in order to compensate science data taken during degraded attitude periods. Temperature data could be used in one of two ways to counteract the TES disturbance. The first methodology would utilize the temperature data in an on-board implementation to provide adjustments in order to accommodate the TES disturbance torque. The second procedure would employ the use of temperature telemetry data to drive a TES disturbance model so *a postori* attitude reconstruction could be performed on the ground.



*Where: 1 count = 0.05 Arseconds, and 1 cycle = 0.512 Seconds

S-019.528

Figure 1: On-Orbit LANDSAT-4 Sunrise/Sunset Disturbance (Roll Axis)



*Where: 1 count = 0.05 Arcseconds and, 1 cycle = 0.512 Seconds

S-019.529

Figure 2: On-Orbit LANDSAT-4 Sunrise/Sunset Disturbance (Yaw Axis)

TES DISTURBANCE MODEL

A mathematical representation for the TES disturbance torque for a typical satellite solar array has been developed by *Dennehy et al.* [1990]. This model is essentially the second time derivative of the array inertia multiplied by the angle through which the array bends. For the TES disturbance model, the solar array is assumed to be a thin, cantilevered beam with mass M and length L. The general form of the disturbance torque is expressed as [*Dennehy et al.*, 1990]:

$$T_b = 2\dot{f}(t-\tau)\dot{u}(t-\tau) + \ddot{f}(t-\tau)u(t-\tau) \tag{1}$$

where $u(t-\tau)$ is a unit *step* function defined as:

$$u(t-\tau) = \begin{cases} 0 & \text{if } t < \tau \\ 1 & \text{if } t \geq \tau \end{cases} \tag{2}$$

and $\dot{u}(t-\tau)$ is a *Dirac* function described as:

$$\dot{u}(t-\tau) = \begin{cases} 1 & \text{if } t = \tau \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

If the array is broken into n pieces, where n is sufficiently large, then a general expression for the function $f(t-\tau)$ can be formulated as:

$$f(t-\tau) = \sum_{i=1}^{n} [\frac{l_i m_i d}{\alpha_{cte}\Delta T}(1 - \cos[\frac{l_i \alpha_{cte}\Delta T}{d}])] \tag{4}$$

where

$$\begin{aligned} d &\equiv \text{thickness of the array} \\ \tau &\equiv \text{point of umbra entrance or exit} \\ \Delta T &\equiv \text{thermal gradient across the array} \\ \alpha_{cte} &\equiv \text{material coefficient of thermal expansion} \end{aligned}$$

The parameters l_i and m_i are determined using the following relations:

$$l_i = \frac{L}{n}(i - \frac{1}{2}) \tag{5}$$

$$m_i = \frac{M}{n} \tag{6}$$

Substituting Equations 2, 3 and 4 into Equation 1 yields a standard expression for the TES disturbance torque:

$$T_b = \begin{cases} 0 & \text{if } t < \tau \\ T_\varsigma + \frac{\ddot{\Delta T} H_b}{\dot{\Delta T}} - \frac{2\dot{\Delta T} H_b}{\dot{\Delta T}} & \\ + \sum_{i=1}^{n} \frac{l_i^3 m_i \alpha_{cte} \dot{\Delta T}^2}{d\Delta T} \cos[\frac{l_i \alpha_{cte}\Delta T}{d}] & \text{if } t \geq \tau \end{cases} \tag{7}$$

where

$$T_\zeta = \begin{cases} 2H_b & \text{if } t = \tau \\ 0 & \text{if } t > \tau \end{cases} \tag{8}$$

and H_b is the array momentum given as:

$$H_b = \begin{cases} 0 & \text{if } t < \tau \\ \sum_{i=1}^{n} -\frac{l_i m_i d\dot{\Delta T}}{\alpha_{cte}\Delta T^2}(1 - \cos[\frac{l_i \alpha_{cte}\Delta T}{d}] & \\ -\frac{l_i \alpha_{cte}\Delta T}{d}\sin[\frac{l_i \alpha_{cte}\Delta T}{d}]) & \text{if } t \geq \tau \end{cases} \tag{9}$$

It is apparent from Equations 7, 8 and 9 that the TES disturbance torque is a function of not only the thermal gradient across the array, but is also dependent on the first and second time derivatives of the thermal gradient, $\dot{\Delta T}$ and $\ddot{\Delta T}$. Thus in order to predict the magnitude of the TES disturbance using the mathematical model, an array temperature gradient profile is necessary.

OPTIMAL TEMPERATURE ESTIMATION

The dependence of the TES disturbance torque model on the successive derivatives of the thermal gradient motivates the need for accurate temperature determination. The thermal response of a solar array may be predicted using numerical techniques such as those presented in *Dennehy et al.* [1990]. Such methods include using the Thermal Radiation Analysis SYStem (TRASYS) and the Systems Improved Numerical Differencing Analyzer (SINDA) software packages, as well as solving the one-dimensional heat equation

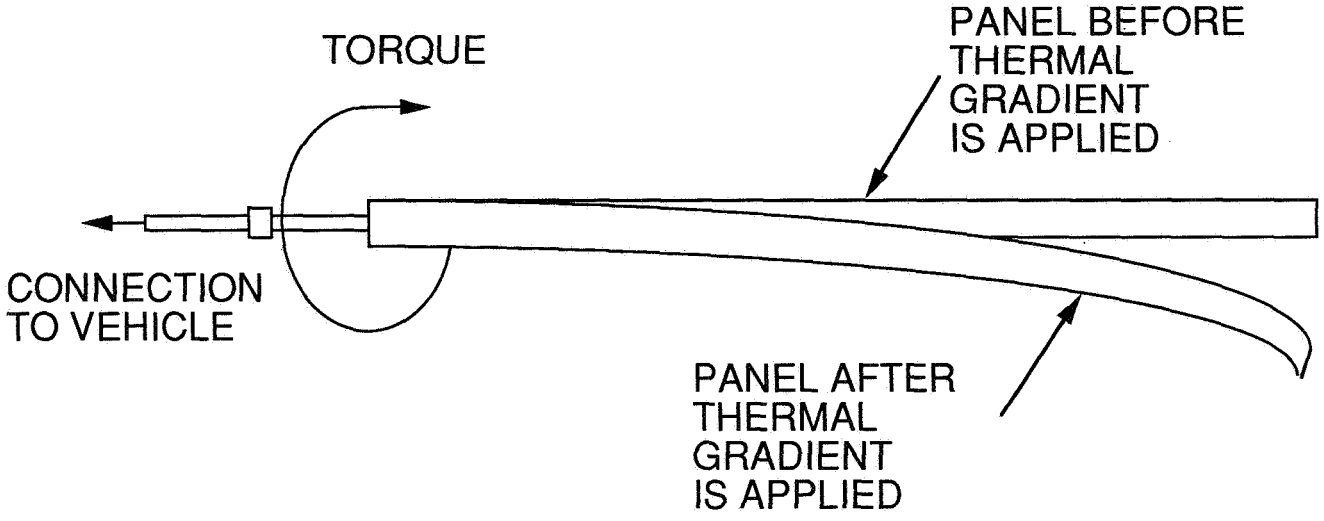


Figure 3: Solar Array Structural Deflection Due to Applied Thermal Gradient

using finite difference approximations. However, to measure the true thermal response of a solar array, thermal sensors are used during on-orbit spacecraft operation. Thermal sensors utilized for spaceflight applications include thermisters and platinum resistance thermometers (PRT). Thermisters are accurate to about $\pm 4°C$ while the accuracy of a PRT is approximately $\pm 1°C$. If a distributed network of thermal sensors exists on each surface of the array, as illustrated in Figure 4 for the front panel, a measurement of the surface temperatures at those prescribed locations is obtained. It would be nice to use all the available surface temperature information to estimate the array surface temperature at a desired location. Thus a technique, based on the Gauss-Markov Theorem, is described to optimally estimate array surface temperatures. Then once the temperature estimate for each surface is determined, the thermal gradient may be predicted by differencing the front and back surface temperatures at a consistant specified location.

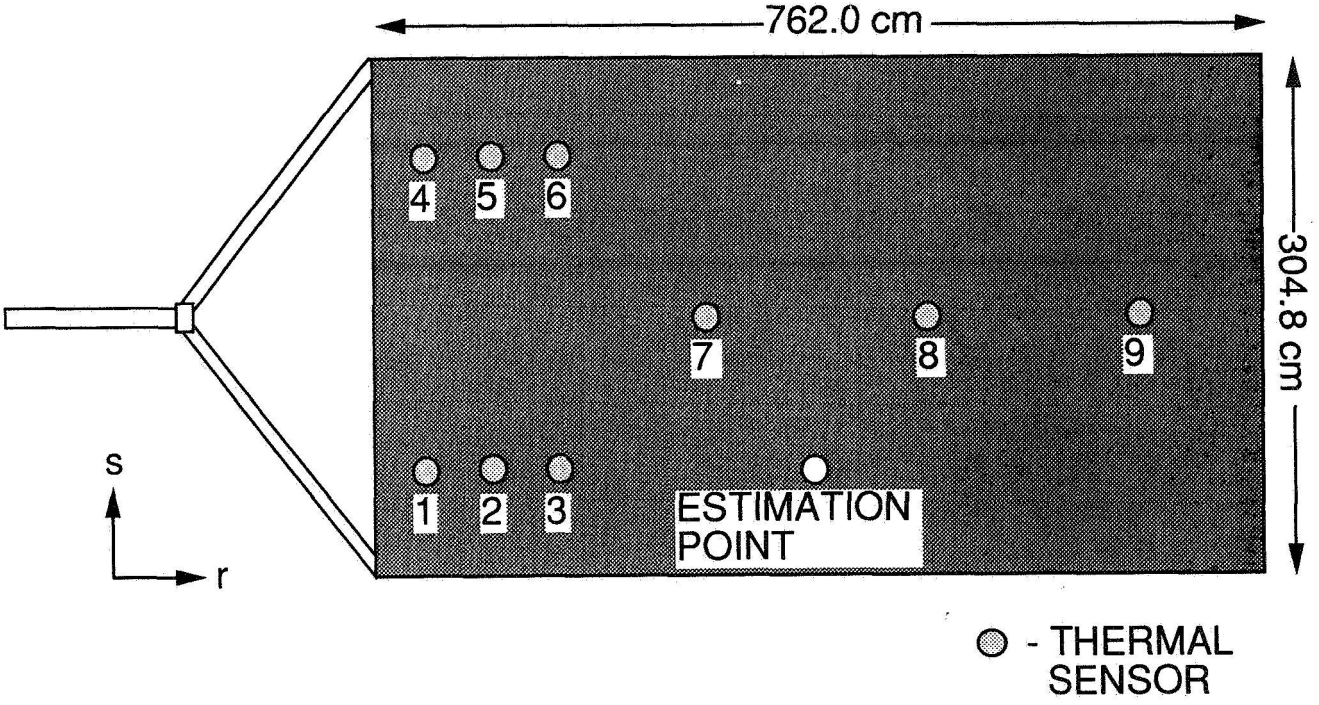


Figure 4: 9-Element Distributed Thermal Sensor Network on a Typical Solar Array (Front Surface)

The Gauss-Markov theorem provides a linear minimum mean square estimate of a vector x with n components given a set of m observations, Θ. The estimator, given by *Liebelt* [1967], is stated as follows:

$$\hat{x} = C_{x\Theta} C_{\Theta}^{-1} \Theta \tag{10}$$

where

$$\begin{aligned} C_{x\Theta} &\equiv E(x\Theta^T)\ (n \times m\ matrix) \\ C_{\Theta} &\equiv E(\Theta\Theta^T)\ (m \times m\ matrix) \\ E &\equiv \text{expected value operator} \end{aligned}$$

The error matrix associated with the estimate of x is given as [*Liebelt*, 1967]:

$$C_e = C_x - C_{x\Theta} C_{\Theta}^{-1} C_{x\Theta}^T \tag{11}$$

with

$$C_x \equiv E(xx^T)\ (n \times n\ matrix)$$

If the expected value of the estimate, $E(\hat{x})$, is equal to x (i.e. the average of the estimate is equal to the true value) then $\hat{x}$ is a linear minimum variance unbiased estimate and C_e is the covariance matrix of the estimate. The error gives a indication of how the estimate is dispersed from the true value. If the error is small then the estimate approaches the true value.

Equations 10 and 11 form the basis of the optimal estimation method developed by *Bretherton et al.* [1976]. For their analysis, the Gauss-Markov theorem was utilized to estimate the value of a two-dimensional scaler variable at a specified location given measurement data at a limited number of positions. A linear form of the observations is assumed and can be expressed as:

$$\varphi_i = \Theta(r,s) + \epsilon_i \tag{12}$$

for $i = 1, ..., N$ where

$$\begin{aligned} \varphi_i &\equiv i^{th} \text{ measurement} \\ \epsilon_i &\equiv i^{th} \text{ measurement error} \\ N &\equiv \text{total number of observations} \\ \Theta(r,s) &\equiv \text{scaler variable at position } (r,s) \end{aligned}$$

Furthermore, the assumption is made that the measurement errors are uncorrelated and independent of φ. Under these assumptions, *Bretherton et al.* [1976] applied the Gauss-Markov theorem, Equations 10 and 11, to obtain the resulting estimation equation given as follows:

$$\hat{\Theta} = \bar{\Theta} + \sum_{i=1}^{N} C_{xi}[\sum_{j=1}^{N} A_{ij}^{-1}(\varphi_j - \bar{\Theta})] \tag{13}$$

where

$$\begin{aligned} \bar{\Theta} &\equiv \text{estimated mean of the observations} \\ A_{ij} &\equiv \text{covariance between all pairs of observations} \\ C_{xi} &\equiv \text{covariance between the estimate and the } i^{th} \text{ observation} \end{aligned}$$

and the associated error matrix, C_e given as:

$$C_e = C_x - \sum_{i=1}^{N}\sum_{j=1}^{N} C_{xi}C_{xj}A_{ij}^{-1} \tag{14}$$

The estimated mean is computed under the condition that the sum of the weighted measurements is zero and is determined by the following equation [*Bretherton et al.*, 1976]:

$$\bar{\Theta} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}^{-1}\varphi_j}{\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}^{-1}} \tag{15}$$

Equations 13, 14 and 15 will be used to provide an optimal estimate of a solar array temperature at a prescribed location and the error associated with the estimate in order to demonstrate the estimation technique.

APPLICATION

The key to implementing the optimal estimation technique is the determination of both the C_x matrix and an analytic weighting function to scale the variance of the data. The weighting function is necessary to compute numerical values for the C_{xi} and A_{ij} matrices. The C_x matrix is generally unknown but can be approximated by the variance of the given data set. The numerical computation of C_x for this analysis was performed using the following equation [*Bretherton et al., 1976*]:

$$C_x = \sigma_\varphi^2 + \frac{(1 - \sum_{i=1}^{N}\sum_{j=1}^{N} C_{xi}A_{ij}^{-1})^2}{\sum_{i=1}^{N}\sum_{j=1}^{N} A_{ij}^{-1}} \tag{16}$$

where σ_φ is the standard deviation of the measurement data given as:

$$\sigma_\varphi = \sqrt{\frac{\sum_{i=1}^{N}(\varphi_i - \bar{\varphi})^2}{N-1}} \tag{17}$$

with

$$\bar{\varphi} = \frac{\sum_{i=1}^{N}\varphi_i}{N} \tag{18}$$

The last term on the right hand side of Equation 16 accounts for uncertainties associated with the estimated mean.

A weighting function was selected, for this study, to weight the measurements according to their spatial location with respect to one another and to the desired position of the estimate. This type of weighting function can be used as a first cut statistical model given no *a priori* knowledge of the data statistics. The estimation technique can, however, easily accommodate more complex statistical models if desired. An analytical expression for the weighting function is given as:

$$W_{ij} = 0.2(\gamma - \bar{r_{ij}}^2 - \bar{s_{ij}}^2)\exp^{-\sqrt{\bar{r_{ij}}^2 + \bar{s_{ij}}^2}} \tag{19}$$

where

γ $\equiv$ measurement degradation factor

$\bar{r_{ij}}$ $\equiv$ scaling parameter between the i^{th} and j^{th} observations in the r direction

$\bar{s_{ij}}$ $\equiv$ scaling parameter between the i^{th} and j^{th} observations in the s direction

The parameter γ is introduced to change the quality of the observations. If γ is set equal to 5.0 then a maximum correlation of 1.0 will exist when the condition $i = j$ is satisfied in

Equation 19. As γ linearly decreases, the maximum attainable correlation also decreases in a linear fashion. The scaling parameters, $\bar{r_{ij}}$ and $\bar{s_{ij}}$, are calculated using the following equations:

$$\bar{r_{ij}} = \frac{r_i - r_j}{r_{scale}} \tag{20}$$

and

$$\bar{s_{ij}} = \frac{s_i - s_j}{s_{scale}} \tag{21}$$

The variables r_{scale} and s_{scale} can be specified to determine an effective range of data influence (decorrelation scale) or set to the dimensions of the spatial area over which the measurements are confined. For the present study, the latter condition is used. The variables r_i and s_i denote the spatial location of the i^{th} observation location while the variables r_j and s_j indicate the spatial position of the j^{th} point. Thus a spatially weighted covariance can be computed between the point of estimation and the measurements, C_{xi}, and between the observations themselves, A_{ij}, given the weighting function. The calculation of C_{xi} can be expressed by:

$$C_{xi} = W_{xi}\sigma_{\varphi}^2 \tag{22}$$

where the subscript x is used to denote the desired location, (r_x, s_x), of the estimate, while the weighted observation matrix, A_{ij}, is determined from the following equation:

$$A_{ij} = W_{ij}\sigma_{\varphi}^2 + \sigma_{\epsilon}^2\delta_{ij} \tag{23}$$

where δ_{ij} is the *Kronecker* delta function expressed as:

$$\delta_{ij} = \begin{cases} 0 & \text{if i} \neq \text{j} \\ 1 & \text{if i} = \text{j} \end{cases} \tag{24}$$

and σ_{ϵ} the standard deviation of the error.

The one drawback of using this technique is the inversion of the observation covariance matrix, A_{ij}. If a large number of observations exist then the dimension of the covariance matrix becomes cumbersome and hard to numerically invert. However, to speed the nu-

Table 1: Thermal Sensor Locations and Temperature Measurements for Cases 1-3

Sensor No.	Sensor Position		Temperature $°C$		
	r (cm)	s (cm)	Case 1	Case 2 (Thermister)	Case 3 (PRT)
1	31.75	76.2	72.8795	75.6214	73.8198
2	95.25	76.2	72.9475	75.1302	72.4815
3	158.75	76.2	72.9530	69.0273	73.5481
4	31.75	228.6	72.8510	76.0959	73.4943
5	95.25	228.6	72.9433	70.5750	72.3178
6	158.75	228.6	72.9525	71.1157	73.8453
7	285.75	152.4	72.9534	72.4241	72.9780
8	476.25	152.4	72.9534	72.7547	72.4672
9	666.75	152.4	72.9534	72.8162	72.7367

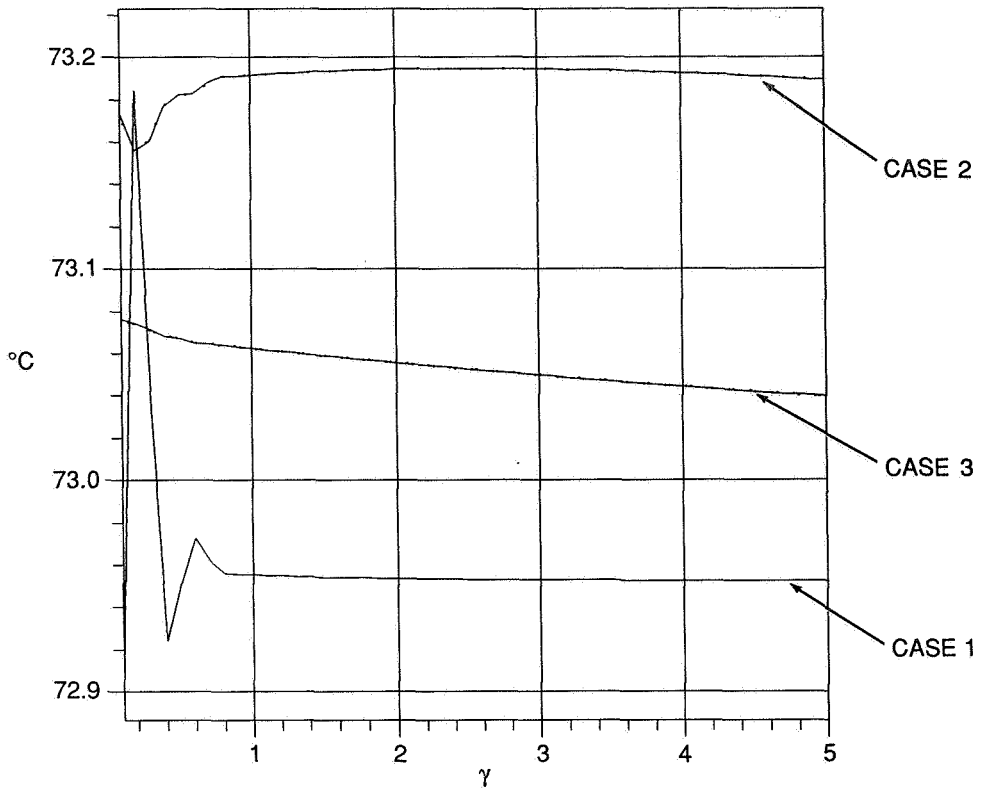


Figure 5: Temperature Estimates Versus Measurement Degradation Factor for Cases 1-3

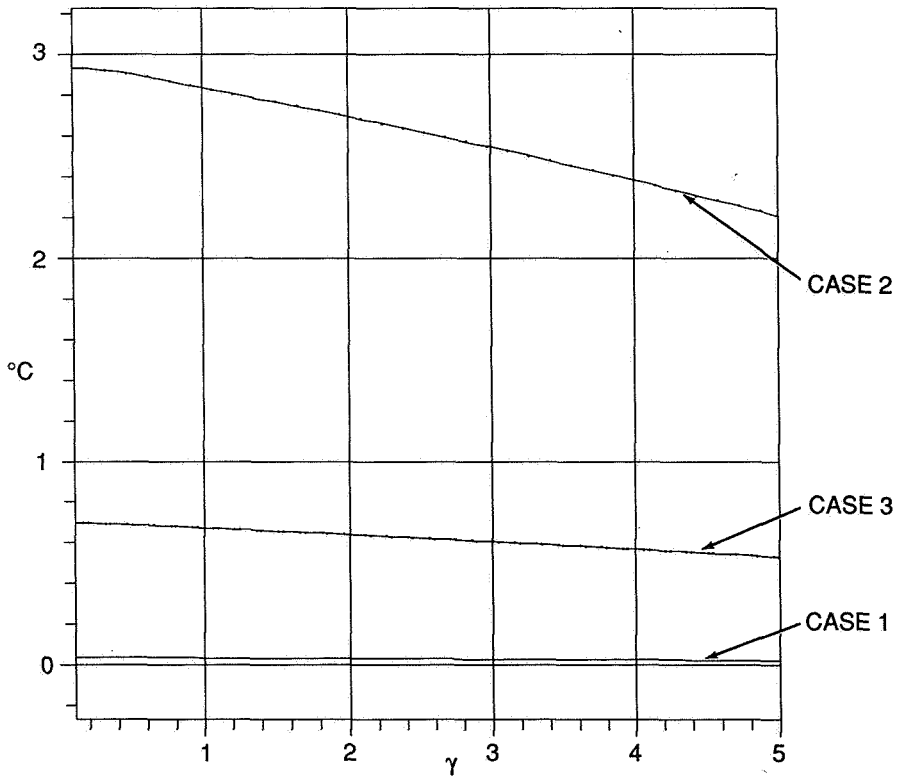


Figure 6: Standard Deviation of Temperature Estimates Versus Measurement Degradation Factor for Cases 1-3

merical inversion process, the order of the covariance matrix could be reduced, keeping only those observations having a correlation value above a specified limit.

RESULTS

For this study, the arrangement of the thermal sensors on a typical solar array front surface, labeled 1 through 9 as shown in Figure 4, represent the spatial location of the temperature observations. The length (762.0 cm) and width (304.8 cm) of the array represent the spatial area scales, r_{scale} and s_{scale} respectively. To demonstrate the technique, temperature measurements at the nine locations, all at a single point in time, are used. The factor γ is set equal to 5.0 to provide a maximum correlation of 1.0. The technique is first used with data that is assumed true with no measurement error, e.g. $\epsilon_i = 0$. Thus the error matrix will represent the covariance of the unbiased estimate. The temperature data and the locations of the thermal sensors are listed in Table 1 under the Case 1 heading. The prescribed point of estimation is positioned at (381,50.8) cm for this and all examples. Applying the technique, the computed estimate is 72.9523°C with a standard deviation of 0.023°C. The estimated temperature is consistant with the observed data and the standard deviation of 0.023°C represents a small deviation from the true temperature at the estimation point. A second example is illustrated by corrupting the true temperature measurements with error. In one case the sensors are considered to be thermisters and in another the PRT sensor is used. The standard deviation of the error is assumed to be 4°C for the thermister and 1°C for the PRT. Case 2 shows the thermister simulated temperature data and Case 3 represents the PRT data as shown in Table 1. The temperature estimates for the thermister and PRT data are 73.1893°C and 73.0396°C with standard deviations of 2.205°C and 0.530°C respectively. The estimates for each of the 3 cases fall within

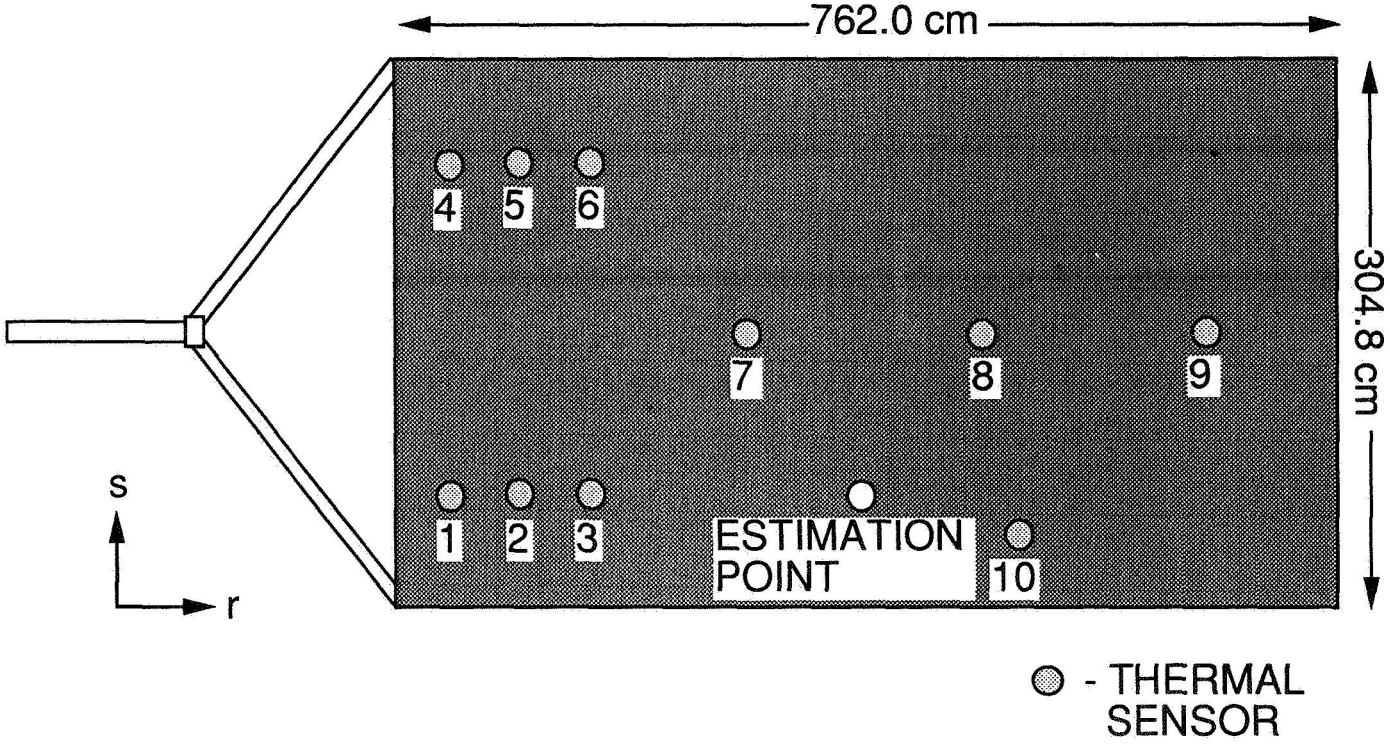


Figure 7: 10-Element Distributed Thermal Sensor Network on a Typical Solar Array (Front Surface)

the range of the observed data with Case 2 exhibiting the largest standard deviation. The standard deviation for Case 2 is a result of the large variance associated with the simulated thermister data set.

If the parameter γ is allowed to approach 0 from its maximum value of 5.0, the quality of the observations is degraded. Furthermore, as γ decreases the estimate is expected to degrade with an increase in the standard devation. To test this hypothesis, estimates and standard deviations for the 3 test cases were recomputed for values of γ ranging from 0.1 to 5.0 in 0.1 increments. Figures 5 and 6 show the estimated temperatures and standard deviations, respectively, for the true data (Case 1), the thermister data (Case 2) and the PRT data (Case 3). An inspection of Figure 5 reveals that the temperature estimate for Case 1 remains essentially constant, $72.955°C$, for $\gamma \geq 0.8$. In the range of $\gamma < 0.8$, the Case 1 temperature estimates reach a maximum of $73.185°C$ at $\gamma = 0.2$ and a minimum of $72.92°C$ at $\gamma = 0.4$. The large deviations occur as the elements of the A_{ij} matrix approach small values resulting in an inverse matrix with large components. Case 2 also shows a non-linear change in the temperature estimate for $\gamma < 0.8$ with a minima at $73.155°C$. The estimate reaches an approximate steady state value of $73.19°C$ with a slight downward trend for $\gamma > 3.5$. Case 3 temperature estimates indicate a linear decrease in temperature from $73.7°C$ to $73.4°C$ with a slope of approximately $0.065\frac{°C}{\gamma_{increment}}$ for $\gamma \geq 0.4$. When $\gamma < 0.4$,the slope increases to $-0.33\frac{°C}{\gamma_{increment}}$. The standard deviations for the 3 Cases, as a function of γ, are illustrated in Figure 6. The standard deviation for the true data (Case 1) remains essentially constant with a value of approximately 0.023°. The thermister data set (Case 2) shows a linear decrease in standard deviation from $2.9°C$ to $2.2°C$ with a

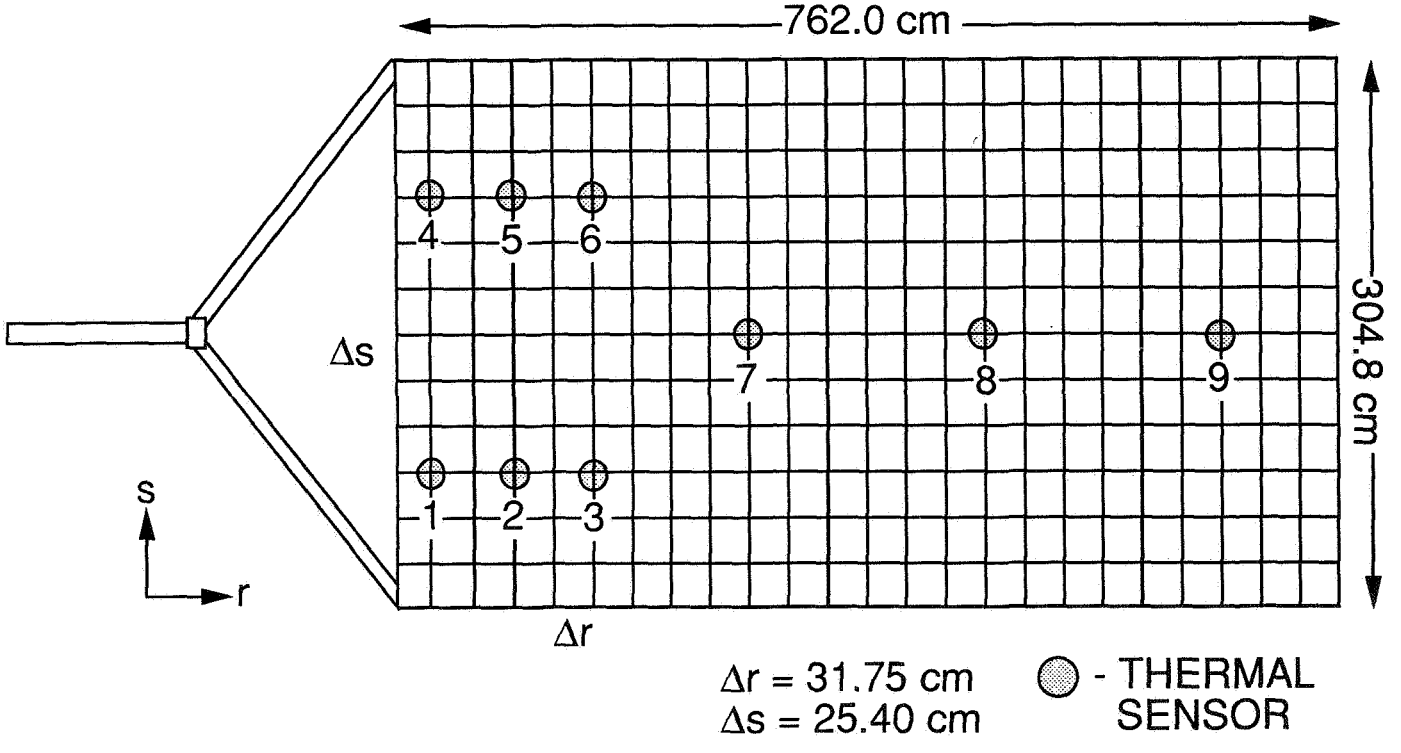


Figure 8: Grid Point Locations Superposed on 9-Element Distributed Thermal Sensor Network

slope of $0.14\frac{^{\circ}C}{\gamma_{increment}}$ as γ increases. The standard deviation of the PRT data also shows a linearly decreasing trend but with a much more subtle slope of $0.04\frac{^{\circ}C}{\gamma_{increment}}$. As expected the estimates for the 3 Cases degrade with an associated increase in the standard deviation as γ approaches a small value.

Another example is illustrated for the case when a hot spot exists on the solar array. For this example a junction box is placed at the location (508,38.1) cm. An additional thermal sensor is also placed at the junction box location. It is assumed that the temperature of the junction box is $82.0^{\circ}C$. Figure 7 shows the new thermal sensor configuration for the front array panel and Table 2 lists the position of the sensors along with the observation data. The parameter γ is equal to 5.0 to provide a maximum correlation of 1.0. The optimal temperature estimation routine provides an estimate of $78.1648^{\circ}C$ and a standard deviation of $1.312^{\circ}C$ for the true temperature, Case 4. Estimates were also computed assuming that the data were measured using a thermister (Case 5) and a PRT (Case 6). The estimates are $75.0298^{\circ}C$ and $77.5123^{\circ}C$ respectively with standard deviations of $2.181^{\circ}C$ and 1.500°. Table 3 summarizes the estimates and the standard deviations for each case.

Since the optimal temperature estimation technique has the ability to provide a temperature estimate at any location, a grid of temperature estimates can be generated to characterize the temperature profile of the entire array using a finite set of data observations. Using only Case 1 from the above discussion this capability will be demonstrated. The grid point locations are illustrated in Figure 8 where the incremental distances between nodes are $\Delta r = 31.75\ cm$ and $\Delta s = 25.4\ cm$. In order to assess the accuracy of the estimation technique, grid points were colocated with the thermal sensor positions as shown in Figure 8. The contoured array front surface temperature profile is shown in Figure 9. The standard deviations associated with the array surface temperature estimates are presented in Figure 10. Table 4 depicts the true and estimated temperatures at the nine thermal sensor locations. A comparison using the percent error between the true and estimated values shows excellent agreement with the largest deviation being 0.309%. The standard deviation for each temperature estimate is also small indicating a small dispersion from the true temperature, thus providing a reasonable estimate for this example.

Table 2: Thermal Sensor Locations and Temperature Measurements for Cases 4-6

Sensor No.	Sensor Position		Temperature $^{\circ}C$		
	r (cm)	s (cm)	Case 4	Case 5 (Thermister)	Case 6 (PRT)
1	31.75	76.2	72.8795	74.4008	72.2307
2	95.25	76.2	72.9475	71.5595	73.5006
3	158.75	76.2	72.9530	71.0906	72.8149
4	31.75	228.6	72.8510	75.2109	73.2364
5	95.25	228.6	72.9433	69.3491	72.4154
6	158.75	228.6	72.9525	76.6270	72.7314
7	285.75	152.4	72.9534	75.1502	73.6916
8	476.25	152.4	72.9534	75.2135	72.6657
9	666.75	152.4	72.9534	75.0416	72.4325
10	508.00	38.1	82.0000	79.8574	82.2281

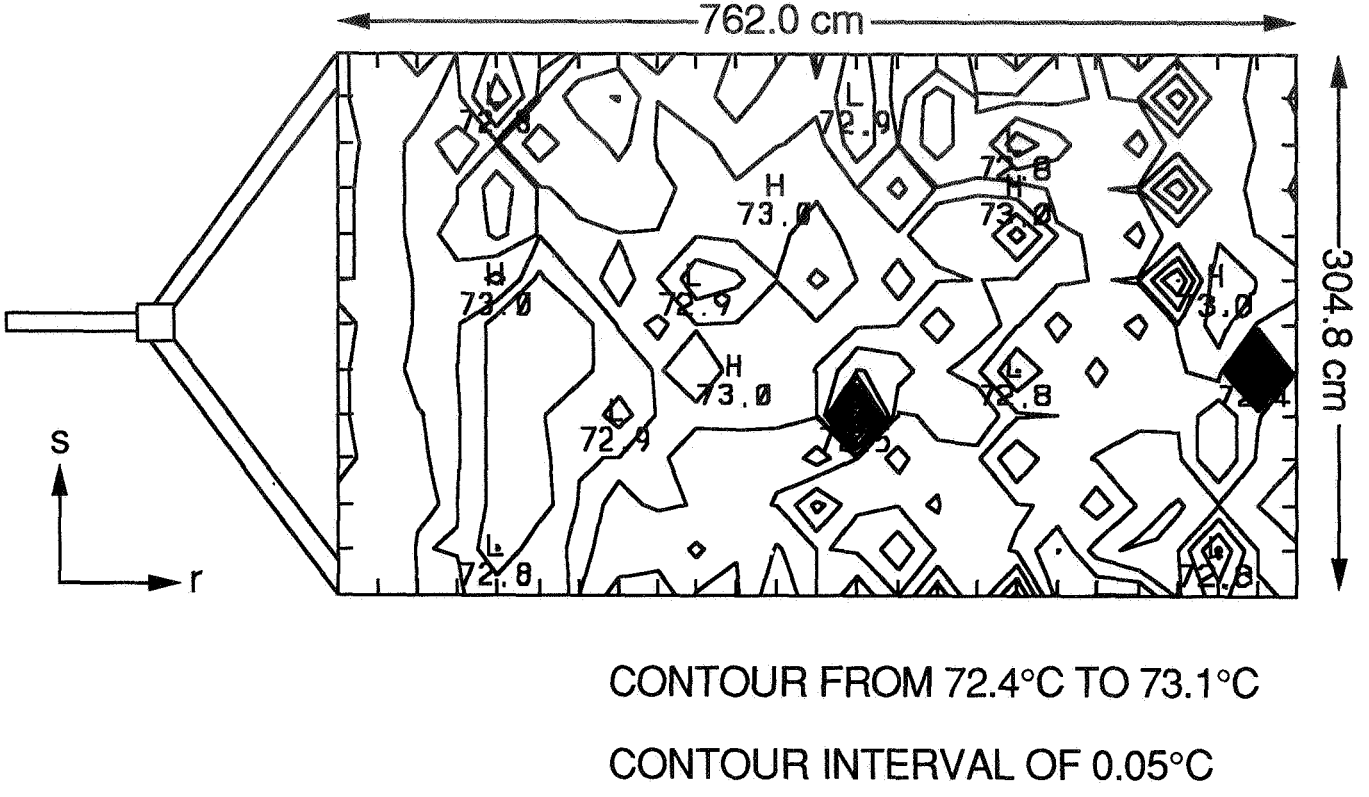


Figure 9: Contoured Array Front Surface Temperature Estimates for 9-Element Distributed Thermal Sensor Network

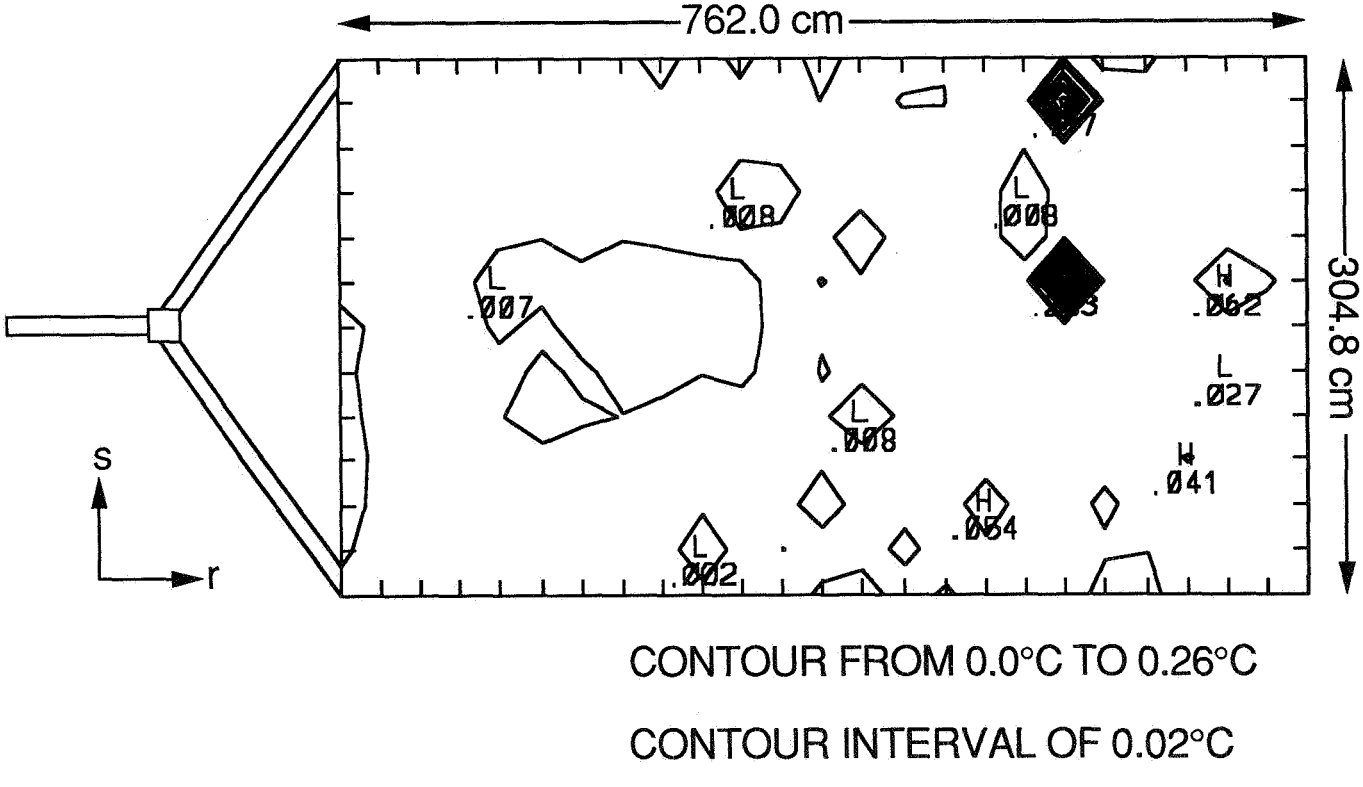


Figure 10: Contoured Standard Deviations Associated with Array Front Surface Temperature Estimates for 9-Element Distributed Thermal Sensor Network

Table 3: Summary of Temperature Estimates and Standard Deviations for Cases 1

Case No.	Temperature Estimate $^{\circ}C$	Standard Deviation $^{\circ}C$
1	72.9523	0.023
2	73.1893	2.205
3	73.0396	0.530
4	78.1648	1.312
5	75.0298	2.181
6	77.5123	1.500

Table 4: Accuracy of Temperature Estimates Compared to True Observations

Sensor No.	Temperature True $^{\circ}C$	Temperature Estimate $^{\circ}C$	Standard Deviation $^{\circ}C$	% Error
1	72.8795	72.8092	0.026	0.096
2	72.9475	72.7262	0.023	0.303
3	72.9530	73.0036	0.022	0.069
4	72.8510	72.8480	0.026	0.004
5	72.9433	72.8636	0.024	0.109
6	72.9525	72.7269	0.022	0.309
7	72.9534	72.8073	0.018	0.200
8	72.9534	72.9120	0.029	0.057
9	72.9534	72.9622	0.033	0.012

CONCLUSIONS

An optimal temperature estimation technique has been described and used to estimate the surface temperature of a satellite solar array at a prescribed location. The technique also provides error information relative to the estimated variable. This technique is capable of determining array surface temperatures at any location, with reasonable accuracy, from a finite set of observational data. Applying the procedure to both surfaces of the array, as a function of time, and differencing the surface temperature estimates will result in an estimated thermal gradient profile. The thermal gradient estimates can then be utilized to drive the TES disturbance model in order to evaluate the true nature of the TES disturbance.

REFERENCES

1. Anonymous, LANDSAT-4 Flight Evaluation Report, 16 October 1982 to 16 January 1983, GE Document 83SDS4203, 16 January 1983, (Contract NAS 5-25300).

2. Bretherton, F.P., R.E. Davis and C.B. Fandry, "A Technique for Objective Analysis and Design of Oceanographic Experiments Applied to MODE-73," *Deep Sea Research*, Vol. 23, pp. 559-582, 1976.

3. Dennehy, C.J., T. Kia and R.V. Welch,"Attitude Determination and Control Subsystem for the TOPEX Satellite," *AIAA Paper* 88-4129, 1988.

4. Dennehy, C., D. Zimbelman and R. Welch,"Sunrise/Sunset Thermal Shock Disturbance Analysis and Simulation for the TOPEX Satellite," *AIAA Paper* 90-0470, 1990.

5. Hamsath, N., and P. M. Bainum,"The Development of an Environmental Disturbance Model for Large Space Structures after the Onset of Thermal Shock," *AIAA Paper* 86-2123, 1986.

6. Liebelt, P.B., *An Introduction to Optimal Estimation*, Addison-Wesley Publishing Company, Massachusetts, 1967.

7. Vigneron, F.R. and R.A. Millar,"Flight Performance of the Stabilization System of the Communications Technology Satellite," *Journal of Guidance and Control*, Vol. 1, 1978, pp. 404-412.

FLIGHT MECHANICS/ESTIMATION THEORY SYMPOSIUM

MAY 22-24, 1990

SESSION 5

The Generalized Tsiolkovsky Equation

V. N. Dvornychenko[1]
South Pasadena, California

Abstract

The performance of a rocket, say a chemical rocket, can be greatly improved by dynamically transferring energy from one part of the propellant to another. Although with present technology the achievable degree of energy concentration is low, payload increases on the order of 20% are theoretically possible. With unlimited energy redistribution increases of a factor of two are possible.

Introduction

Some eighty-five years have passed since K.E.Tsiolkovsky published the equation for the ideal-velocity-gain of a rocket:

$$V_i = V_e \ln(m_o / m_1) \qquad \text{(Tsiolkovsky)} \qquad \ldots\ldots(1)$$

Here V_i is the ideal-velocity-gain.
V_e is the effective exhaust velocity of the mass ejecta.
m_o and m_1 are the initial and final rocket masses, respectively.

It is customary to introduce two derived quantities, the mass-ratio (μ) and the specific-impulse (Isp):

$$\mu \overset{d}{=} m_0/m_1 \ , \ I_{sp} \overset{d}{=} V_e/g \qquad \ldots\ldots(2)$$

where g is the earth's gravitational constant.

Typical values for I_{sp} are 200 to 450, corresponding to V_e of 1.96 to 4.41 km/sec. The upper range is achieved by high-performance LOX-LH (liquid-oxygen/liquid-hydrogen) propulsion systems such as the Space Shuttle. Exotic oxydizers such as ozone or fluorine can push this into the low 500's, but this seems to be about the chemical propulsion limit. Usable and practical mass-ratios are in the range 5 to 10. For structural reasons values beyond 20 are very difficult to achieve.

The important thing about (1) (and this was most emphatically noted by the early space pioneers) is that the mass-ratio, and therefore the size of the rocket, increases exponentially with the required ideal-velocity-gain. A single-stage-to-orbit using LOX-LH requires a mass-ratio of about 10 (after performance losses are taken into account). This is difficult but achievable. But a low performance propellant (Isp = 200) would require a mass-ratio of 150, which is essentially impossible.

Let us now look at an apparently unrelated situation: the fission of a particle (m_0) into two particles (m_1,m_2) with a release of energy (E), which goes wholly into kinetic energy. The result, familiar from undergraduate physics, is

$$V_1 = \sqrt{\frac{2E/m_1}{1 + m_1/m_2}} \qquad \ldots\ldots(4)$$

where V_1 is the velocity of particle one.

---

1 The author is employed by Northrop ESD, 500 E. Orangethorpe Avenue, Anaheim, CA. 92801.
His duties at Northrop are in the area of image understanding. Member AIAA.

Suppose now that all the energy resides initially in particle two, and that this energy has a specific energy (energy per unit mass) of

$$\sigma = E/m_2 = V_e^2/2 \qquad \ldots.(5)$$

Substituting this into (4) and introducing the mass-ratio we obtain

$$V_1 = V_e(\sqrt{\mu} - 1/\sqrt{\mu}) \qquad \ldots..(6a)$$

as opposed to the classical (Tsiolkovsky) equation

$$V_i = V_e \ln(\mu) \qquad \ldots.(6b)$$

As expected, (6a) always produces a greater value than (6b), as shown in Figure 1.

A natural question is the reason for this discrepancy. The usual explanation is that the classical rocket suffers by having to drag its propellant along. But this is only partly correct. A better answer lies in the inability of the classical rocket to transfer energy from one part of the propellant to another.

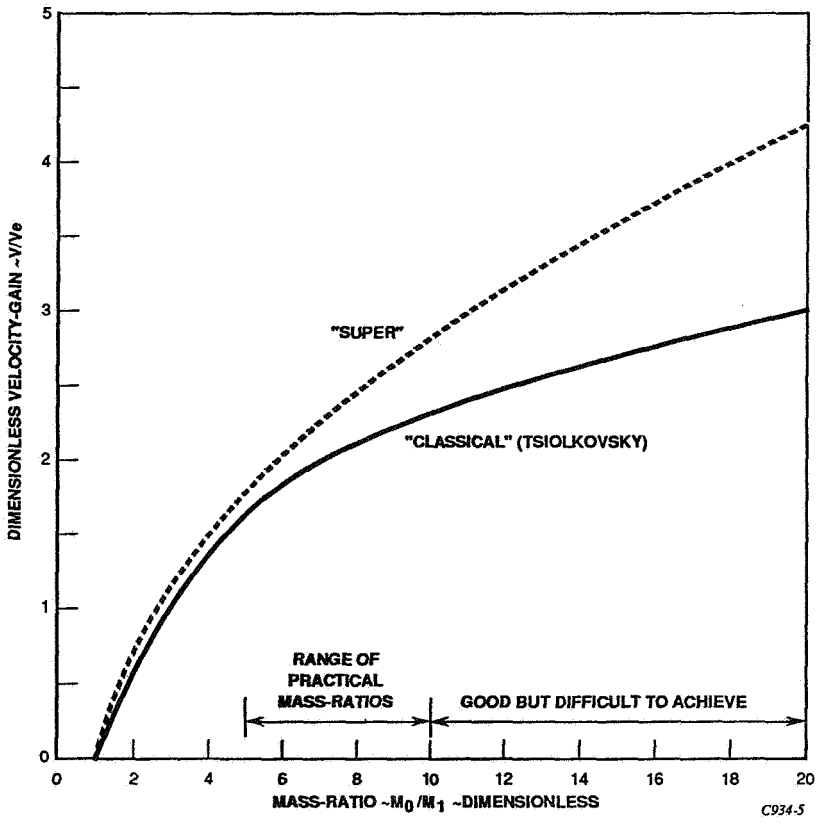


Figure 1. Performance Comparison of Classical and "Super" Rockets

emphatically, it does not need to occur instantaneously. Let us agree to call the rocket which obeys (6a) the "super" rocket. Figure 2 provides additional performance comparisons between the classical and super rocket. The first curve shows how much additional specific impulse is needed by the classical rocket to match the super, as a function of mass-ratio. The second curve shows how much the payload (final mass) is increased on the super with the same initial energy. For a mass-ratio of 20 the super has an 84% higher payload, which is equivalent to increasing the specific impulse in the classical rocket by 42%.

Methods of Energy Modulation

Do practical methods exist for transferring energy from one part of the propellant to another? The answer is, yes and no. Ways certainly exit, some may even be "practical". The problem is that it is not possible to achieve the high exhaust-velocity modulation required by the previous theory. Let us consider some of the ways available:

1. mechanical -- the energy is directly stored as rotational kinetic energy.
2. thermal -- the energy is stored as heat.
3. chemical -- the energy is concentrated by creating certain high-energy compounds or species; for instance, making ozone out of oxygen.
4. mixture-ratio shift
   -- the relative ratio of two substances is changed. This may be considered to be a special case of the "chemical" method.

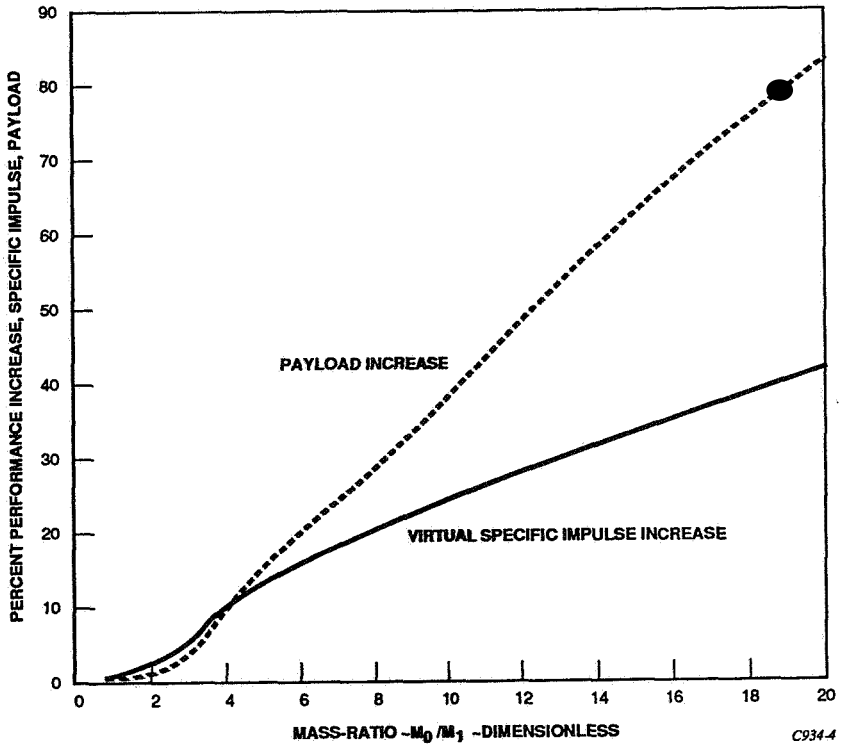


Figure 2. Percent Performance Increase for "Super"

We now show two things: 1) that (6a) is the upper limit of the velocity attainable by mass ejection of mass m_2 and expended energy E, and 2), that this limit is in fact attainable with unlimited energy redistribution.

Let V_e (m) be the variable exhaust velocity, taken as a function of the propellant mass remaining, and measured with respect to the space-fixed system in which the rocket is initially at rest. The total propellant mass is m_2 (as before) and the mass of the all-burned rocket is m_1. As before, we assume the total energy expended is E.

Then we have from the momentum and energy conservation equations :

$$\int_0^{m_2} V_e(m)dm = V_1 m_1 \qquad \text{....(7a)}$$

and

$$\int_0^{m_2} V_e^2\,(m)dm + V_1^2\, m_1 = 2E \qquad \text{....(7b)}$$

Let

$$\bar{V}_e = V_1 m_1/m_2 \qquad \text{.....(7c)}$$

and introduce the "loss function", L, via

$$L \overset{\text{d}}{=} \int_0^{m_2} (V_e(m) - \bar{V}_e)^2 dm \qquad \text{....(7d)}$$

From these we derive :

$$V_1 = V_e^*(\sqrt{\mu} - 1/\sqrt{\mu})\,\sqrt{1 - L/(V_e^{*2} m_2)} \qquad \text{.....(8)}$$

where V_e^* is essentially as given by (5), the star having been added to differentiate it from other uses of V_e .

Since L $\geqslant$ 0 it is clear that the velocity is maximized when L=0, in which case (8) becomes identical to (6a).Equation (7d) shows that to achieve the condition L=0 we must have the exhaust velocity constant in a space-fixed frame, rather than with respect to the rocket. The exhaust velocity must therefore be of the form

$$|V_e(t)| = |V_{e;0}| + |V_1(t)| \qquad \text{...(9)}$$

where $V_1(t)$ is the instantaneous rocket velocity and $V_{e;0}$ is the initial exhaust velocity.

Note that, as in the classical equation, there are no restrictions on the rate at which mass is expended. We can take all the time we want. We can even have bursts of mass ejection and pauses. Most

5. "exotic" -- this is a catch-all to include everything else, such as, electrical energy in capacitors, plasma, etc.

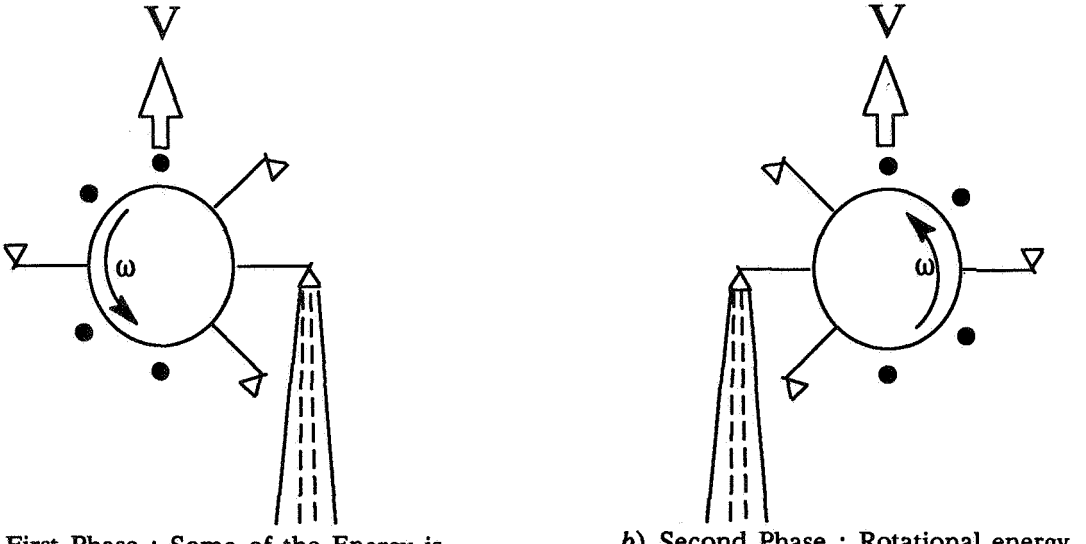
Surprisingly a conceptually simple mechanical method exists. Figure 3 shows the basic concept. The rocket is "flying-saucer-shaped", with rocket engines mounted on booms at the periphery. These are fired sequentially as the rocket rotates, so that the thrust maintains a space-fixed direction. During the initial phase energy is transferred into kinetic energy of rotation. During the subsequent phase opposing rockets are fired so that the energy is returned to the mass-ejecta stream. Because of structural limitations it is difficult to achieve peripheral rotation speeds beyond about 300 m/sec. Thus the velocity modulation is limited to about 7%, corresponding to about 14% modulation in energy density.

Comparable energy modulation is achievable using heat storage. Here the problem is not only a structural one -- high temperatures cause high pressures, requiring strong vessels -- but also a thermodynamic one: as the temperature is elevated disassociation causes energy extraction to drop.

Chemical methods do not provide a much higher energy modulation than the previous, mainly because the types of reactions in rockets are already among the most energetic. Using reactions such as the previously mentioned oxygen to ozone conversion a modulation of 20 to 30% might be possible.

Mixture-ratio shift is interesting because it is comparable in its simplicity to the mechanical method. The basic idea is to mix the ratio of two "reactants". These need not be fuel and oxidizer. Indeed the separation is a conceptual one, and they need not be physically isolated. The two species could be, 1) propellant, and 2) an unreactive tamper; or else 1) the stoichiometric part, and 2) the excess part. The modulation is generally lower than any of the previous.

As for "exotic" methods, I know of no practical ones. Perhaps electrical plasmas, in which atoms are stripped of their electrons, provide some potential.



a) First Phase : Some of the Energy is siphoned off into rotational Energy

b) Second Phase : Rotational energy returned to mass ejecta.

Figure 3 - Mechanical Means of Energy Modulation

The picture which emerges is this: the kinds of energy density modulation feasible are in the order of a few percent to a maximum of some 30%, rather than the factors of 25 to 100 required by the simple theory expounded in the first part of this paper. We thus need to modify the theory.

Case of Limited Energy Modulation

Suppose the total propellant (m_2) consists of two parts (m_3, m_4), where m_3 is the lower energy portion (say), and m_4 the higher. It is not necessary that these two parts be physically separated, as in different tanks; the separation may be a virtual one, in which m_3 represents that part of the propellant which is depleted in energy, while m_4 represents the energy-enhanced part. Let σ_3 and σ_4 be the corresponding specific energies. If an amount $dm_2 = dm_3 + dm_4$ of mass is ejected the amount of energy available is :

$$dE = \sigma_3 \ dm_3 + \sigma_4 \ dm_4 \qquad .. (10)$$

and the total momentum of the (differential) mass ejecta will be

$$dp = \sqrt{2dE * dm_2} \qquad .... (11)$$

and this amount of momentum will be imparted to the rocket.

Introduce now the mixture-ratio

$$r \overset{d}{=} dm_4/(dm_3 + dm_4) \qquad ....(12)$$

The velocity-gain of the rocket is now given by :

$$V_i = \sqrt{2}\int_0^{m_2} \sqrt{(1-r)\sigma_3 + r\sigma_4}\ dm/m \qquad ....(13)$$

where dm is the change in mass of the rocket and is the same as dm_2.

Our objective is to find the optimal mixture-ratio profile which maximizes (13). To do this we need to add the constraint

$$m_4 = \int_o^{m_2} rdm \qquad ....(14)$$

which says nothing more than at all-burned all of m_4 is used up.

In addition we need to observe the restraint

$$0 \leqslant r \leqslant 1 \qquad ....(15)$$

The optimization of (14) can now be accomplished using either the optimal control theory formulation or the older calculus of variations. The gist of the solution is that there are three phases:

1. during the first phase r=0.

2. during the second phase r is variable and given by the recipe.

$$r(m) = (\bar{\sigma}M_0M_1/m^2 - \sigma_3)/(\sigma_4 - \sigma_3) \qquad \ldots.(16)$$

where $\bar{\sigma}$ is the mean specific energy during this phase, and M_0 , M_1 are the initial and final masses during this phase.

3. during the third phase r=1.

Depending on the values of m_3 ,m_4 ,σ_3 and σ_4 some of these phases may be missing. In the case where all phases are present the total velocity gain is given by

$$V = \sqrt{2\sigma_4}\ \ln((m_0/m_1) * \sqrt{\sigma_3/\sigma_4}) \ + (\sqrt{2\sigma_4} - \sqrt{2\sigma_3})\ \ln\ (e * (m_1 + m_4 * (\sqrt{\sigma_4/\sigma_3} + 1))/(2m_0)) \qquad \ldots.(17)$$

In (11) we tacitly assumed that all the mass ejecta are imparted the same velocity. However, when gasses are mixed the energy divides on a molecular basis (equipartition principle). If this principle is pursued we need to slightly modify (13) by introducing into the integrand the factor Q defined by

$$Q(r) = \frac{(1-r)/\sqrt{w_3} + r/\sqrt{w_4}}{\sqrt{(1-r)/w_3 + r/w_4}} \qquad \ldots.(17)$$

where w_3 and w_4 are the average molecular weights of the (combustion) products of dm_3 and dm_4 . It can be shown that $Q(r) \leqslant 1$, so that this kind of process produces a smaller V_i . However, not only does this complicate the analysis, but as it is unlikely to accurately model the thermodynamics we do not investigate this approach further.

Some Numerical Results

Consider the case where all the propellant is originally the same, with specific energy σ_n. We assume it is possible to enrich some of this to a level σ_4 , in the process depleting some of it to level σ_3 . However, the degree of enrichment is limited to a level $\sigma_4 = \sigma_n\ (1 + \epsilon)$, where $\epsilon << 1$. (The degree of depletion is not assumed limited, and can go all the way to zero.) The theory of the previous section may be used to show that the resulting fractional velocity gain is given by ϵ /2. Or if the velocity is kept the same the fractional payload increase is given by $\epsilon * \ln(\mu)/2$. For example, if ϵ=0.2 (maximal enrichment is 20%) then the realizable velocity gain is 10%, or a payload increase of 23% (μ=10). This is a meaningful increase.

Historical Postscript on K.E.Tsiolkovsky

Of the three great space pioneers, R. Goddard, H. Oberth[2], K. E. Tsiolkovsky (alphabetical order), Konstantin Eduardovich Tsiolkovsky has the honor of being the first. He was born near Ryzan, Russia in 1857. At the age of ten he was taken seriously ill, resulting in almost total hearing loss. Isolated in this way, he turned inward and became very studious and contemplative. He took his education upon himself and was

2. During the preparation of this paper a notice appeared in the newspapers announcing the death of Hermann Oberth. He was the last survivor of the three, and the only one to actually witness the space-age.

almost totally self-taught. By profession he was a school teacher in the city of Kaluga. His scientific studies were his avocation. At at early age he was influenced by Jules Verne toward speculations on space exploration. In this he was spectacularly successful.

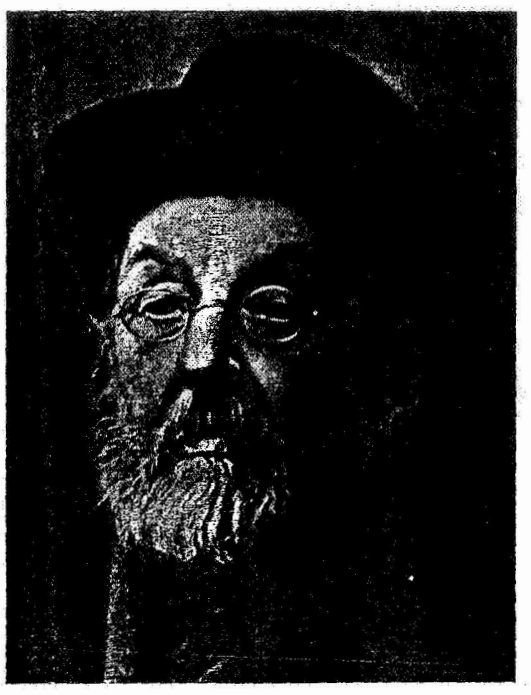
Equation 1 -- the Tsiolkovsky equation -- appears to have been derived by him as early as 1895, though it was not published until 1903. He was apparently the first to have derived it, though the Soviet literature gives credit to a certain Meshchersky (1859-1935) for having done work in the dynamics of variable-mass bodies. It does not appear, though, that he derived eq. 1. (A very early rocket theoretician was the English scientist Benjamin Robins (1707-1751). He belongs to the generation following Newton. Interestingly enough, he was also largely self-taught. He did much good work on ballistics, including the determination that the exhaust velocity of blackpowder was in the order of 7000 fps.)

When a man achieves a very great station his mistakes become noteworthy. Perhaps Tsiolkovsky's greatest error occurred when he initially discounted the step or stage principle, announced by Goddard. He soon realized his mistake and became an avowed supporter. It is interesting to note that his reason for discounting staging was a too literal interpretation of eq. (1)! A more venial mistake came about as follows. Tsiolkovky had the ingenious idea of circling the exhaust gasses, via looped pipes, prior to ejection. This, he reasoned, would provide desired stability through gyroscopic action. Realizing that a gyroscope can stabilize in only two axis, he considered two such orthogonal loops. But of course two gyroscopes rigidly tied together act as a single (or no) gyro. He died in his native country, greatly honored and revered, in 1935.

References:

1. K. E. Tsiolkovsky Selected Works, compiled by V. N. Sokolsky, MIR Publishers (Moscow), English Translation, 1968.

2. Sloop, John L., Liquid Hydrogen as a Propulsion Fuel, 1945-1959, NASA History Series, NASA Publication SP-4404, 1978. This monograph cannot be praised too highly: besides giving an excellent summary of the performance and history of liquid hydrogen, it provides excellent short biographics of the space pioneers.



K. E. TSIOLKOVSKY

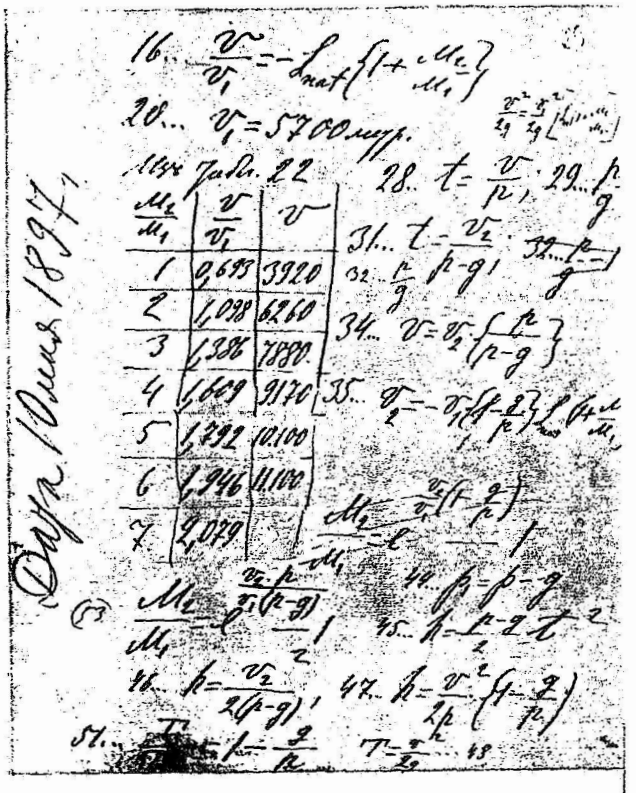


Figure 4. TSIOLKOVSKY'S FAMOUS FORMULA OF ROCKET MOTION (THE PHOTO DEPICTS A MANUSCRIPT PAGE DATED 1897)

CALCULATION OF DOUBLE-LUNAR SWINGBY TRAJECTORIES: II. NUMERICAL SOLUTIONS IN THE RESTRICTED PROBLEM OF THREE BODIES*

S. Stalos
Computer Sciences Corporation

ABSTRACT

The double-lunar swingby trajectory is a method for maintaining alignment of an Earth satellite's line of apsides with the Sun-Earth line. From a Keplerian point of view, successive close encounters with the Moon cause discrete, instantaneous changes in the satellite's eccentricity and semimajor axis. This paper identifies numerical solutions to the planar, restricted problem of three bodies as double-lunar swingby trajectories. The method of solution is described and the results compared to the Keplerian formulation.

*This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, Contract NAS 5-31500.

I. FORMULATION OF THE PROBLEM

Farquhar and Dunham (Reference 1) first recognized the utility of the double-lunar swingby trajectory. This trajectory uses encounters with the Moon to "rotate the line of apsides" at a rate equal to that of the Earth about the Sun (Figure 1). At the end of a complete double-lunar swingby cycle, the spacecraft, Sun, Earth, and Moon have the same relative position as at the beginning of the cycle. Hence, the trajectory is said to be periodic in the "lunar-rotating frame" as well as in the "solar-rotating frame." (The lunar-rotating frame is that coordinate system in which the Earth and Moon appear to be at rest with each other. That is, the coordinate axes rotate at the same rate as the rotation of the Moon about the Earth. Similarly, for the solar-rotating frame, the Sun and Earth are at rest with respect to each other. The coordinate axes then rotate at the rate of the mean solar motion.) Figure 1 depicts a typical double-lunar swingby trajectory in an "Earth-inertial" coordinate system; the Moon, Sun, and spacecraft motions are shown relative to axes centered on the Earth and at rest with respect to the stars. Figure 2 shows the same spacecraft trajectory viewed in the solar-rotating frame. Figure 3 gives the same spacecraft trajectory viewed in the lunar-rotating frame.

Dunham and Davis, in Reference 2, presented extensive tables describing double-lunar swingby trajectories in terms of Keplerian orbital elements. Reference 3 documented the method used to calculate these elements. In each case, the trajectory is viewed as two elliptic Earth orbits: an inner-segment loop orbit and an outer-segment loop orbit. Transfer is considered to be instantaneous between the two orbits and occurs at the lunar radius. Physically, of course, the change is not instantaneous. Other simplifications include the neglect of the Sun's gravity, the assumption of a circular lunar orbit, and the assumption that the Sun, Moon, and Earth are coplanar. How well, then, does the Keplerian formulation describe the double-lunar swingby trajectory? Further, how may we conveniently characterize such trajectories without resorting to Keplerian elements? Howell (Reference 4) addressed several of these issues and obtained numerical solutions, discussed in the following paragraphs.

One might integrate the equations of motion of the Earth, Moon, Sun, and spacecraft. A double-lunar swingby trajectory would then depend only on the spacecraft state vector (position and velocity) at a given time. This, in fact, is the procedure for a real mission. But how is that state vector determined? By satisfying the necessary equations given in Reference 3. How does one guess an initial state vector that might satisfy those equations? The answer is to build the state vector from the parameters given in Reference 2. Thus, the utility of the Keplerian formulation is seen: the model does not, however, depend on the Moon's true anomaly or argument of perigee or on the Earth's true anomaly in its orbit around the Sun. Little utility might be found in preparing massive tables that add these variables to our model.

We still may examine the Keplerian formulation's assumption of an instantaneous transfer between inner-segment and outer-segment orbits. We know that the transfer is not instantaneous. In fact, no Keplerian elements describe the possible behavior of a spacecraft under the influence of two gravitational forces. The simplest model that will describe such behavior is referred to as "the restricted problem of three bodies" (References 5

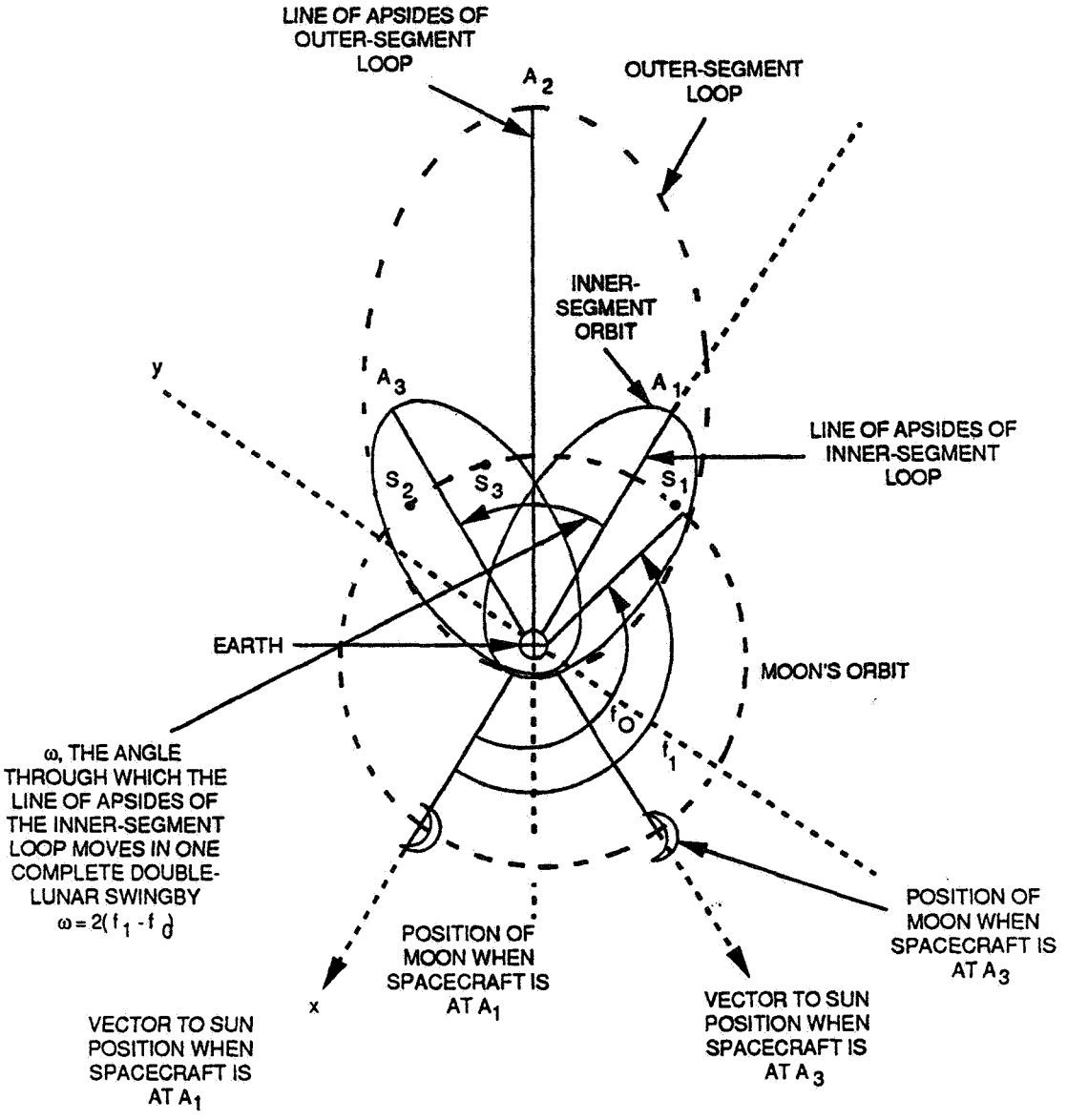


Figure 1. A Double-Lunar Swingby Trajectory

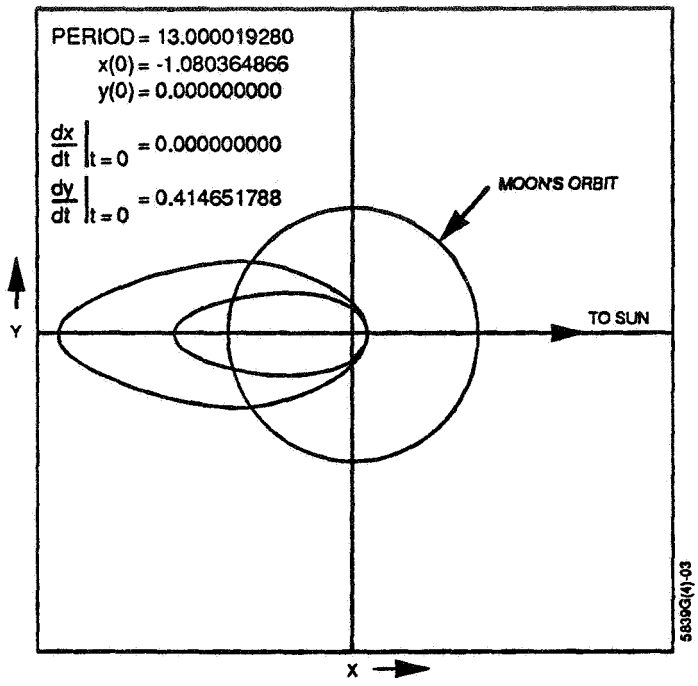


Figure 2. Double-Lunar Swingby Trajectory As Viewed in the Solar-Rotating Frame

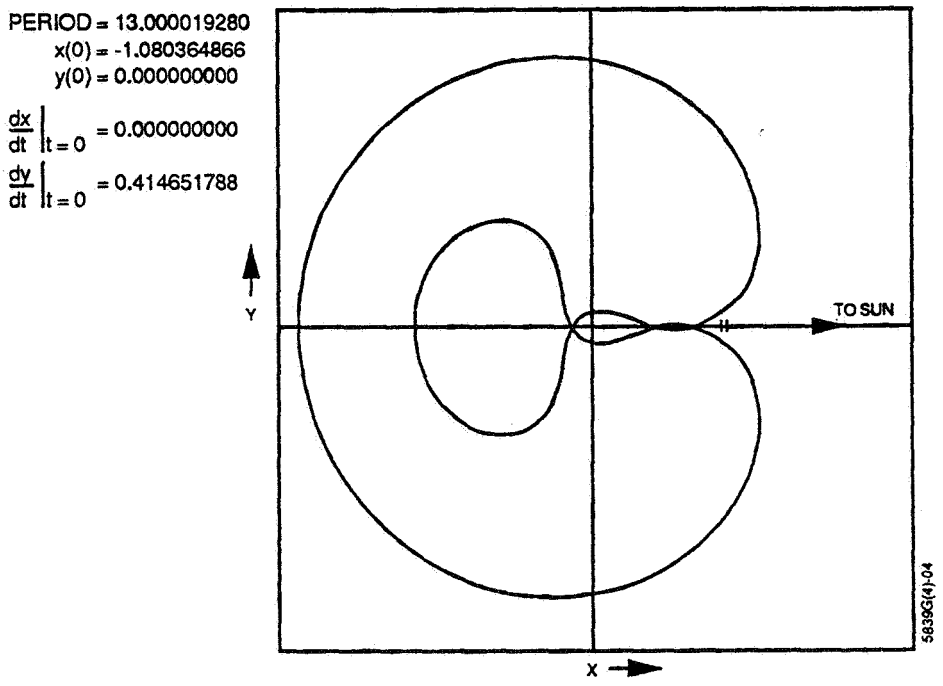


Figure 3. Double-Lunar Swingby Trajectory As Viewed in the Lunar-Rotating Frame

through 8). In this model, a body of zero mass (the spacecraft) is acted upon by the gravitational attraction of two bodies (here, the Earth and the Moon). The two bodies of finite mass are considered to be in a circular orbit about their center of mass, and the body of zero mass moves in the plane of that orbit. That is, we require that the initial out-of-plane position and velocity be zero. Hence, any trajectory depends only on five numbers: the spacecraft position and velocity and the relative masses of the finite bodies.

Our coordinate system is shown in Figure 4. Since the x and y axes rotate in the direction of the Moon's orbit about the Earth, both the Earth and the Moon appear motionless in this coordinate system: the lunar-rotating frame. The coordinate system's origin is at the center of mass of the Earth and the Moon. The equations of motion for the spacecraft, in this coordinate system, are

$$\frac{d^2x}{dt^2} = x + 2\frac{dy}{dt} - \frac{(1-\mu)(x+\mu)}{r_E^3} - \mu\frac{(x-1+\mu)}{r_M^3}$$

$$\frac{d^2y}{dt^2} = y - 2\frac{dx}{dt} - \frac{(1-\mu)y}{r_E^3} - \frac{\mu y}{r_M^3} \qquad (1\text{-}1)$$

where

r_E = spacecraft-Earth distance = $[(x+\mu)^2 + y^2]^{1/2}$

r_M = spacecraft-Moon distance = $[(x-1+\mu)^2 + y^2]^{1/2}$

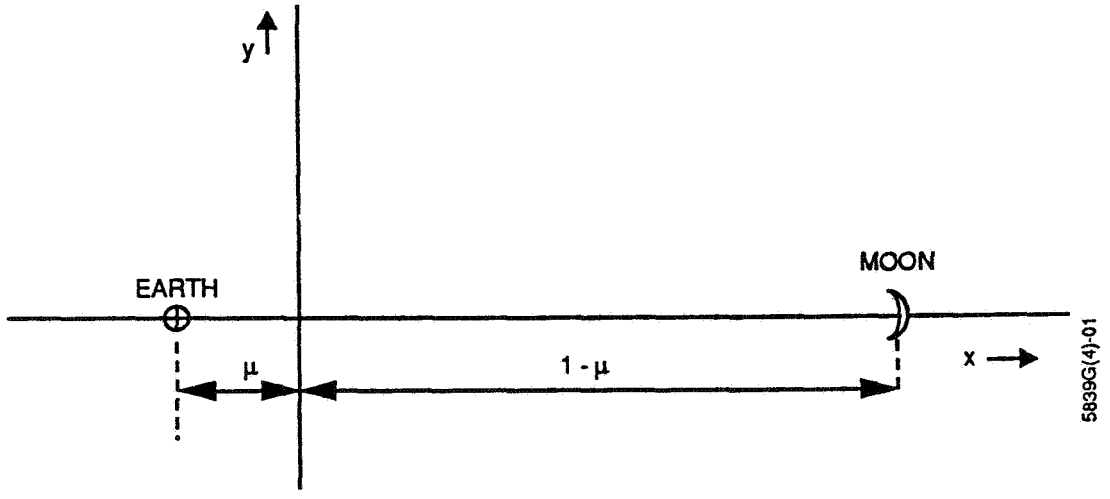


Figure 4. Lunar-Rotating Frame Used for Three-Body Equations

These equations are derived in pages 277-280 of Reference 6. Here, the mass unit is chosen such that the sum of the Earth's and Moon's masses is 1. The length unit is chosen such that the distance between the Earth and Moon is 1. The time unit is chosen such that the mean angular motion of the Moon about the Earth is 1. The Moon's mass is denoted by μ, and the Earth's mass is then $(1 - \mu)$. The (x, y) coordinates of the Earth are $(-\mu, 0)$. The coordinates for the Moon are $(1 - \mu, 0)$.

2. DEFINITION OF PERIODIC SOLUTIONS AND METHOD OF SOLUTION

Let the period of a double-lunar swingby be T_{DLS}. Then, by the definition of periodicity, the spacecraft coordinates and velocity at time t are equal to those at time $t + T_{DLS}$. For simplicity, we choose our time $t = 0$ to occur at apogee of the outer-segment loop. From Figure 3, we note that the symmetry of the orbit requires

$$y(0) = y(T_{DLS}/2) = y(T_{DLS}) = 0$$

$$\left.\frac{dx}{dt}\right|_{t=0} = \left.\frac{dx}{dt}\right|_{t=T_{DLS}/2} = \left.\frac{dx}{dt}\right|_{t=T_{DLS}} = 0 \qquad (2\text{-}1)$$

Suppose we integrate the equations of motion from $t = 0$ to $t = T_{DLS}/2$. Since y and the x velocity are known at $t = 0$, the values of y and of the x velocity at $t = T_{DLS}/2$ are a function of only two variables: the values of x and of the y velocity at $t = 0$. Thus, we must find the solution of a set of two equations (the values of y and of the x velocity integrated from $t = 0$ to $t = T_{DLS}/2$) in two unknowns (the initial x and the initial y velocity), or

$$y(T_{DLS}/2) = \int_0^{T_{DLS}/2} \frac{dy}{dt}\,dt \equiv f_1\left(x(0), \left.\frac{dy}{dt}\right|_{t=0}\right) = 0$$

$$\left.\frac{dx}{dt}\right|_{t=T_{DLS}/2} = \int_0^{T_{DLS}/2} \frac{d^2x}{dt^2}\,dt \equiv f_2\left(x(0), \left.\frac{dy}{dt}\right|_{t=0}\right) = 0 \qquad (2\text{-}2)$$

Any method used for the iterative solution of nonlinear equations may be applied to Equation (2-2). Both the Newton-Raphson method used in Reference 3 and the multivariate Halley method (Reference 9) were used to solve Equation (2-2). Note that Equation (2-1) is the same for $t = T_{DLS}/2$ or for $t = T_{DLS}$, so we may also have approached the problem by integrating from $t = 0$ to $t = T_{DLS}$. Another approach, suitable

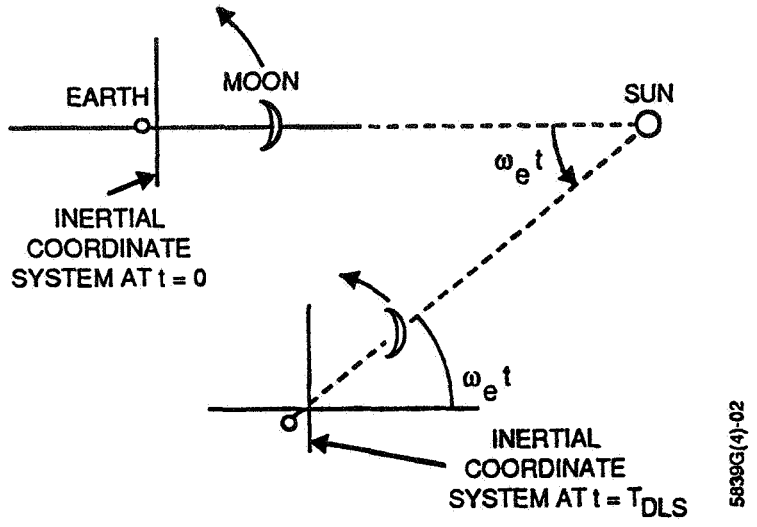


Figure 5. Determination of the Period of a Double-Lunar Swingby, T_{DLS}

even for asymmetric periodic solutions, is to form the norm of all four variables evaluated at t = 0 and at t = T_{DLS}

$$F\left(x(0),\ y(0),\ \left.\frac{dx}{dt}\right|_{t=0},\ \left.\frac{dy}{dt}\right|_{t=0}\right) \equiv \left\{[x(0) - x(T_{DLS})]^2 + [y(0) - y(T_{DLS})]^2 + \left(\left.\frac{dx}{dt}\right|_{t=0} - \left.\frac{dx}{dt}\right|_{t=T_{DLS}}\right)^2 + \left(\left.\frac{dy}{dt}\right|_{t=0} - \left.\frac{dy}{dt}\right|_{t=T_{DLS}}\right)^2\right\} \tag{2-3}$$

One then seeks to minimize this norm by any nonlinear optimization algorithm. Both a discrete Newton and the Davidon-Fletcher-Powell algorithms were implemented (References 9 and 10).

We now determine T_{DLS}. Consider Figure 5, which shows the Earth-inertial coordinate system at two different times. The Earth, moving with constant angular velocity ω_e around the Sun, has traveled the angle $\omega_e t$. The Moon, moving with constant angular

velocity ω_m around the Earth, returns to its original position with respect to the Sun and the Earth when it moves through the angle

$$\omega_m t = 2\pi n + \omega_e t \tag{2-4}$$

in the time

$$t = \frac{2\pi n}{\omega_m - \omega_e} \tag{2-5}$$

For the case n = 1

$$T_{DLS} = \frac{2\pi}{\frac{2\pi}{27.32} - \frac{2\pi}{365.25}} = 29.5287 \text{ days} \tag{2-6}$$

For the case n = 2

$$T_{DLS} = \frac{4\pi}{\frac{2\pi}{27.32} - \frac{2\pi}{365.25}} = 59.057 \text{ days} \tag{2-7}$$

Knowing T_{DLS}, we may now seek numerical solutions to Equation (2-2).

To find the three-body solution for the double-lunar swingby of Figure 1, we use initial values of x and the y velocity taken from Reference 1. The inner-segment loop has an apogee of 549889 kilometers (km) and a perigee of 37436 km. Using these values, the initial x and the y velocity values in the lunar-rotating frame are calculated to be

$$x(0) = -1.41836564, \left.\frac{dy}{dt}\right|_{t=0} = +1.1216628 \tag{2-8}$$

These initial conditions are then used to begin the iterative solution of Equation (2-2). The three-body solution is found to be in close agreement with the Keplerian solution. In the units of the lunar-rotating frame, the solution is

$$x(0) = -1.42050598244, \left.\frac{dy}{dt}\right|_{t=0} = +1.09755070684 \qquad (2\text{-}9)$$

(The period of the trajectory, in the units of the lunar-rotating frame, is 13.600878. This corresponds to 59.057 days. We consistently use a distance unit of 384399 km and μ = 0.0121505649405.)

This solution of the three-body problem was first obtained by Howell in Reference 4, for μ = 0.012.

Table 1 compares the three-body solution with the Keplerian solution. The values are so close, one may question the necessity for the three-body calculation. However, the three-body trajectory has no instantaneous jump in the velocity at the lunar orbit, as does the Keplerian solution. Further, of course, no simple parameterization akin to the Keplerian elements is possible for the general solution to three-body motion.

Table 1. Comparison of Three-Body Solution With Keplerian Solution of Reference 1

	Three-Body	Reference 1
Inner-segment loop apogee	541,370 km	549,889
Inner-segment loop perigee	42,350 km	37,438
Outer-segment loop apogee	892,710 km	898,915

Not every solution to the equations given in Reference 3 has a corresponding solution in the three-body problem. For example, the [1, 1, 1, 1] double-lunar trajectory shown in Figure 6, while certainly periodic in the lunar-rotating frame, is not quite periodic in the solar-rotating frame (Figure 7).

Also, the Keplerian formulation is not amenable to analysis of orbits such as that of Figure 8. Here the line of apsides is rotated *within* the lunar orbit. The apogee and perigee of this orbit are 248400 km and 120983 km, giving a period of approximately 9.14 days. The periodicity in the solar- and lunar-rotating frames is 29.5287 days for this orbit (n = 1).

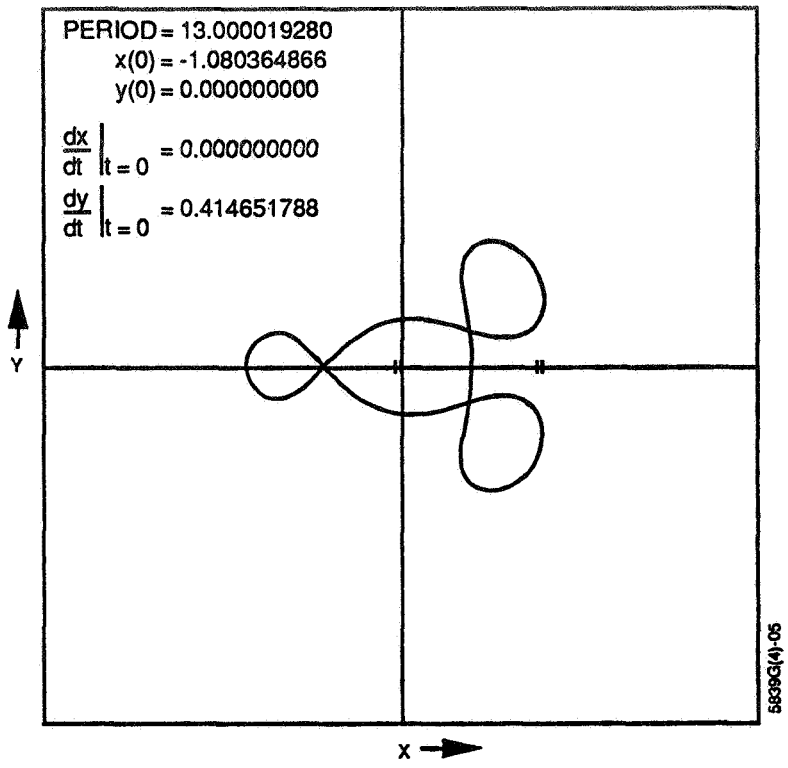


Figure 6. A Double-Lunar Swingby Trajectory Periodic Only in the Lunar-Rotating Frame

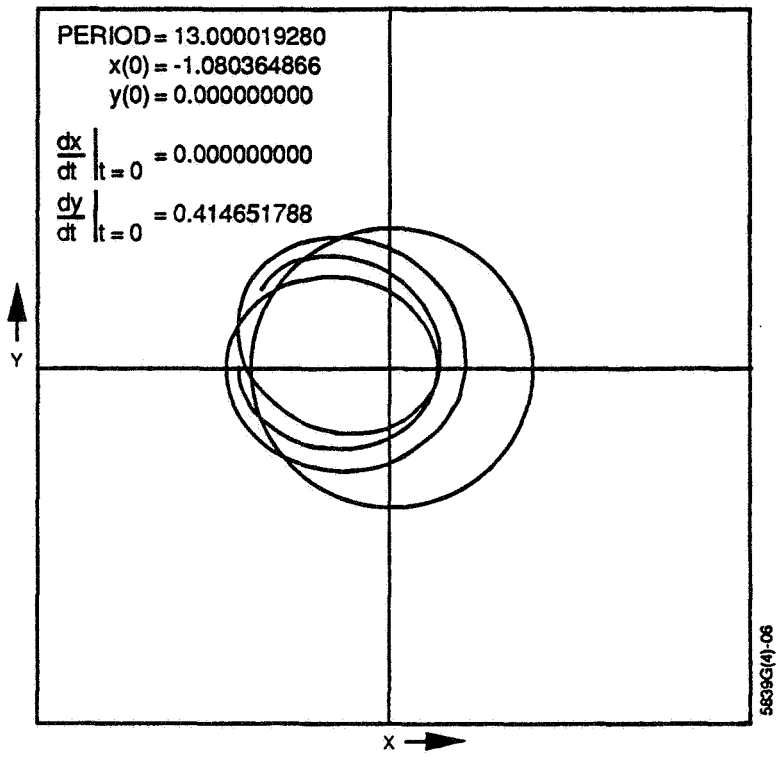


Figure 7. The Trajectory of Figure 6, Viewed in the Solar-Rotating Frame

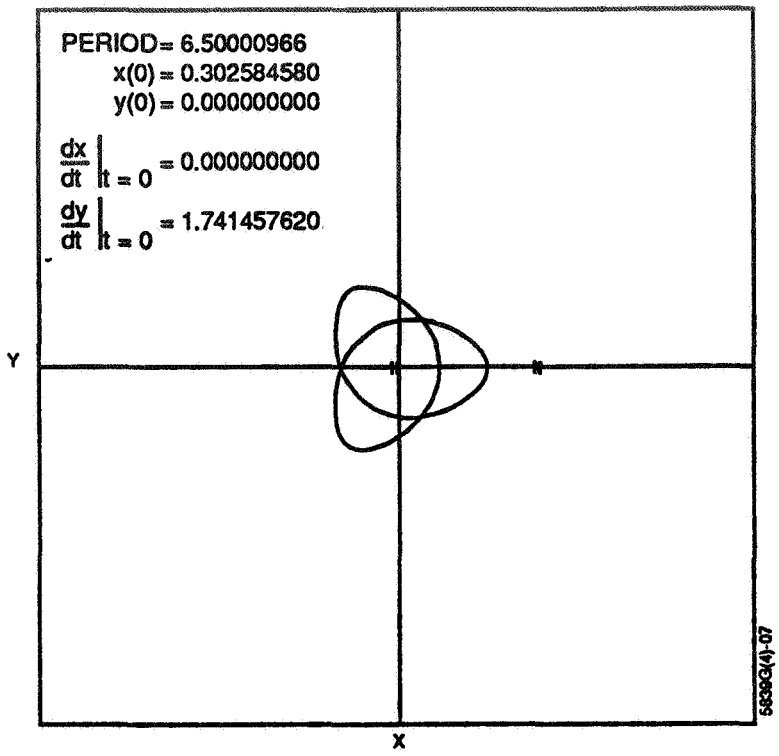


Figure 8. A Cislunar Periodic Solution of the Three-Body Problem

REFERENCES

1. R. W. Farquhar and D. W. Dunham, "A New Trajectory Concept for Exploring the Earth's Geomagnetic Tail," *Journal of Guidance and Control*, vol. 4, 1981, pp. 192-196

2. D. W. Dunham and S. A. Davis, "Catalog of Double-Lunar Swingby Orbits for Exploring the Earth's Geomagnetic Tail," Computer Sciences Corporation, CSC/TM-80/6322, October 1980

3. S. Stalos, "Calculation of Double-Lunar Swingby Trajectories: I. Keplerian Formulation," Flight Mechanics/Estimation Theory Symposium, May 23-24, 1989, Goddard Space Flight Center, Greenbelt, Maryland

4. K. C. Howell, "Halo Orbit Studies, Final Report," Purdue University, June 1985, pp. 56-60

5. V. Szebehely, *Theory of Orbits*, Academic Press, 1967

6. F. R. Moulton, *An Introduction to Celestial Mechanics*, Second Revised Edition, Dover Press, 1970

7. J. M. A. Danby, *Fundamentals of Celestial Mechanics*, Second Edition, Willman-Bell, 1988

8. E. T. Whittaker, *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Fourth Edition, Cambridge University Press, 1988

9. A. M. Cuyt, "Computational Implementation of the Multivariate Halley Method for Solving Nonlinear Systems of Equations," *ACM Transactions on Mathematical Software*, vol. 11, no. 1, March 1985, pp. 20-36

10. R. W. Daniels, *An Introduction to Numerical Methods and Optimization Techniques*, Elsevier/North Holland, 1978, pp. 226-232

11. W. E. Gill, R. T. Murray, and C. Wright, *Practical Optimization*, Academic Press, 1984

OPTIMAL AEROASSISTED ORBITAL TRANSFER WITH PLANE CHANGE USING COLLOCATION AND NONLINEAR PROGRAMMING+

By

Yun. Y. Shi*, R. L. Nelson, and D. H. Young*****

McDonnell Douglas Space Systems Company
5301 Bolsa Avenue
Huntington Beach, California 92647

ABSTRACT

The fuel optimal control problem arising in the non-planar orbital transfer employing aeroassisted tehnology is addressesd. The mission involves the transfer from high energy orbit (HEO) to low energy orbit (LEO) with orbital plane change. The basic strategy here is to employ a combination of propulsive maneuvers in space and aerodynamic maneuvers in the atmosphere. The basic sequence of events for the aeroassisted HEO to LEO transfer consists of three phases. In the first phase, the orbital transfer begins with a deorbit impulse at HEO which injects the vehicle into an elliptic transfer orbit with perigee inside the atmosphere. In the second phase, the vehicle is optimally controlled by lift and bank angle modulations to perform the deisred orbital plane change and to satisfy heating constraints. Because of the energy loss during the turn, an impulse is required to initiate the third phase to boost the vehicle back to the desired LEO orbital altitude. The third impulse is then used to circularize the orbit at LEO. The problem is solved by a direct optimization technique which uses piecewise polynomial representation for the state and control variables and collocation to satisfy the differential equations. This technique converts the optimal control problem into a nonlinear programming problem which is solved numerically. Solutions were obtained for cases with and without heat constraints and for cases of different orbital inclination changes. The method appears to be more powerful and robust than other optimization methods. In addition, the method can handle complex dynamical constraints.

---

* Staff Manager, ** Principal Scientist, *** Manager, Flight Mechanics/Advanced Systems Analysis

+ The senior author wishes to thank Professor P.E.Gill for helpful discussions on the structure of the nonlinear programming codes, i.e., NZSOL and NPSOL. He also wishes to thank Greg Badum for computing supports and graphics.

NOMENCLATURE

A : $S\rho_s/2$
C_D : drag coefficient
C_{DO} : zero-lift drag coefficient
C_L : lift coefficient
C_{LR} : lift coefficient for maximum lift-to-drag rati
D : drag force
g : gravitational acceleration
g_s : gravitational acceleration at surface level
H : altitude
J : performance index
K : induced drag factor
L : lift force
m : vehicle mass
R : distance from Earth center to vehicle center of gravity
R_a : radius of the atmospheric boundary
R_c : radius of the low Earth orbit (LEO)
R_d : radius of the high Earth orbit (HEO)
R_E : radius of Earth
S : aerodynamic reference area
t : time
V : velocity
T : Thrust
β : inverse atmospheric scale height
γ : flight path angle
ψ : heading angle
σ : bank angle
θ : down range angle or longitude
ϕ : cross range angle or latitude
μ : gravitational constant of Earth
ρ : density
ΔV : characteristic velocity

subscripts

c : subscript for circularization or reorbit
d : subscript for deorbit
s : subscript for surface level

1. INTRODUCTION

In order to have a viable and affordable space program, advanced technology must be exploited and new design concepts must be developed to reduce the size and cost of transportation elements for supporting new mission requirements. One of the new concepts that has evolved in recent years to advance the cost effectiveness of space transportation systems is the aerodynamically assisted orbit transfer. Such an orbital transfer vehicle is designed with an aerodynamic configuration which can utilize the planetary atmosphere for the purpose of energy management. Numerous studies have demonstrated that the use of the aerobraking can significantly reduce the

propulsive velocity requirements for certain class of orbit transfers. Excellent review papers were given by Warberg (Reference 1) and Mease (Reference 10).

In this paper, the fuel optimal control problem arising in a typical orbit transfer with plane change from HEO to LEO as discussed in most recent publications is addressed. In this case, as discussed in Reference 2, the aeroassisted orbit transfer vehicle (AOTV) maneuver involves three propulsive burns or impulses as sketched in Fig.1. In the first phase , the transfer begins with a deorbit impulse at HEO which injects the vehicle into an elliptic transfer orbit with the perigee inside the atmosphere. In the second phase, the vehicle is inside the atmosphere and is optimally controlled by the lift and bank angle modulations to perform the desired orbital plane change and to satisfy the heating rate and other physical constraints. Because of the the energy loss during the atmospheric maneuvers, an impulse is required to initiate the third phase to boost the vehicle back to the final orbital altitude. Finally, the third impulse is applied to circularize the orbit at LEO. In summary, there are three propulsive burns and an aeroassist plane change inside the atmosphere. Simulation results similar to those obtained in the draft paper of Reference 2 have been obtained here by using the Hermite polynomial and collocation technique to convert the optimal control problem into a nonlinear programming (NP) problem which is solved numerically using the optimization code, NZSOL (cf. Reference 12) provided by Gill, which is an improved version of NPSOL (cf. Reference 6), developed at Stanford. This solution method is different from the indirect method such as those discussed in Reference 2,4,7 and 8. The above simulation results have been extended to cases with heating constraints and cases for different orbit inclination changes. The details are presented and discussed in this paper. It is important that in the future these simulations be extended to include all other realistic flight constraints and to establish baseline optimum trajectory characteristics for GEO to Space station or shuttle, lunar and Mars missions.

2. DIRECT TRAJECTORY OPTIMIZATION WITH COLLOCATION AND HERMITE POLYNOMIALS

In the direct collocation with nonlinear programming approach, the trajectory is approximated by piecewise polynomials, which represent the state and control variables at a number of discrete time points, i.e., nodes. For a given state variable, the state trajectory over a given "segment" between two nodes is taken to be the unique Hermite cubic which goes through the end points of the segments with the appropriate derivatives that are dictated by the differential equations of motion at the endpoints. This is the "Hermite cubic" since it is determined by the states and their derivatives. A collocation is taken at the center of the segment where the derivative given by the Hermite cubic is compared to the derivative obtained from the evaluation of the equations of motion. The difference is termed the "defect" and is a measure of how well the equations of motion are satisfied over the segments. If all the defects are zero, then the differential equations are satisfied at the center collocation points as well as at the endpoints. Figure 2 shows the typical defects between node 1 and node 2.

Let the system of equations of motion be given as

$$X' = f(X,U) \tag{2-1a}$$

where X is the state vector and U is the control vector and (') denotes the differentiation with respect to the time. Let the time over a given segment be T. For the problem discussed here, one can show that

$$X = (x, y, z, x', y', z', m)$$
$$U = (C_L, \sigma) \tag{2–1b}$$

Then the Hermite interpolated x-component of the state vector X at the center point is

$$x_c = (1/2)\,(x_1 + x_r) + (T/8)\,[f(X_1,U_1) - f(X_r,U_r)] \tag{2-2}$$

where x_1 and x_r are respectively the x-component of the state vector X at the left and the right nodes. The derivative of the interpolating Hermite cubic at the center point is

$$x_c' = -3/(2T)\,(x_1 - x_r) - (1/4)\,[f(X_1,U_1) + f(X_r,U_r)] \tag{2-3}$$

The defect vector is then calculated as

$$d = f(X_c,U_c) - x_c' \tag{2-4}$$

If x_1, u_1, x_r, and u_r are chosen such that the elements of the defect vector, *d*, are sufficiently small, the "Hermite polynomials" become an accurate approximation to the solution of the differential equations of motion (by implicit integration). With the above approach, the differential equations are converted into nonlinear algebraic equations and the optimal control problem can then be solved using the nonlinear programming techques.

3. APPLICATION TO OPTIMAL AEROASSISTED ORBITAL TRANSFER WITH PLANE CHANGE

The aeroassited orbital transfer can be analyzed in three phases , i.e., deorbit, aeroassist (or atmospheric flight), boost and reorbit (or circularization). In each of the phases, a particular set of equations of motion apply.

3.1 Deorbit

Initially, the spacecraft is moving with a circular velocity $V_d = \sqrt{\mu / R_d}$ in a circular orbit of radius R_d, well outside the Earth's atmosphere. Deorbit is accomplished at point D by means of an impulse ΔV_d, to transfer the vehicle from a circular orbit to an elliptic orbit with perigee low enough for the trajectory to intersect the dense part of the atmosphere. Since the elliptic velocity at *D* is less than the circular velocity at *D*, the impulse ΔV_d is executed so as to oppose the circular Velocity V_d. In other words, at point *D*, the velocity required to put the vehicle into elliptic orbit is less than the velocity required to maintain it in circular orbit. The deorbit impulse ΔV_d causes the vehicle to

enter the atmosphere at radius R_a with a velocity V_e and flight path angle γ_e. It is known that the optimal energy loss maneuver from the circular orbit is simply the Hohmann transfer and the impulse is parallel and opposite to the instantaneous velocity vector.

After applying the deorbit impulse and before entering the atmosphere at R_a, the deorbit trajectory is a coasting arc and known integrals of the equations of motion can be used to relate the state vectors at R_a, the entry into atmosphere to the state vectors right after the deorbit impulse at R_d. Using the principle of conservation of energy and angular momentum at the deorbit point D and the atmospheric entry point E, we get

$$V_e^2/2-\mu/R_a=(V_d-\Delta V_d)^2/2-\mu/R_d \tag{3-1}$$

$$R_a V_e \cos(-\gamma_e)=R_d(V_d-\Delta V_d) \tag{3-2}$$

from which we can solve for ΔV_d to get

$$\Delta V_d=\sqrt{\mu/R_d}-\sqrt{2\mu(1/R_a-1/R_d)/\left[(R_d/R_a)^2/\cos^2\gamma_e-1\right]} \tag{3-3}$$

It is easily seen that the minimum deorbit impulse ΔV_{dm} obtained for $\gamma_e = 0$, corresponds to an ideal transfer with the space vehicle grazing the atmospheric boundary. To ensure proper atmospheric entry, the deorbit impulse ΔV_d must be higher than the following minimum deorbit impulse ΔV_{dm}

$$\Delta V_{dm}=\sqrt{\mu/R_d}-\sqrt{2\mu(1/R_a-1/R_d)/\left[(R_d/R_a)^2-1\right]} \tag{3-4}$$

Physically, the second term of the above equation corresponds to the apogee velocity of an ellpitic transfer orbit with perigee radius R_a and apogee radius R_d. This elliptic transfer orbit is tangent to the atmosphere boundary at perigee. It will be shown later that the nonlinear constraint equations (3-15) at the atmospheric entry point can also be derived from equations (3-1 and 2).

3.2 Aeroassist

During the atmospheric flight, the vehicle is optimally controlled by the lift and bank angle modulations to achieve the necessary velocity reduction (due to the atmospheric drag) and the plane change. In the present formulation, only the aeroassisted atmospheric flight need be solved by using the collocation and nonlinear programming techniques discussed earlier in this paper. The solutions in the other phases are provided by the known integral relations of the equations of motion because these arcs are coasting arcs.

Consider a vehicle with the point mass m, moving about a rotating spherical planet. The atmosphere surrounding the planet is assumed to be at rest, and the central gravitational field obeys the usual inverse square law. The equations of motion for the vehicle are given by (Figure 1),

$$\dot{r} = V \sin\gamma \tag{3-5a}$$

$$\dot{\theta} = \frac{V \cos\gamma \cos\psi}{r \cos\phi} \tag{3-5b}$$

$$\dot{\phi} = \frac{V \cos\gamma \sin\psi}{r} \tag{3-5c}$$

$$\dot{v} = \frac{(\eta T \cos\varepsilon - D)}{m} - \frac{\mu \sin\gamma}{r^2} + \omega^2 r \cos\phi (\sin\gamma \cos\phi - \cos\gamma \sin\psi \sin\phi) \tag{3-5d}$$

$$\dot{\gamma} = \frac{(\eta T \sin\varepsilon + L)\cos\sigma}{mV} - \frac{\mu \cos\gamma}{Vr^2} + \frac{V \cos\gamma}{r} + 2\omega \cos\psi \cos\phi$$

$$+ \frac{\omega^2 r \cos\phi}{V} (\cos\gamma \cos\phi + \sin\gamma \sin\psi \sin\phi) \tag{3-5e}$$

$$\dot{\psi} = \frac{(\eta T \sin\varepsilon + L)\sin\sigma}{mV \cos\gamma} - \frac{V \cos\gamma \cos\psi \tan\phi}{r} + 2\omega(\tan\gamma \sin\psi \cos\phi - \sin\phi)$$

$$+ \frac{\omega^2 r \cos\psi \sin\phi \cos\phi}{V \cos\gamma} \tag{3-5f}$$

$$\dot{m} = -f(r, V, \eta) \tag{3-5g}$$

where for a given vehicle, the drag D and the lift L are

$$D = \frac{S}{2m} \rho V^2 C_D \tag{3-5h}$$

$$L = \frac{S}{2m} \rho V^2 C_L \tag{3-5i}$$

and the drag and lift coefficients obey the drag-polar relation

$$C_D = C_{DO} + K C_L^2 \tag{3-5j}$$

Also, for an exponential atmosphere, one has

$$\rho = \rho_s \exp(-H\beta) \quad \text{and} \quad H = R - R_E \tag{3-5k}$$

Simulation results obtained here were using the U.S. standard Atmosphere 1976.

For the problem considered here, one assumes that, inside the atmosphere, the vehicle is optimally controlled by the aerodynamic forces only. Thus it is assumed that the thrust T is absent and the point mass is constant in this region. Furthermore, no earth rotation was assumed. The later is equivalent to consider the motion with respect to an earth fixed inertial coordinate system (ECI). The plane change or the orbit inclination, *i*, is related to the cross range ϕ and the heading angle ψ as

$$\cos i = \cos\phi \cos\psi \qquad t_e \le t \le t_f \tag{3-6}$$

The orbit inclination changes throughout the atmospheric flight and must end up with the required value at exit. For small values of cross range angle ϕ, *i* is given by the heading angle ψ itself.

3.3 Boost and Reorbit

During the atmospheric flight, the vehicle undergoes the plane change using the lift and bank angle modulation . Because of the loss of energy during the atmospheric maneuver, a second impulse is required at the exit from the atmosphere to boost the vehicle back to the final orbital altitude at LEO.

The vehicle exits the atmosphere at point *F*, with a velocity V_f and the flight path angle γ_f. The additional impulse ΔV_b, required at the exit point *F* for boosting the vehicle into an elliptic transfer orbit with apogee radius R_c, and the reorbit (or circularization) impulse ΔV_c, required to insert the vehicle into a circular orbit, are obtained by using the principle of conservation of energy and angular momentum at the exit point *F* and the reorbit or circularization point *C*. Thus, we have

$$(V_f + \Delta V_b)^2 / 2 - \mu / R_a = (V_c - \Delta V_c)^2 / 2 - \mu / R_c \tag{3-7}$$

$$(V_f + \Delta V_b) R_a \cos\gamma_f = R_c (V_c - \Delta V_c) \tag{3-8}$$

Solving for ΔV_b and ΔV_c from the above equations (3-7) and (3-8) yields

$$\Delta V_b = \sqrt{2\mu(1/R_a - 1/R_c) / \left[1 - (R_a/R_c)^2 \cos^2\gamma_f\right]} - V_f \tag{3-9}$$

$$\Delta V_c = \sqrt{\mu / R_c} - \sqrt{2\mu(1/R_a - 1/R_c) / \left[(R_c/R_a)^2 / \cos^2\gamma_f - 1\right]} \tag{3-10}$$

It is interesting to note that the second term of equation (3-10) is maximum for $\gamma_f = 0$ and therefore the reorbit impulse ΔV_c ,is minimum for $\gamma_f = 0$. It will be shown later that boundary conditions and nonlinear constraint equations at the exit point F, can be derived in terms of the final orbit characteristics and the final state vectors at the exit as shown in (3-16,17,& 18).

3.4 Performance Index

It is known that the change in speed, ΔV, also called the characteristic velocity, is a convenient parameter to measure the fuel consumption. For minimum-fuel maneuver, the objective is then to minimize the total characteristic velocity. A convenient performance index is the sum of the characteristic velocities for deorbit, boost, and reorbit, as

$$J = \Delta V_d + \Delta V_b + \Delta V_c \tag{3-11}$$

Where, ΔV_d, ΔV_b, and ΔV_c are the deorbit, boost, and reorbit characteristic velocities respectively, and are related as

$$\Delta V_d = \sqrt{\mu / R_d} - (R_a / R_d) V_e \cos(-\gamma_e) \tag{3-12}$$

$$\Delta V_c = \sqrt{\mu / R_c} - (R_a / R_c)(V_f + \Delta V_b)\cos\gamma_f \tag{3-13}$$

Alternatively, ΔV_d, ΔV_b, and ΔV_c are also given by (3-3, 9, and 10) respectively. Note that for a given final circular orbit, the impulses ΔV_b and ΔV_c are completely determined by the state variables V_f and γ_f at the exit of the atmospheric portion of the trajectory. The velocity V_e and the flight path angle γ_e at the atmospheric entry point are dependent only on the magnitude of the deorbit impulse ΔV_d. It follows that the optimal control problem needs to consider only the trajectory segment within the atmosphere subject to the nonlinear constraints and boundary conditions at the atmospheric entry and exit points. In addition, other path constraints such as the peak heating rate have to be satisfied.

3.5 Boundary conditions and constraints

The boundary conditions and constraints for the optimal control problem can be summarized as follows:

- At the entry into atmosphere, the following initial constraints.must be satisfied.

$$R = R_a \;;\; \gamma_e \le 0 \quad , \qquad \phi_e = 0, \;\; \psi_e = 0 \tag{3-14}$$

$$\frac{V_e^2}{2}\left[1 - \left(\frac{R_a}{R_d}\right)^2 \cos^2(\gamma_e)\right] - \mu\left(\frac{1}{R_a} - \frac{1}{R_d}\right) = 0 \tag{3-15}$$

The first initial constraint is required to ensure the vehicle enters the atmosphere. In the present formulation, the initial velocity and the flight path angle are unknown and to be determined by the optimization processes subject to the second constraint.

- At the exit from atmosphere, the following constraints must be satisfied.

$$R = R_a \;;\; \gamma_f \ge 0 \tag{3-16}$$

$$\frac{(V_f + \Delta V_b)^2}{2}\left(1 - \frac{R_a^2}{R_c^2}\cos^2\gamma_f\right) - \mu\left(\frac{1}{R_a} - \frac{1}{R_c}\right) = 0 \tag{3-17}$$

$$\cos i_f - \cos\phi_f \cos\psi_f = 0 \tag{3-18}$$

Equation (3-16) is required to ensure the vehicle exit the atmophere. The second costraint can be used to compute ΔV_b, and if ΔV_b,is assumed to be zero as in the case of aerobraking without orbit plane change, the above constraint must be imposed to determine the correct V_f and γ_f. The third constraint is required to perform the desired orbital plane change.

- In addition, there are other path constraints ,i.e., constraints must be satisfied along the trajectory such as
  - a) Stagnation Point Heating Rate Constraints
  - b) Altitude Constraints
  - c) Bounds on the Control Variables
  - d) Others

4. STRUCTURE AND SOLUTION OF THE NONLINEAR PROGRAMMING PROBLEM

The direct collocation and Hermite polynomial procedures described above convert optimal control problems into corresponding nonlinear programming problems. Ordinary differential equations are converted into corresponding nonlinear algebraic equations (or nonlinear "defects" constraint equations). These problems can then be solved using nonlinear programming codes.

The variables for the nonlinear programming problem are the collected state vectors and control vectors at the nodes and the time duration of phases. These quantities are assembled into the NLP state vectors

$$X^T = \left[X_1^T, U_1^T, \cdots\cdots X_n^T, U_n^T, t_1, t_2 \cdots\cdots t_k\right] \tag{4-1}$$

where n is the number of nodes and k is the number of phases on the trajectory. The defects and other physical and mathematical constraints are collected into the NLP constraint vector C

$$C^T = \left[d_1^T, d_2^T, \cdots\cdots d_n^T, w_1^T, w_2^T, w_3^T, \cdots\cdots w_j^T\right] \tag{4-2}$$

where d_i is the defect vector and w is a vector of additional problem constraints.

The nonlinear programming code used here is the NZSOL (Reference 12). The NZSOL is an improved version of the NPSOL (Reference 6) , developed by the Stanford Optimization Laboratory and designed to minimize a smooth nonlinear function subject to a set of constraints which may include simple bounds on the

variables, linear constraints, and smooth nonlinear constraints The problem is assumed to be stated in the following form:

NP

$$\begin{aligned} &\text{minimize} \quad F(x) \\ &x \in R^n \end{aligned}$$

$$\text{subject to} \quad \ell \le \begin{Bmatrix} x \\ A_L x \\ c(x) \end{Bmatrix} \le u, \tag{4-3}$$

where the objective function $F(z)$ is a nonlinear function, A_L is an m_L, x n constant matrix of general linear constraints, and $c(x)$ is an m_N - vector of nonlinear constraint functions. the objective function F and the constraint functions are assumed to be smooth, i.e., at least twice-continuously differentiable. (The method of NPSOL will usually solve NP if there are only isolated discontinuities away from the solution).

Note that upper and lower bounds are specified for all the variables and for all the constraints. This form allows full generality in specifying other types of constraints. In particular, the *i-th* constraint may be defined as an equality by setting $\ell_i = u_i$. If certain bounds are not present, the associated elements of ℓ or u can be set to special values that will be treated as $-\infty$ or $+\infty$.

Here we briefly summarize the main features of the method of NZSOL and NPSOL as discussed in Reference 6 because Reference 12 is not available to general public. At a solution of NP, some of the constraints will be active, i.e., satisfied exactly. An active simple bound constraint implies that the corresponding variable is fixed at its bound, and hence the variables are partitioned into fixed and free variables. Let C denote the m x n matrix of gradients of the active general linear and nonlinear constraints. The number of fixed variables will be denoted by n_{FX}, with n_{FR} ($n_{FR} = n - n_{FX}$) the number of free variables. The subscripts *"FX"* and *"FR"* on a vector or matrix will denote the vector or matrix composed of the components corresponding to fixed or free variables. The details are discussed in Reference 11.

A point x is a first-order Kuhn-Tucker point for NP if the following conditions hold:

(i) x is feasible;

(ii) there exist vectors ζ and λ (the Lagrange multiplier vectors for the bound and general constraints) such that

$$g = C^T \lambda + \zeta, \tag{4-4a}$$

where g is the gradient of F evaluated at x, and $\zeta_j = 0$ if the *j-th* variable is free.

(iii) The Lagrange multiplier corresponding to an inequality constraint active at its lower bound must be non-negative, and non-positive for an inequality constraint active at its upper bound.

Let Z denote a matrix whose columns form a basis for the set of vectors orthogonal to the rows of C_{FR}; i.e., $C_{FR}Z = 0$. An equivalent statement of the condition in terms of Z is

$$Z^T g_{FR} = 0 \tag{4-4b}$$

The vector $Z^T g_{FR}$ is termed the projected gradient of F at x. Certain additional conditions must be satisfied in order for a first-order Kuhn-Tucker point to be a solution of NP.

4.1 The Quadratic Programming Subproblem

Similar to NPSOL, the basic structure of NZSOL involves major and minor iterations. The major iterations generate a sequence of iterates (x_k) that converge to x^*, a first-order Kuhn-Tucker point of NP. At a typical major iteration, the new iterate $\bar{x}$ is defined by

$$\bar{x} = x + \alpha p, \tag{4-5a}$$

where x is the current iterate, the non-negative scalar α is the step length, and p is the search direction. Also associated with each major iteration are estimates of the Lagrange multipliers and a prediction of the active set.

The search direction p is the solution of a quadratic programming subproblem of the form

$$\begin{aligned} &\underset{p}{\text{minimize}} && g^T p + \frac{1}{2} p^T H p \\ &\text{subject to} && \bar{\ell} \le \begin{Bmatrix} p \\ A_L p \\ A_N p \end{Bmatrix} \le \bar{u}, \end{aligned} \tag{4-5b}$$

where g is the gradient of F at x, the matrix H is a positive-definite quasi-Newton approximation to the Hessian of the Lagrangian function and A_N is the Jacobian matrix of c evaluated at x.

The estimated Lagrange multipliers at each major iteration are the Lagrange multipliers from the subproblem (and similarly for the predicted active set) and provide information about the the sensitivity of these NLP problems.

Certain matrices associated with the QP subproblem are relevant in the major iterations. Let the subscripts *"FX"* and *"FR"* refer to the predicted fixed and free variables, and let C denote the $m \times n$ matrix of gradients of the general linear and nonlinear constraints in the predicted active set. First, we have available the TQ factorization (Reference 11) of C_{FR} :

$$C_{FR}\, Q_{FR} = (0 \quad T), \tag{4-6}$$

where T is a nonsingular $m \times m$ reverse-triangular matrix (i.e., $t_{ij} = 0$ if $i + j < m$), and the non-singular $n_{FR} \times n_{FR}$ matrix Q_{FR} is the product of orthogonal transformations. Second, we have the upper-triangular Cholesky factor R of the transformed and re-ordered Hessian matrix

$$R^T R = H_Q \equiv Q^T \bar{H} Q, \tag{4-7}$$

where $\bar{H}$ is the Hessian H with rows and columns permuted so that the free variables are first, and Q is the $n \times n$ matrix

$$Q = \begin{pmatrix} Q_{FR} & \\ & I_{FX} \end{pmatrix}, \tag{4-8}$$

with I_{FX} the identity matrix of order n_{FX}. If the columns of Q_{FR} are partitioned so that

$$Q_{FR} = (Z \quad Y), \tag{4-9}$$

the n_Z $(n_Z \equiv n_{FR} - m)$ columns of Z form a basis for the null space of C_{FR}. The matrix Z is used to compute the projected gradient $Z^T g_{FR}$ at the current iterate.

As discussed in Reference 6 and 11, a theoretical characteristic of SQP methods is that the predicted active set from the QP subproblem is identical to the correct active set in a neighborhood of x^*. In NPSOL, this feature is exploited by using the QP active set from the previous iteration as a prediction of the active set for the next QP subproblem, which leads in practice to optimality of the subproblems in only one iteration as the solution is approached. Separate treatment of bound and linear constraints in NPSOL also saves computation in factorizing C_{FR} and H_Q.

4.2 The merit function

Detailed discussions of the merit function are given in Reference 14. In NZSOL and NPSOL, once the search direction p has been computed, the major iteration proceeds by determining a steplength α that produces a "sufficient decrease" in the augmented Lagrangian merit function

$$L(x,\lambda,s) = F(x) - \sum_i \lambda_i (c_i(x) - s_i) + \frac{1}{2} \sum_i \rho_i (c_i(x) - s_i)^2, \tag{4-10}$$

where x, λ and s vary during the line search. The summation terms involve only the nonlinear constraints. The vector λ is an estimate of the Lagrange multipliers for the nonlinear constraints of NP. The non-negative slack variable $\{s_i\}$ allow nonlinear inequality constraints to be treated without introducing discontinuities. The solution of the QP subproblem (4-5) provides a vector triple that serves as a direction of search for the three sets of variables.

4.3 The quasi-Newton updated

Before going into the detailed discussions, it is important to point out that both the NZSOL and NPSOL start by initializing the Hessian matrix H = Identity matrix. Thus at the beginning, the search direction is in the steepest decent direction. No initial curvature information is computed and the curvature information is accumulated through the BFGS quasi-Newton updates. The matrix H in (4-5) is a positive-definite quasi-Newton approximation to the Hessian of the Lagrangian function. At the end of each major iteration, a new Hessian approximation H is defined as a rank-two modification of H. In NPSOL the BFGS quasi-Newton update is used:

$$\bar{H} = H - \frac{1}{s^T H s} H s s^T H + \frac{1}{y^T s} y y^T, \tag{4-11}$$

where $s = \bar{x} - x$ (the change in x).

Rather than modifying H itself, the Cholesky factor of the transformed Hessian H_Q (4-7) is updated, where Q is the matrix from (4-8) associated with the active set of the QP subproblem. The update (4-11) is equivalent to the following update to H_Q :

$$\bar{H}_Q = H_Q - \frac{1}{s_Q^T H_Q s_Q} H_Q s_Q s_Q^T H_Q + \frac{1}{y_Q^T s_Q} y_Q y_Q^T, \tag{4-12}$$

where $y_Q = Q^T y$ and $s_Q = Q^T s$. This update may be expressed as a rank-one update to R and is used to incorporate new curvature information obtained in the move from x to $\bar{x}$.

4.4 NZSOL, NPSOL 4.02, and NPSOL 2.1

For those who are interested in applying these NLP codes, there are two publised versions of NPSOL. The NPSOL 4.02 was developed after the NPSOL 2.1 and therefore more reliable and efficient algorithm were incorporated according to Gill (Reference 12). However, in updating the Cholesky factor, the NPSOL 4.02 updates the whole or complete R while the NPSOL 2.1 updates only the part associated with the Z-space or null space of R. For the problem formulated here ,usually several hundred varibles are involved and the NPSOL 2.1 converges in less computing time. The NZSOL (Reference 12) incorporates not only latest efficient and reliable algorithm but also updates only the part of R associated with the null space of R only. In addition to improve the algorithm of NPSOL, it also adopts the best parts of both NPSOL 2.1 and 4.02.

Finally, it may be interesting to point out that the matrices in the present formulation using collocation and Hermite polynomial are large and fairly sparse. For computational efficiency, it is important to incorporate NLP codes such as MINOS (Reference 13) to take advantage of the special characteristic of the collocation formulation discussed here.

5. NUMERICAL RESULTS AND DATA

The data used in the numerical experiments presented here (c.f. Reference 2 and 9) are summarized as follows:

$$C_{DO} = 0.1 \quad ; \quad K = 1.111 \quad ; \quad m/S = 300 \text{ kg/m}^2 \tag{5-1}$$

and the drag polar is

$$C_D = C_{DO} + K * C_L^2 \qquad (5\text{-}2)$$

and other data are

$$\rho_a = 1.225\ kg/m^3; \mu = 3.986x10^{14}\ m^3/sec^2$$
$$\beta = 1/6900\ m^{-1};\ R_E = 6378\ km$$
$$H_a = 120\ km; R_d = 12996\ km;\ R_c = 6558\ km \qquad (5\text{-}3)$$

Using the above mentioned data, simulations were carried out. The optimal solution for the reference case (shown in figures as Case 1) has the following entry and exit status.

Entry status: $H_e = 120$ km; $V_e = 9034.74$ m/sec
$\gamma_e = -4.36$ degrees; $\phi_e = 0$; $\psi_e = 0$ (5-4)

Exit status: $H_f = 120$ km; $V_f = 7028.95$ m/sec
$\gamma_f = 0.0$ deg; $\phi_f = -6.69$ deg
$\psi_f = 18.891$ deg; total flight time = 478 sec (5-5)

Characteristic velocities:

Deorbit characteristic velocity	$\Delta V_d = 1031.59$ m/sec
Boost characteristic velocity	$\Delta V_b = 821.49$ m/sec
Reorbit characteristic velocity	$\Delta V_c = 17.98$ m/sec
Total characteristic velocity	$\Delta V = 1871.07$ m/sec

(5-6)

Time histories of altitude, velocity, flight path angles, heading angles, dynamical pressure, atmospheric density, orbit inclination and heating rate are shown in Figure 3-10. Figures 11-13 show lift coefficient, bank angle, and lift to drag ratio as a function time for several simulation runs for the reference case(.i.e., Case 1). These simulation runs show that at high altitude the control may be different for different simulation runs depending upon initial guesses. This is really not a surprise because at high altitude, the aerodynamic forces or the controls are ineffective. In fact, the problem has a weak optimium with respect to the controls at high altitude.

Without going into the details, the characteristic velocities for the cases with orbital inclination changes for 15, 20 and 25 degrees are summarized in Table 2.

The heating rate Q_r , along the atmospheric trajectory, is computed for the stagnation point of a sphere of radius of one meter, according to the following relation (Reference 2 and 6)

$$Q_r = K_r \rho^{0.5} V^{3.08} \qquad (5\text{-}7)$$

where the ρ is the atmospheric density in kg/km^3, V is the velocity in km/sec and the K_r is the proportionality constant equal to 0.000308. The peak heating rate for the reference case is about 239 W/cm^2. Simulation results for cases with peak heat rate constrained to 203 W/cm^2 and 170 W/cm^2 were also obtained and shown in Figures 3-10 in comparison with the reference case without heat constraints as Case 2 and 3 respectively. These two cases reduces the peaking heating rate of the reference case by 15 and 30 percent respectively. Simulation results presented here provide the sensitivity of trajectory and associated physical variables as the heating constraints are imposed.

Similarly, the characteristic velocities for the cases with heating constraints are summarized in Table 2. The percent reductions are with respect to the peak heating rate of the reference case without heat constraints. As shown here, one needs less thermal protection materials and more fuel consumption to fly the heat constrained trajectories and therefore by taking into account the weight of thermal protection materials one may find an optimal design to minimize the total vehicle weight.

Another interesting observation from the data given in Table 1 and 2 is that the deorbit impulse is almost the same for all the cases simulated here. The total characteristic velocity for a given optimal trajectory is almost completely determined by the boost and the recirculation. In fact, the boost velocity contributes the most to the variation of the total characteristic velocity. Physically, it is obvious as the vehicle makes a larger turn it also loses more energy and therefore needs more velocity to boost it back to the final orbital altitude. Although the total characteristic velocity is insentive to the magnitude of deorbit impulse, the optimal trajectory is very sensitive to ΔV_d.

6. CONCLUDING REMARKS

An excellent survey of the subject was given in Reference 1. Walberg reviewed the problem of synergetic plane change for optimal orbital transfer. In a recent paper by Naidu (c.f. Reference 2), fuel optimal trajectories of aeroassisted orbital transfer with plane change were presented using the so-called multiple shooting method for the case without heat constraints and under the assumption that all the synergetic plane change was performed entirely in the atmosphere. A brief review of the progress made in this field was also given in Reference 2. In our paper, a similar problem for cases with and without stagnation point heating rate constraints was solved using the collocation and nonlinear programming technique. This method is especially suitable for parametrical studies because of its relative insensitivity to initial guesses. Once a solution for a reference case is obtained, solutions for other cases such as different orbital inclination change can be obtained easily.

Finally, the present problem can also be formulated under a more general assumption that not all the plane changes are entirely made in the atmosphere. It must be noted that the AOTV transfer can be made more efficient propulsively if the plane change is performed partly in the atmosphere and partly in space and the propulsive plane change in space is subdivided into components associated with various impulsive points. For the more general formulation discussed, the desired plane change may consist of more than one plane change, i.e.

$$\text{Total orbital plane change} = i_1 + i_2 + i_3 + i_a$$

where i_1, i_2, and i_3 are plane changes at the deorbit, reboost, and reorbit respectively and i_a is the aeroassisted plane change and all these plane changes will be determined by the optimization processes discussed here. Preliminary simulation results were obtained and are to be published in the near future.

It should be mentioned that the collocation and nonlinear programming technique discussed here was recently applied to another group of orbital transfer problem by Enright and Conway in Reference 3 and the relative insensitivity of this method to the initial guesses was also observed by them. Our basic simulation test bed is the OTIS codes (Reference 5) with an improved and updated nonlinear programming code (NZSOL). All physical models used were documented in Reference 5. Of course, necessary modifications and corrections have to be incorporated to simulate the aerobraking problems discussed here.

It may be worthwhile mentioning that the present problem was actually solved by guessing the initial state and control varibles at four selected points , i.e., the initial point, the final point and two other nodal points along the trajectory inside the atmosphere. The initial state and control variables at other nodes or grid points were simply obtained by linear interpolation. These initial guesses do not have to satisfy either the governing equations or the nonlinear constriants including the defects. Only roughly guesses are needed at these four points. Converged solutions were obtained with relative ease. However, it is important to point out proper scaling of the defects, constraints and variables are essential to get converged solutions. Although our results were compared with the draft paper of Reference 2, the solution presented by Naidu was not actually optimal because the final flight path angle $\gamma_f = -0.6217$ degrees is negative. For simulations discussed here, converged solutions were obtained by using as little as 20 nodes. However, in some cases, converged solutions were obtained using 60 nodes, In the later case, the problem has more than 660 independent variables and more than 400 nonlinear "defects" equations. For cases with heating rate constraints, the problem has more than 500 nonlinear constraint equations includinding the "defect" equations. As far as we know, this may be the first time converged soultions were obtained for so many independent variables and nonlinear constraint equations. This also illustrates how powerful the nonlinear programming code and the collocation and Hermite polynomial technique are.

Finally, it is important to mention again that aeroassisted orbital transfer introduces a strong coupling between the vehicle design and the trajectory design as indicated by the simulation data. A trajectory that minimizes fuel mass, without attention to heating , may require the vehicle to have a heavy thermal protection systems. As shown here, an optimal design for the total vehicle weight may be obtained as discussed earlier. However, if the aeroassisted transfer is to be prefered to all propulsive transfer, it must offer a reduction in fuel mass greater than the increase in thermal protection mass.

7. REFERENCES

1. Walberg, G. A., "A Survey of Aeroassisted Orbit Transfer", Journal of Spacecraft and Rockets, Vol. 22, Jan-Feb. 1985, pp 3-18.

2. Naidu, D. S., "Fuel-optimal Trajectories of Aeroassisted Orbital Transfer with Plan Change" AIAA Guidance, Navigation & Control Conference, Boston, MA., August 14-16, 1989.

3. Enright, P. J. and Conway, B. A., "Optimal Finite Thrust Spacecraft Trajectories Using Collocation and Nonlinear Programming", Paper AAS 89-350, AAS/AIAA Astrodynamics Specialist Conference, Stowe, Vermont, August 7-10, 1989.

4. Shi, Y. Y.," Matched Asymptotic Solution For Optimum Lift Controlled Atmospheric Entry ", AIAA JOURNAL, VOL. 9, 1971, PP. 2229-2238.

5. Hardgraves, C. R.and Paris, S. W., " OTIS-Optimal Trajectories by Implicit Integration ", Boeing Aerospace Co., Contract No. F33615-85-C-3009,1988.

6. Gill, P.E., Murray, W., Sanders, M.A., and Wright, M.H., " User's Guide For NPSOL (Version 4.0) : A Fortran Package For Nonlinear Programming ", System Optimization Laboratory, Department of Operation Research, Standford University, Stanford, Claifornia, 1986.

7. Shi, Y. Y., Pottsepp, L., and M. C. Eckstein. " Optimal lift Control of a Hypersonic Lifting Body During Atmospheric Entry ", AIAA Journal, Vol.7, No.12, December,1969..

8. Shi, Y. Y. and Eckstein, M. C., "An Exact Solution for Optimum Controlled Soft Lunar Landing", Astronautica Acta, Vol.16, pp. 9-18, Pergamon Press, New York, 1971.

9. Mease,K.D.and Vinh, N. X.," Minimum-Fuel Aeroassisted Coplanar Orbit Transfer Using Lift-Modulation', AIAA Journal of Guidance and Control, Vol.8,No.1, Jan-Feb. 1985.

10. Mease,K.D., "Optimization of Aeroassisted Orbital Transfer : Current Status", The Journal of the Astronautical Sciences, Special Issue On Hypervelocity Flight,Volume 36, Nos.1/2, January-June,1988.

11. Gill, P.E., and Wright, M.H., " Practical Optimization ", Academic press, London and New York, 1981.

12. Gill, P.E.," NZSOL: An Improved Version of NPSOL", Private Communications , Huntington Beach, California,1989-90.

13. Murtagh, B. A. and Saunders, M.A., " MINORS 5.0 User's Guide", Report SOL 83-20, Department of Operation Research, Standford Research, California, 1983.

14. Gill, P.E., Murray, W., Sanders, M.A., and Wright, M.H.," Some Theoretical Properties of An Augmented Lagragian Merit Function ", Technical Report SOL 86-6R, System Optimization Laboratory, Department of Operation Research, Standford University, Stanford, Claifornia, 1986.

TABLE 1

EFFECTS OF HEAT CONSTRANTS ON CHARACTERISTIC VELOCITIES

PEAK HEATING RATES	ΔV_1	ΔV_2	ΔV_3	TOTAL
(WATTS/cm)	(m/sec)	(m/sec)	(m/sec)	(m/sec)
UNCONSTRAINED $\dot{Q} = 239$	1031	821	18	1870
DR. NAIDU (BNDSCO) UNCONSTRAINED	1034	816	43	1893
HEAT CONSTRAINT $\dot{Q} = 203$	1028	855	18	1901
HEAT CONSTRAINT $\dot{Q} = 170$	1026	930	19	1974

TABLE 2

CHARACTERISTIC VELOCITIES FOR DIFFERENT ORBIT INCLINATION CHANGES

Δi	ΔV_1	ΔV_2	ΔV_3	TOTAL ΔV
(degree)	(m/sec)	(m/sec)	(m/sec)	(m/sec)
15	1029	312	21	1362
20	1031	821	18	1871
25	1035	1270	31	2336

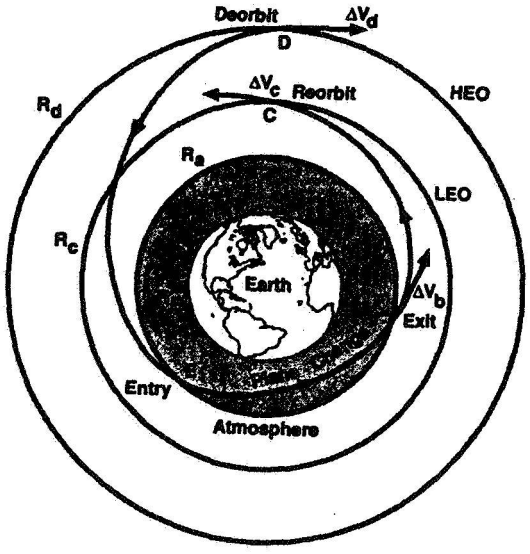


Fig. 1 Aeroassisted Orbital Plane Change

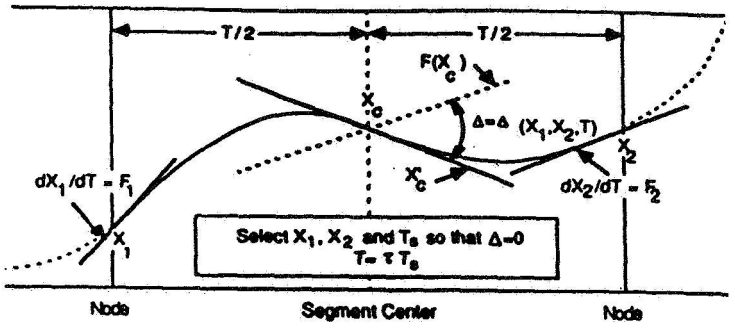


Fig. 2 Collocation and Hermite Approximation

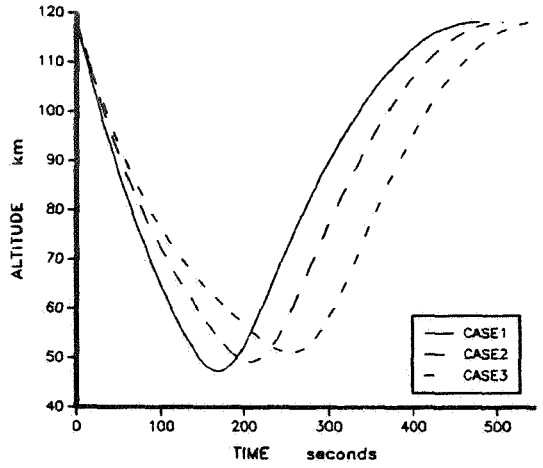


Fig.3 Time History of Altitude

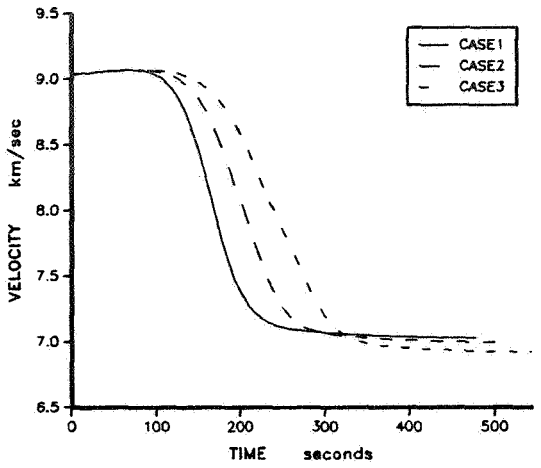


Fig.4 Time History of Velocity

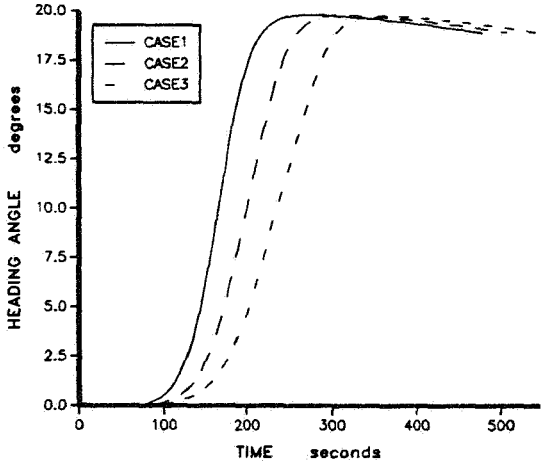


Fig.5 Time History of Heading Angle

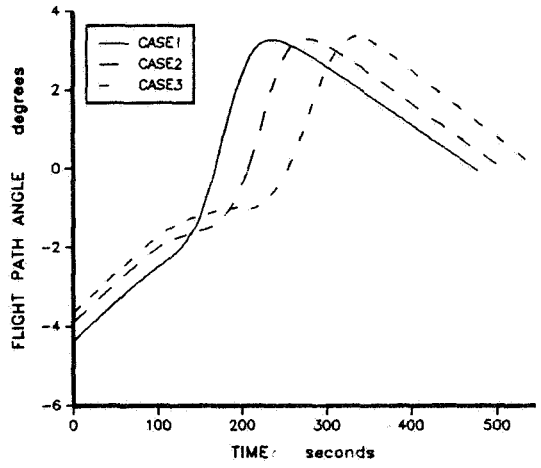


Fig.6 Time History of Flight Path Angle

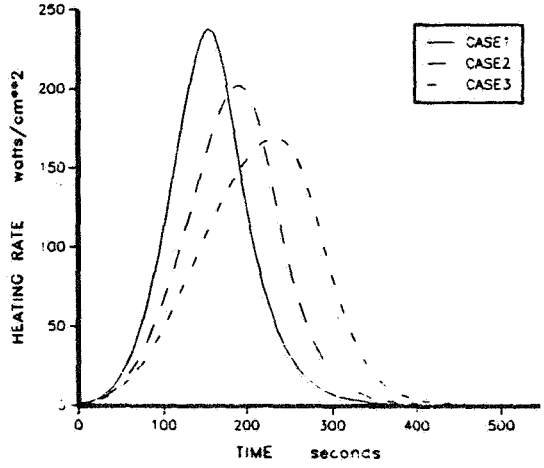


Fig.7 Time History of Heating Rate

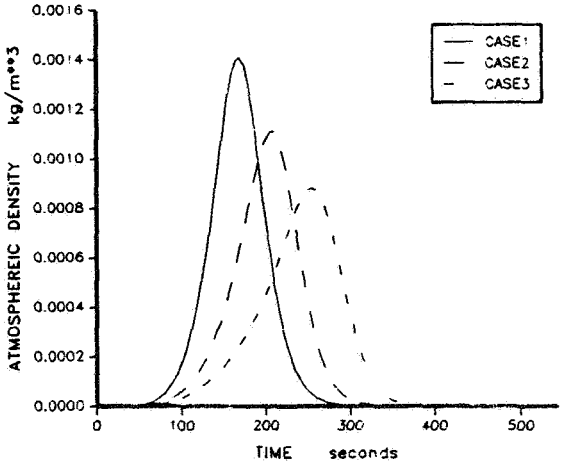


Fig.8 Time History of Atmospheric Density

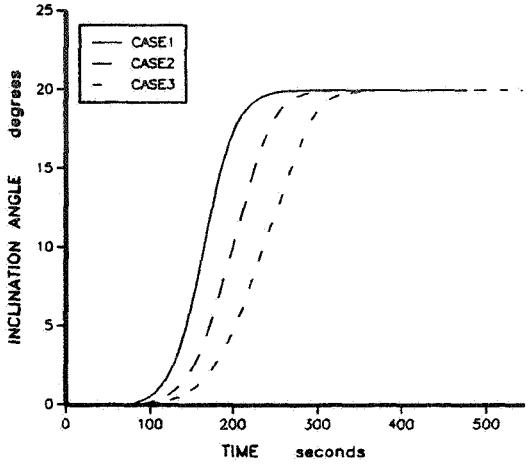


Fig.9 Time History of Inclination Angle

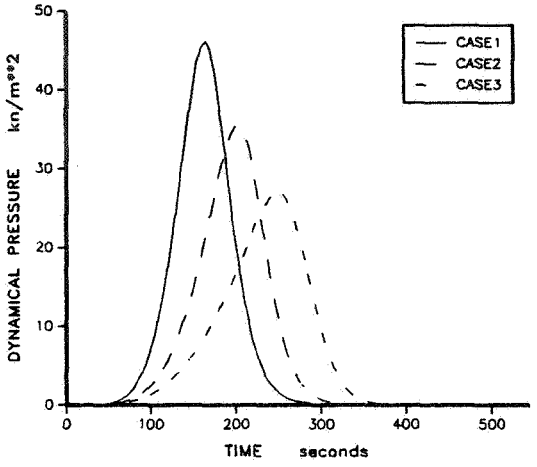


Fig.10 Time History of Dynamical Pressure

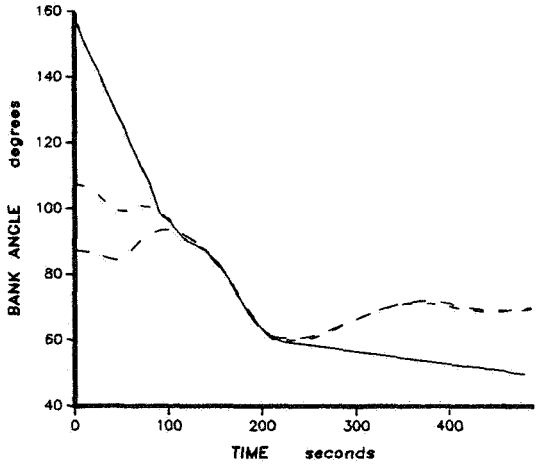


Fig.11 Bank Angle History for Reference Case

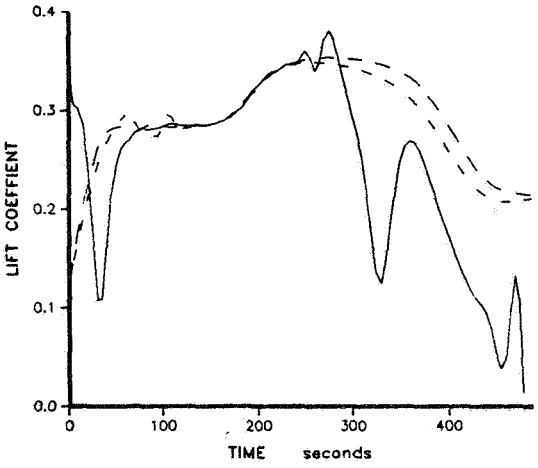


Fig.12 Lift Coefficient History for Reference Case

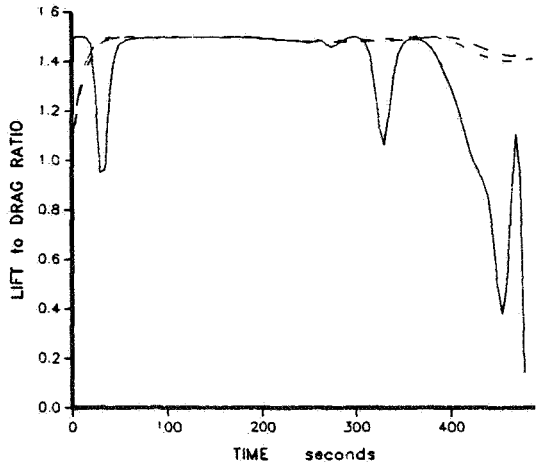


Fig.13 Lift to Drag Ratio History for Reference Case

AEROBRAKING FOR CAPTURE INTO MARTIAN ORBIT

by

William T. Suit
Mary C. Lee
Paul V. Tartbini
Barbara S. Tomlinson

ABSTRACT

The paper is a summary of several studies into problems associated with aerobraking a manned vehicle into a Martian capture orbit. The problems investigated are:

1) The establishment of entry flight path angle windows that allow aerocapture.

2) The determination of the sensitivity of the entry trajectory to initial flight path angle.

3) The determination of the effect on aerocapture of the assumed Martian atmosphere model.

4) The determination of the effect of random atmosphere disturbances on adaptive guidance systems that may be used for aerocapture.

As a result of investigating the above problem areas, entry windows were established for three different vehicle configurations. Sensitivities to changes in initial flight path angle were also obtained for these three configurations. One configuration was chosen to determine the effect of Martian atmospheric model changes and random variations of density within a specific atmospheric model. Of particular interest was the effect of random density variations on adaptive guidance techniques. The effect of entry velocity on the size of the entry window was also examined.

INTRODUCTION

Aerobraking has been identified as an enabling technology for manned Mars missions because of weight savings in propulsion fuel (ref.1). The accuracies required for Martian entry that guarantee aerocapture must be established.

Guidance techniques to accomplish aerocapture must also be determined. The guidance must be adaptable and robust enough to compensate for a wide range of atmospheric disturbances among which density variation has the greatest impact on guidance performance. In this paper only density variations will be considered.

For the types of vehicles envisioned for the manned Mars mission the guidance technique generally employed is bank angle lift modulation. Lift modulation is achieved by rotating the vehicle through various bank angles to change the magnitude of the lift component in the vertical plane of the entry trajectory. The vehicle will penetrate into the atmosphere to a given altitude and then the drag will slow the vehicle enough to ensure capture. After capture velocity is reached, the vehicle tends toward lift up attitude so that it can escape from the atmosphere of Mars and go into orbit. Lift modulation is used to adjust the level of penetration into the atmosphere and how quickly the vehicle exits the atmosphere so that operational constraints can be satisfied.

Since this is a preliminary study, a range of possible manned Mars vehicles was considered. Basically, all of the vehicles will accommodate a 6- to 8- person crew and be designed for missions of up to 2 years. Within this framework vehicles with a range of ballistic coefficients were studied to establish combinations of initial flight path angle and bank angle required for capture into Martian orbit. To establish these angles, entry flight path angle windows and the sensitivity to entry flight path angle for aeropass trajectory parameters were determined. Next,bank angle profiles required for aerocapture for different vehicles were investigated. Simulated guidance profiles using fixed bank angle sequences were used to determine the sensitivity to bank angle for various vehicles and for different Martian atmospheres.

Using the results of the studies with fixed bank commands two adaptive guidance techniques were developed. In this paper adaptive guidance is a procedure for continuously changing the bank commands to the vehicle control system. So that these guidance techniques could be tested in a realistic environment, a random Martian atmosphere was developed. By incorporating the random atmosphere into the program used to simulate Martian entries, the adaptive guidance techniques were tested under conditions of random density variation. Monte Carlo techniques were used to establish success boundaries for the various guidance techniques so that their adaptability to random density variations could be demonstrated.

This paper will discuss the entry windows for several potential manned Mars mission vehicles , the sensitivity to the entry flight path angle, the bank angle profiles required for Martian aerocapture and the "survivability" of adaptive guidance techniques in a randomly varying Martian atmosphere.

SYMBOLS

A	area, m^2
a	acceleration, m/sec^2
C_D	drag coefficient
C_L	lift coefficient

h_p	perigee altitude,km
Δh_p	change in perigee altitude, km
$h_{p,p}$	predicted perigee altitude, km
$h_{p,T}$	target perigee altitude, km
k	feedback gain
L/D	lift-to-drag ratio
M	mass, kg
S_c	current vehicle state
S_N	vehicle state from baseline trajectory
V_I	velocity, m/sec
ΔV_I	change in velocity, m/sec
ϕ	commanded bank angle, degrees
ϕ_N	bank angle from baseline trajectory, degrees
γ_I	initial flight path angle, degrees
$\Delta\gamma_I$	change in initial flight path angle, degrees

ABBREVIATIONS

ALTITO	altitude, m
BNKANG	bank angle, degrees
DENS	density, kg/m^3
ENERGY	energy per unit mass, m^2/sec^2
POST	program to optimize simulated trajectories
TIME	time, sec
VELI	velocity, m/sec
ASMG	acceleration, "g" units

APPROACH

The studies discussed in this paper were made using the Program to Optimize Simulated Trajectories (POST). This program can be used to determine initial parameters and control parameters throughout a trajectory to accomplish stated mission objectives. The simulations were started at 300,000 meters and at an entry angle selected by the user. Initially, a fixed number of bank angles were chosen to represent an entry guidance system. By varying the magnitudes of these bank angles the lift force is modulated to control the trajectory of the vehicle. Runs were considered successful if capture was achieved, the maximum acceleration was less than 5 "g"s, and the minimum altitude was

greater than 30 kilometers.

Studies were run for the vehicles described in Table I which contains the characteristics required to calculate the ballistic coefficients of the three vehicles used in this study. A large number of bank angles was used to ensure capture. To establish the entry window the maximum and minimum flight path angles for which capture was possible were determined. The difference in these flight path angles was the entry window. During the runs the sensitivity of the trajectory to flight path angle and bank angle was calculated by POST.

Once the trends in bank angle magnitude for various vehicles and entry conditions were determined, then an attempt was madeto establish the minimum number of bank commands that were required for aerocapture. One of these "fixed" bank profile trajectories became the nominal for the adaptive guidance.

Various Martian atmosphere models obtained from David Pitts et al. at Johnson Space Center were used to determine the effects of varying Martian atmospheres. When the random Martian atmosphere was developed, these same data were used as the nominal for the perturbed atmosphere. The random atmosphere subroutine allowed different levels of density variation.

The two adaptive guidance techniques to be discussed in this paper are presented in reference 3. Adaptive guidance 1 is a trajectory following technique and adaptive guidance 2 is a predictor-corrector technique In summary, the adaptive techniques adjust the bank angle to change the orientation of the lift vector and control the entry trajectory. Adaptive guidance 1 follows a nominal trajectory that gives an acceptable entry. The guidance used in this study compared the actual and nominal energies at the current velocity. The form of the bank angle command equation was $\phi = \phi_N + k * (E_C - E_N)$. This technique tried to correct to a nominal energy when the trajectory was perturbed by density variations.

Adaptive guidance 2 predicted the perigee altitude based on current conditions and adjusted the bank angle to try to attain a desired perigee altitude. The form of the bank angle command was $\phi = \phi_N + k * (h_T - h_p)$. Once perigee was reached the vehicle was rolled to a specified bank angle until the Martian capture velocity was attained and then the vehicle was rolled to full lift up for escape from the atmosphere. The adaptive guidance techniques implemented were not optimal but were used to determine the effect of a random atmosphere on representative guidance systems.

A program for establishing requirements for and evaluating the performance of guidance techniques was developed by combining the guidance subroutines, the random Martian atmosphere program, and the basic POST software. The results and discussion section, which follows, will discuss applications of this program.

RESULTS AND DISCUSSION

The initial investigations to determine Martian aerocapture characteristics were conducted using a hypothetical manned vehicle with a $M/C_DA = 620.5$ and $L/D = 1$ (ref. 4). Since the sensitivity to bank angle was unknown, the initial entries were made using a simulated guidance that would allow 16 bank changes during the aeropass. Using this simulated guidance, the maximum and minimum entry flight path angles for which capture was possible were determined. Typical time histories of these aerobraking trajectories are shown as figures 1 and 2. As can be seen from examination of the bank angle time histories, about nine of the possible bank angle changes were required for the maximum flight path angle entry and six for the minimum flight path angle entry. This gave an indication of the amount of maneuvering an entry might require.

The POST program calculates the sensitivity of entry trajectory parameters to entry flight path angle. Typical sensitivities to entry flight path angle and the maximum and minimum flight path angles possible for two proposed manned vehicles are shown in figure 3. The sensitivity is generally larger for the maximum flight path angle entries since they fly higher in the atmosphere, have smaller lift components and are, therefore, less able to correct for disturbances.

The entry windows $\Delta\gamma_i$ (difference between the maximum and minimum entry angles) are shown as Table II for two potential manned Mars entry vehicles. Table II gives the maximum and minimum entry flight path angles and $\Delta\gamma_i$ for several L/D ratios determined by assuming a fixed drag coefficient and changing the lift coefficient. The entry flight path angles for which aerocapture was possible showed almost no change with ballistic coefficient; however, the change in entry flight path angle with L/D was significant. As the lift that was available to be modulated to control the vehicle was reduced, the size of the entry window decreased.

The sensitivities to entry flight path angle for several vehicles with different L/D values and ballistic coefficients are given as Table III taken from reference 5. This table was generated by taking an entry flight path near the center of the entry window and varying this angle by .001 degrees from the chosen entry angle. The results of the runs with the modified angle were compared with a run made using the original entry angle and the Δ parameter to $\Delta\gamma_i$ ratios were obtained. These sensitivities and window sizes can be used to establish navigation and guidance accuracy requirements.

Thus far, only aerobrake shapes with large nose radii have been discussed. However, for completeness a more streamlined vehicle with significantly less drag and a much larger ballistic

coefficient was examined (Table I). The entry windows for this vehicle are given in Table IV. Since this vehicle penetrated deeper into the atmosphere the minimum entry flight path angle was constrained by altitude considerations and the entry windows were smaller for this vehicle.

Aerobraking entries were run with several other Martian atmospheric models. The aerobraking trajectories showed very little impact due to the change in atmospheric model. These results are shown in references 4 and 5 .

Results presented earlier indicated that the number of bank commands required for successful aerobraking could be greatly reduced. The vehicle referred to as blunt vehicle type 1 (Table I) was shown to require six or less bank commands when a L/D of .5 was assumed (figure 4). Since nominal entries will be flown near the center of the entry corridor to allow as much margin as possible before capture trajectory limits are encountered, runs were made with entry flight path angles near the middle of the entry windows. These show that less guidance activity was required for the nominal runs, but these runs are only for a deterministic atmosphere. Using a deterministic atmosphere and flying near the middle of the entry corridor, capture trajectories were generated that required only two commanded bankangles. One of these was chosen as the nominal trajectory of adaptive guidance 1.

So that the more realistic case of the effect of random disturbances on a guidance system could be tested, a random Martian atmosphere generator was implemented as a subroutine to POST. This combination enabled the adaptive guidance techniques to be tested in a realistic environment. When random density variations of up to 50 percent were allowed, both adaptive guidance routines gave acceptable capture trajectories. Typical entry trajectories with 50 percent random density variations as compared to a trajectory using a deterministic density profile are shown as figures 5 and 6. Adaptive guidance 2 also gave acceptable results for maximum density variations of up to 86 percent. This assessment, discussed in reference 5, was based on a limited number of runs.

Since adaptive guidance 2 seemed the most tolerant to large density variations, Monte Carlo runs were made to establish the success of the guidance for a large number of runs. One hundred runs were made at each of two density variation levels for adaptive guidance 2. The variations were from the deterministic density of figure 6 and the results are comparisons to various parameters from figure 6. When the first variation level wasused, the maximum density variation was such that 90 percent of the maximum densities fell within a plus/minus 60 percent band about the maximum deterministic density. In spite of large density variations, the final periods of the capture orbits fell within a plus/minus 20 percent band of the deterministic density orbital period 88 percent of the time. The acceleration was over 5"g"s only 6 percent of the time and at no time was the acceleration over 5.5 "g"s. All of the perigee altitudes were greater than 34 kilometers. All of the entries resulted in capture.

To see how far the maximum density could be varied and the guidance still be successful, an additional one hundred runs were made with the second level of maximum density variation. In this case 96 percent of the maximum densities fell inside a plus/minus 120 percent band about the deterministic maximum density. The final periods of the resulting orbits fell within plus/minus 20 percent of the deterministic density orbital period value 44 percent of the time and within plus/minus 30 percent 77 percent of the time.

In spite of the fact that all the perigee altitudes were greater than 34 kilometers because of the larger maximum densities, many of the maximum accelerations were large. The maximum acceleration was greater than 5 "g"s 40 percent of the time and greater than 5.5 "g"s 23 percent of the time. In all cases, capture was achieved.

Although 40 percent of the cases exceeded the 5 "g" limit imposed on the capture trajectory, the density variations in the order of 100 percent are probably extreme. The results using the 60 percent variations are probably more realistic. The fact that a very simple guidance performed well with large density variations implies that a more optimal guidance should be very successful.

All of the runs were made using an entry velocity of 6.7 kilometers/sec. This entry velocity is on the low end of possible entry velocities for manned Mars missions (ref. 1).Several additional runs were made using vehicle described as a blunt vehicle type 1 (Table I), and having a L/D of .5. Entry velocities of 7.5 kilometers/sec and 8.0 kilometers/sec were tested. The results are shown in Table V . The entry velocity had no effect on the trajectories that flew lowest in the atmosphere (minimum γ_i). However, the entry angles for which capture was possible for trajectories passing higher in the atmosphere (maximum γ_i) became much more negative as velocity increased. Because of the added energy the vehicle had to pass lower in the atmosphere to be captured, thereby, reducing the entry window.

CONCLUDING REMARKS

Potential manned Mars mission vehicles with large nose radii and L/D's of .5 were found to have entry flight path angle windows of over 1 degree. For these windows the sensitivity to change in initial flight path angle tended to be greater for trajectories that flew higher in the atmosphere. The size of the windows changed very little with ballistic coefficient, but were smaller as L/D decreased.

Vehicles with an L/D of .5 that flew near the middle of the entry corridor required very little guidance activity to obtain capture for the deterministic case. However, when flying near the top of the entry corridor or when random disturbances were encountered, more guidance response was required. To determine the

extent of the guidance response required to adapt to density disturbances, a random Martian density generator and two adaptive techniques were developed.

The guidance techniques were tested at several levels of maximum random density variation for a limited number of runs. Both adaptive guidance techniques gave acceptable results for random variations of up to 50 percent. Adaptive guidance 2 had better performance at higher levels of random density variation, so it was tested for 100 cases using Monte Carlo techniques. For maximum density variations of up to 60 percent, 96 percent of the runs resulted in acceptable entries.

Increasing the entry velocity reduced the entry flight path angle window for which capture was possible.

REFERENCES

1. Walberg,Gerald D.: A Review of Aerobraking for Mars Missions. IAF PAPER No. IAF-88-196, 39th Congress of the International Astronautical Federation, Bangalore, India, October 1988.

2. Braner,G.L.; Cornick,D.E.; and Stevenson,R.: Capalities and Applications of the Program to Optimize Simulated Trajectories (POST), NASA CR-2770, February 1977.

3. Tomlinson, Barbara S.; and Suit, William T.: Use of Random Martian Atmosphere to Evaluate Potential Entry Guidance Schemes, NASA TM 102606, January 1990.

4. Lee, Mary C.; and Suit, William T.: Preliminary Investigation of Parameter Sensivities for Atmospheric Entry and Aerobraking at Mars, NASA TM 101607, September 1989.

5. Tartabini, Paul V.; and Suit, William T.: Aerobraking Characteristics for Several Potential Manned Mars Entry Vehicles, TM 101669, November 1989.

Table I. Vehicle Characteristics

Blunt Vehicles		
Vehicle Type 1	$M = 226,378$ kg $A = 182.415$ m^2	$M/C_DA = 919.3$ kg/m^2 $C_D = 1.35$
Vehicle Type 2	$M = 226,378$ kg $A = 182.415$ m^2	$M/C_DA = 620.5$ kg/m^2 $C_D = 2$
Streamlined Vehicle		
Vehicle Type 1	$M = 136,116.2$ kg $A = 79$ m^2	$M/C_DA = 2970.7$ kg/m^2 $C_D = 0.58845$

Table II. Entry Angle Windows for Blunt Vehicles

L/D	M/C_DA (kg/m^2)	C_D	Max γ_i	Min γ_i	$\Delta\gamma_i$
0.3	620.5	2	−18.3193°	−19.1109°	0.7916°
0.3	919.3	1.35	−18.4461°	−19.2344°	0.7883°
0.5	620.5	2	−18.2415°	−19.5880°	1.3465°
0.5	919.3	1.35	−18.3289°	−19.7263°	1.3974°
0.75	919.3	1.35	−18.2432°	−20.0423°	1.7991°
1.0	620.5	2	−18.0646°	−19.6747°	1.6101°
1.0	919.3	1.35	−18.3492°	−20.2000°	1.8508°

Table III. Sensitivity of Velocity, Altitude, and Acceleration to γ_i at Perigee for Two Potential Manned Mars Vehicles.

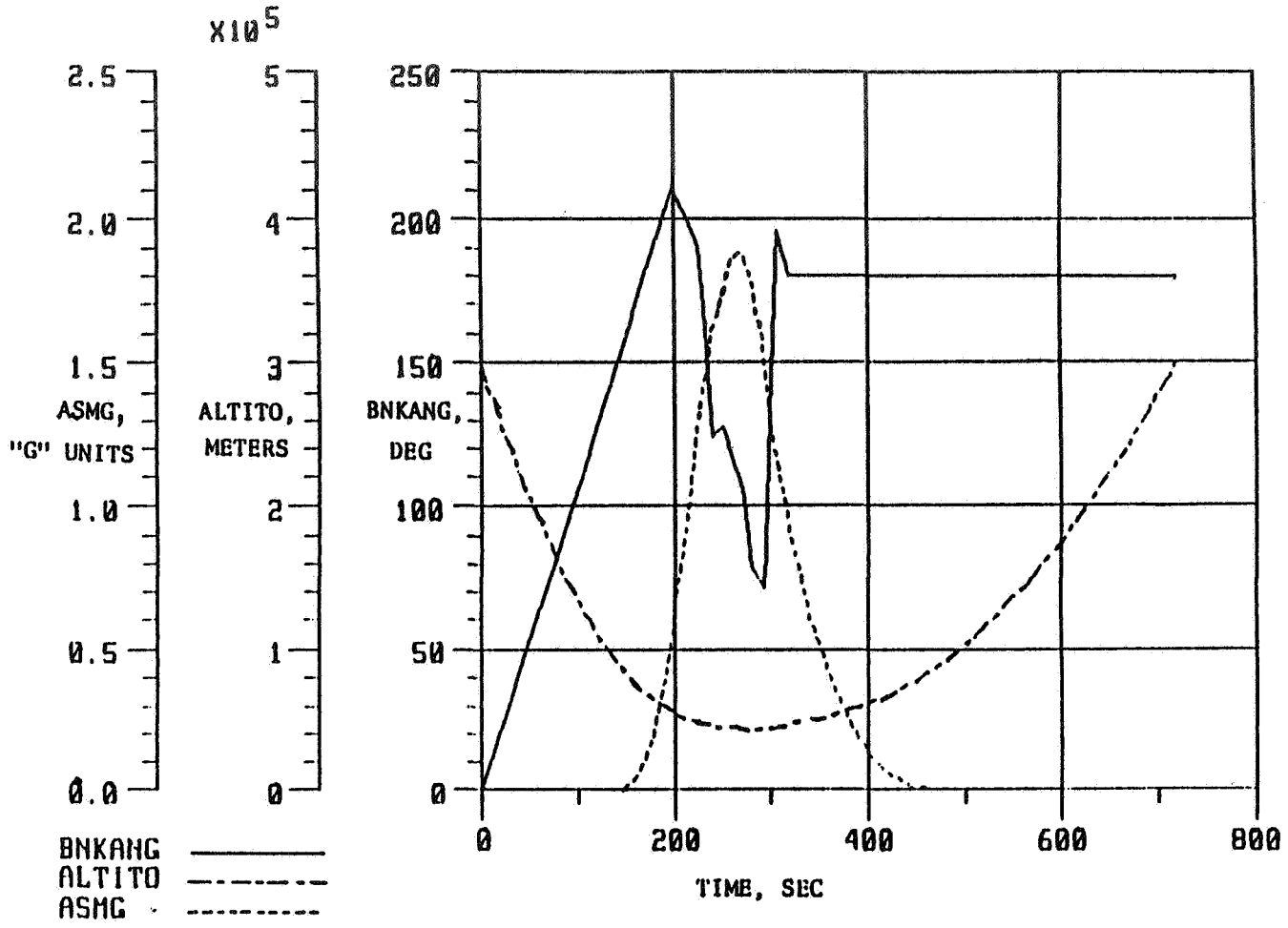
L/D	$M/C_D A$ (kg/m^2)	$\Delta\gamma_i$ (deg)	$\Delta V/\Delta\gamma_i$ (m/s/deg)	$\frac{\Delta h}{\Delta\gamma_i}$ (m/deg)	$\frac{\Delta a}{\Delta\gamma_i}$ (m/s^2/deg)
0.3	919.3	+0.001	2.0×10^3	2.44×10^4	45.5
0.3	919.3	−0.001	2.0×10^3	2.38×10^4	44.6
0.5	919.3	+0.001	1.5×10^3	1.30×10^4	31.0
0.5	919.3	−0.001	1.5×10^3	1.20×10^4	29.5
0.75	919.3	+0.001	1.1×10^3	9.80×10^3	32.2
0.75	919.3	−0.001	1.1×10^3	9.80×10^3	31.7
1.0	919.3	+0.001	1.2×10^3	1.26×10^4	36.0
1.0	919.3	−0.001	1.2×10^3	1.26×10^4	36.0
0.5	2970.7	+0.001	2.3×10^3	2.84×10^4	43.3
0.5	2970.7	−0.001	2.3×10^3	2.86×10^4	86.5
0.75	2970.7	+0.001	3.8×10^3	7.33×10^4	48.8
0.75	2970.7	−0.001	3.8×10^3	7.63×10^4	48.8
1.0	2970.7	+0.001	3.1×10^3	6.09×10^4	52.8
1.0	2970.7	−0.001	3.1×10^3	1.23×10^5	53.0

Table IV. Entry Angle Windows for Streamlined Vehicle

L/D	h_p(km)	a_{max}(m/sec^2)	γ_i(deg)	Comments	$\Delta\gamma_i$(deg)
.5	33.0	12.5	-18.637	Lift down	--
.5	30.0	18.3	-20.508	32 km limit	.663
.75	35.6	10.8	-18.613	Lift down	--
.75	32.1	17.0	-19.197	32 km limit	.684
1.0	37.9	10.0	-18.424	Lift down	--
1.0	32.1	19.8	-19.314	32 km limit	.890

Table V. Entry Angle Windows for Several Entry Velocities

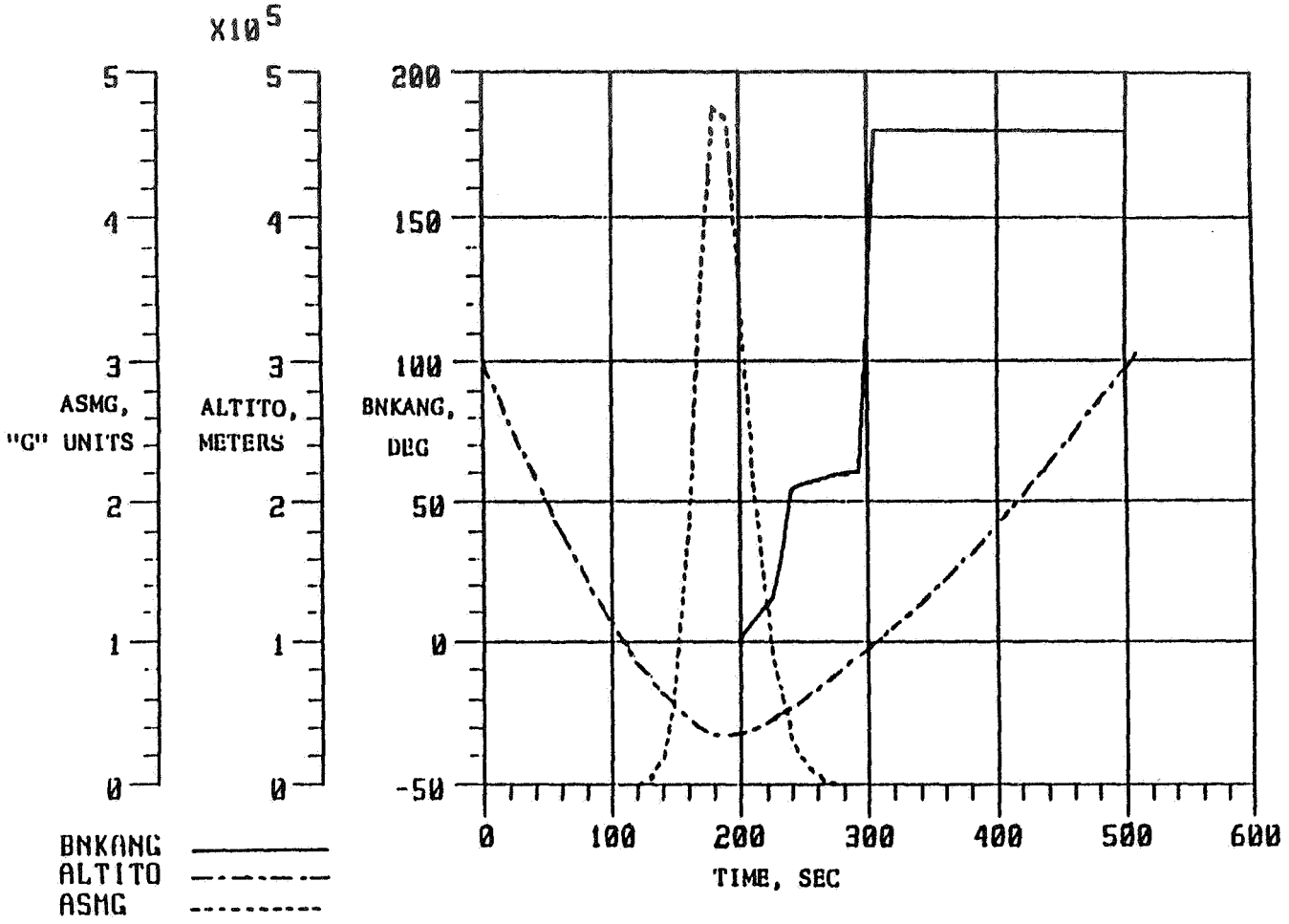
V_I	Max γ_i	Min γ_i	$\Delta\gamma_i$
6.7 km/sec	-18.3 deg	-19.8 deg	1.5 deg
7.5 km/sec	-18.9 deg	-19.8 deg	0.9 deg
8.0 km/sec	-19.2 deg	-19.8 deg	0.6 deg



Vehicle Type 2

M/C_DA = 620.5 L/D = .5

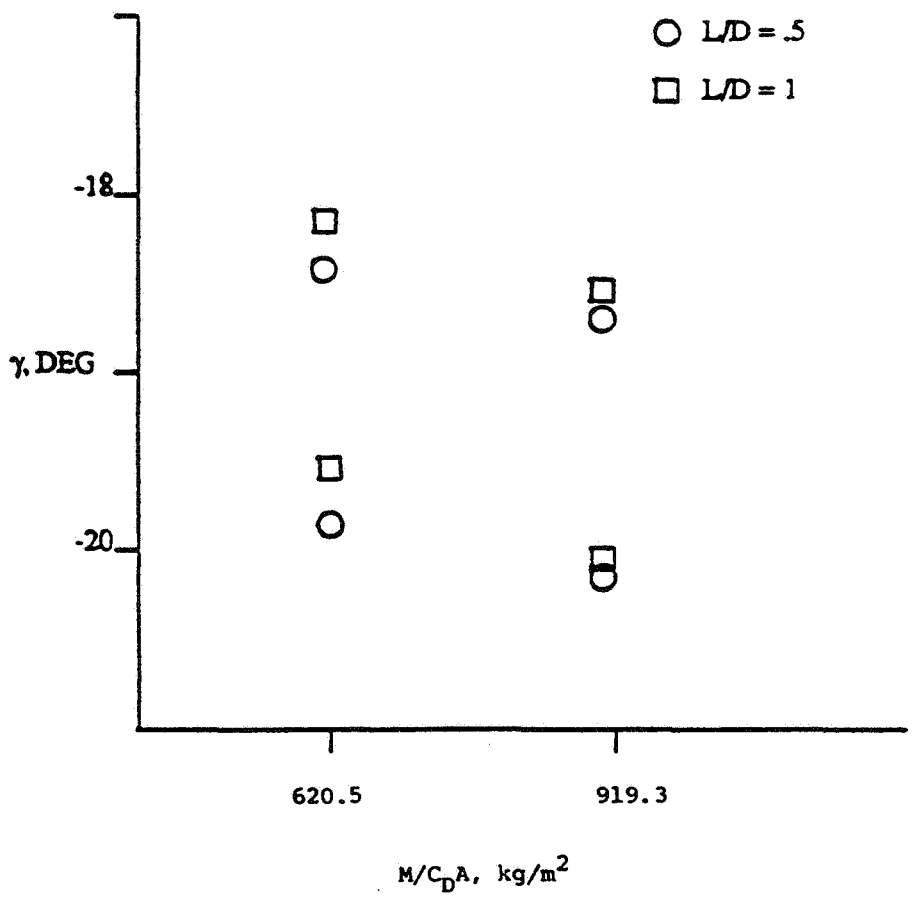
Figure 1. Time Histories of Bank Angle, Altitude, and Acceleration for a Maximum Flight Path Angle Martian Entry.



Vehicle Type 2

M/C_DA = 620.5 L/D = .5

Figure 2. Time Histories of Bank Angle, Altitude, and Acceleration for a Minimum Flight Path Angle Martian Entry.



TYPICAL SENSITIVITIES

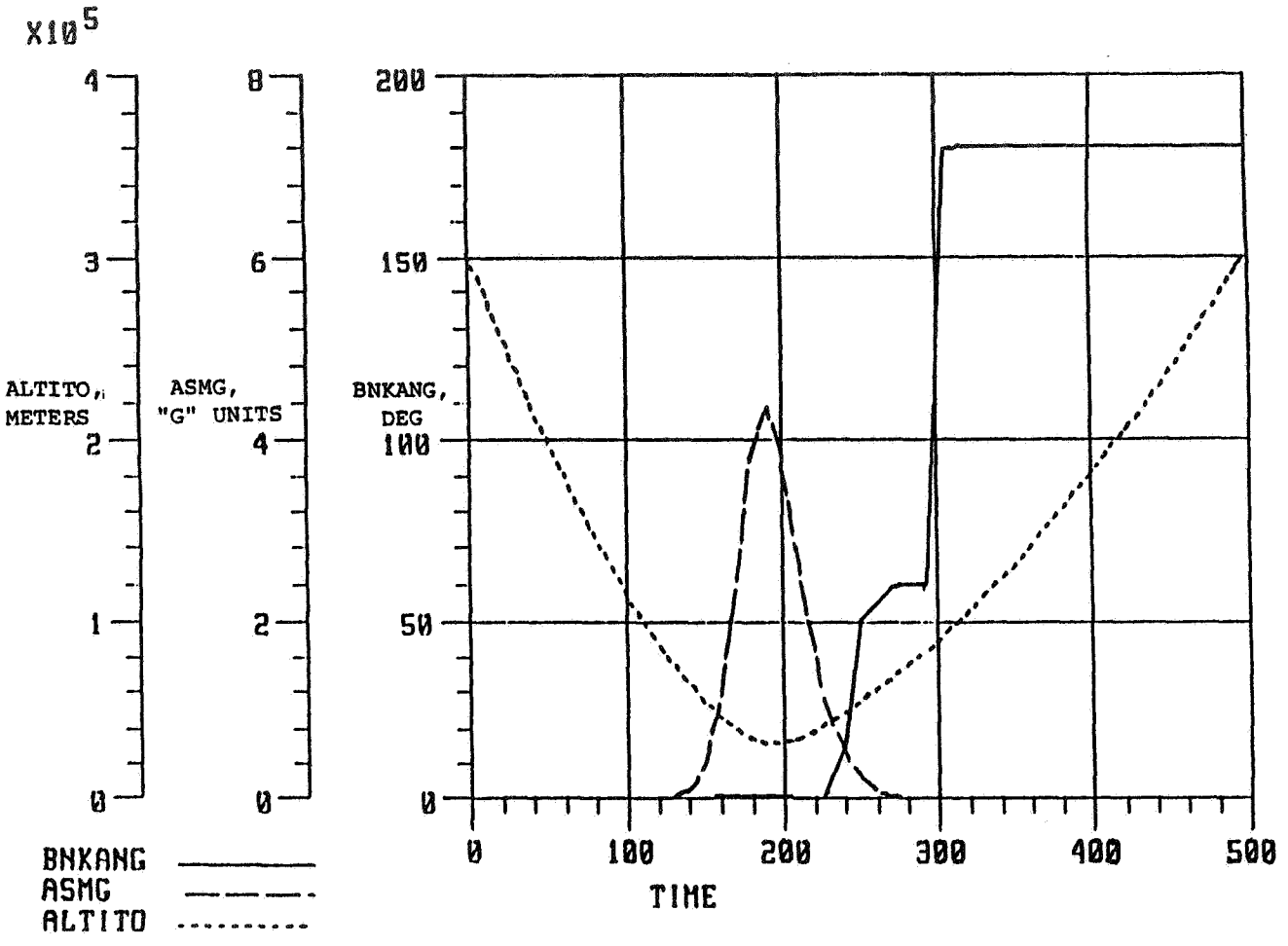
MAXIMUM FLIGHT PATH ANGLES

$\left\lvert \frac{\Delta V}{\Delta \gamma_I} \right\rvert$	L/D = 1	L/D = .5
	1.37E4	1.0E4

MINIMUM FLIGHT PATH ANGLE

$\left\lvert \frac{\Delta V}{\Delta \gamma_I} \right\rvert$	L/D = 1	L/D = .5
	3.85E3	.93E3

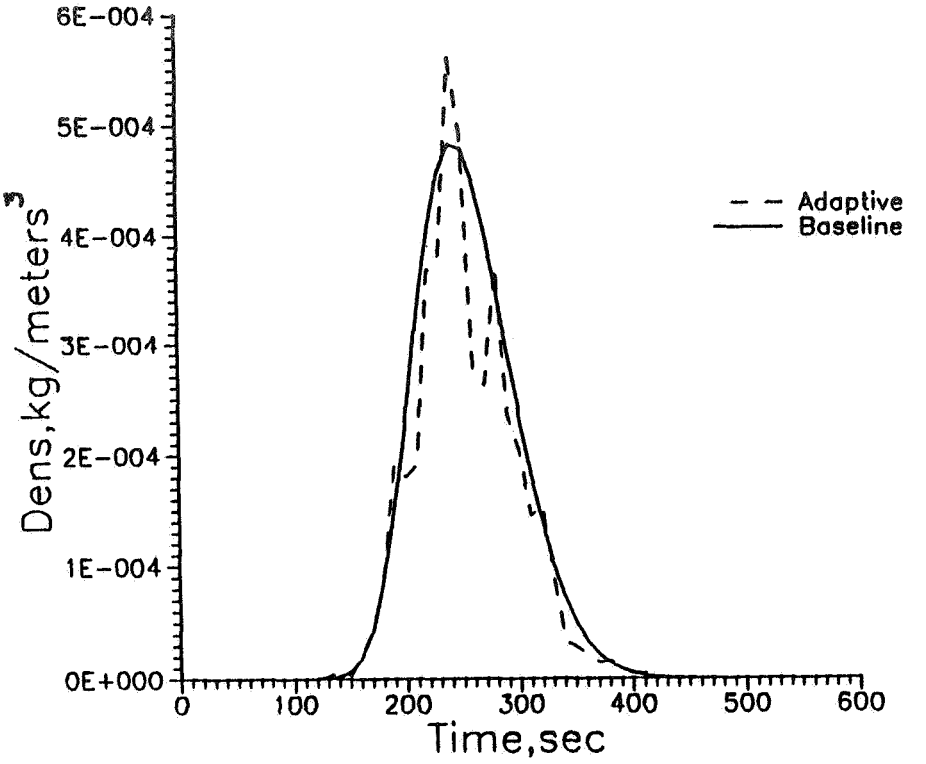
Figure 3. Flight Path Angle Windows and Typical Sensitivities for Two Proposed Mars Entry Vehicles.

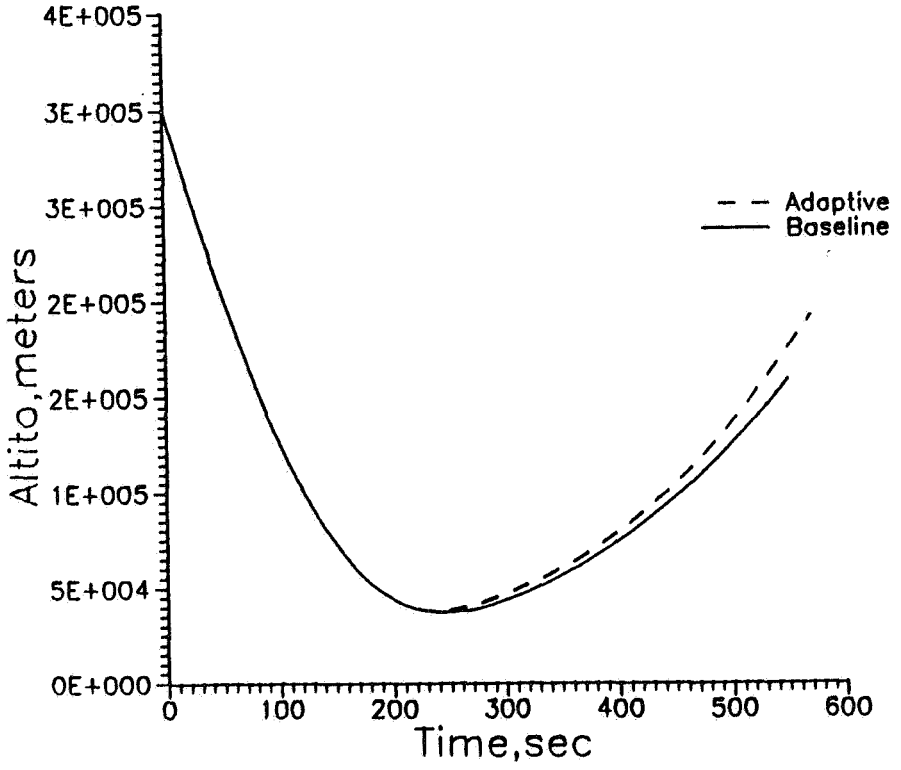


Vehicle Type 1

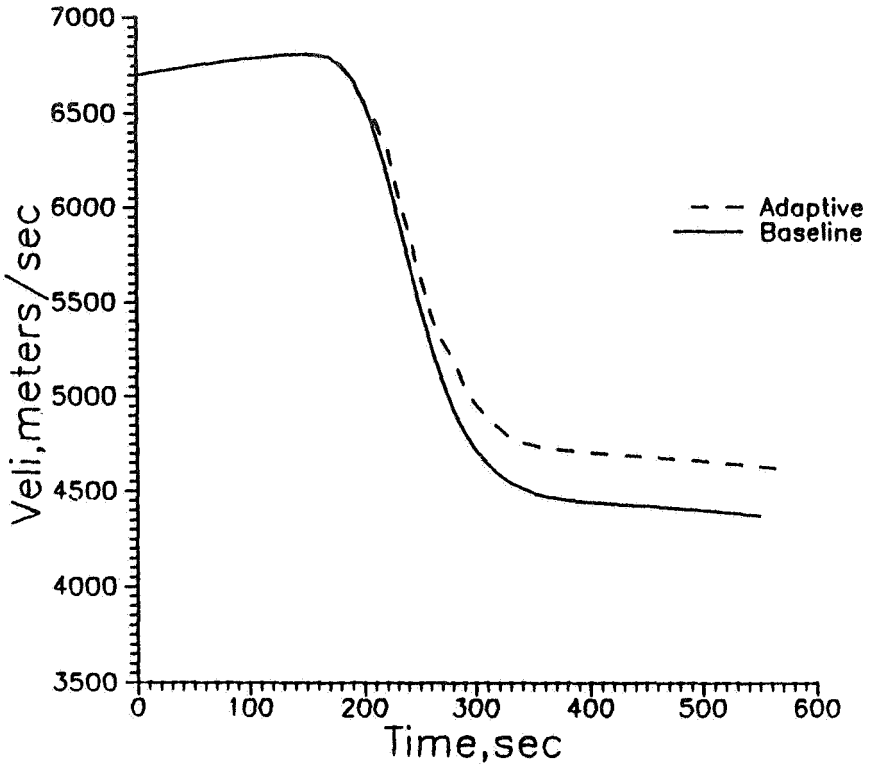
M/C_DA = 919.3 L/D = .5

Figure 4. Typical Time Histories of Bank Angle, Acceleration, and Altitude for a Martian Entry with Reduced Bank Commands.





igure 5. State and Atmosphere Time Histories for a Martian Aerocapture using Adaptive Guidance 1, with a Reduced Gain of $1x10^{-4}$, Flying Through a Random Atmosphere with Maximum Density Variations of 50 percent and for the Baseline Trajectory.



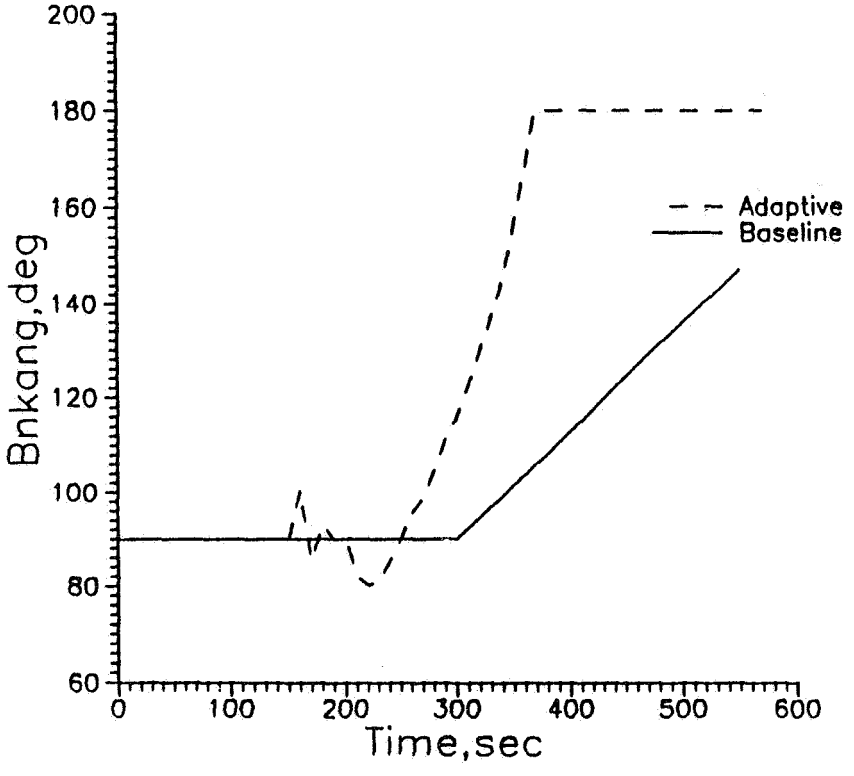


Figure 5. Continued.

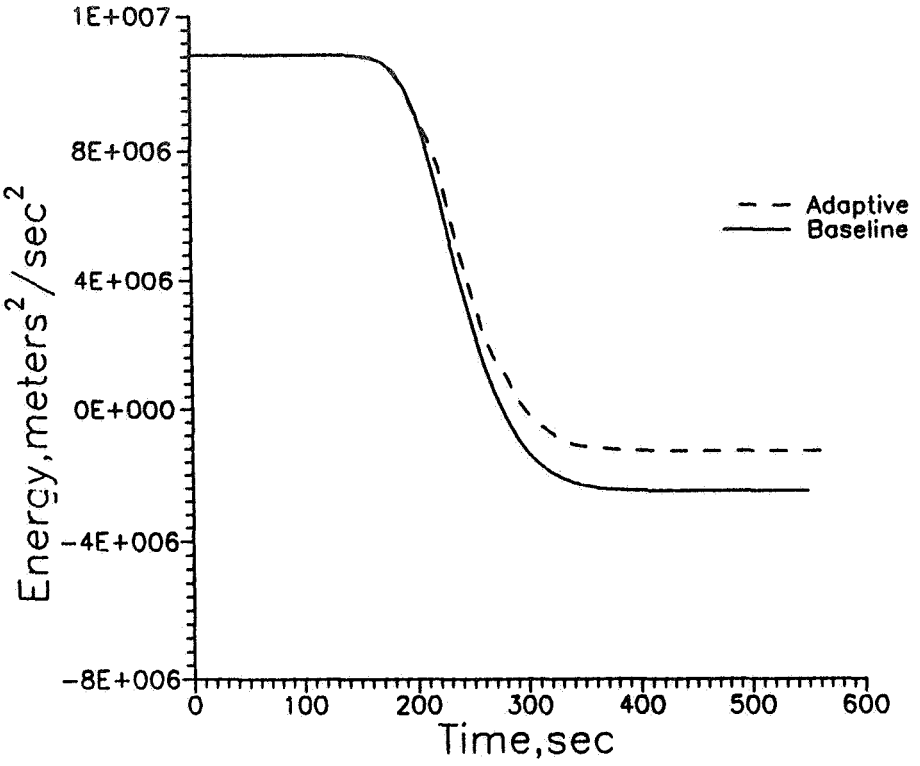


Figure 5. Concluded.

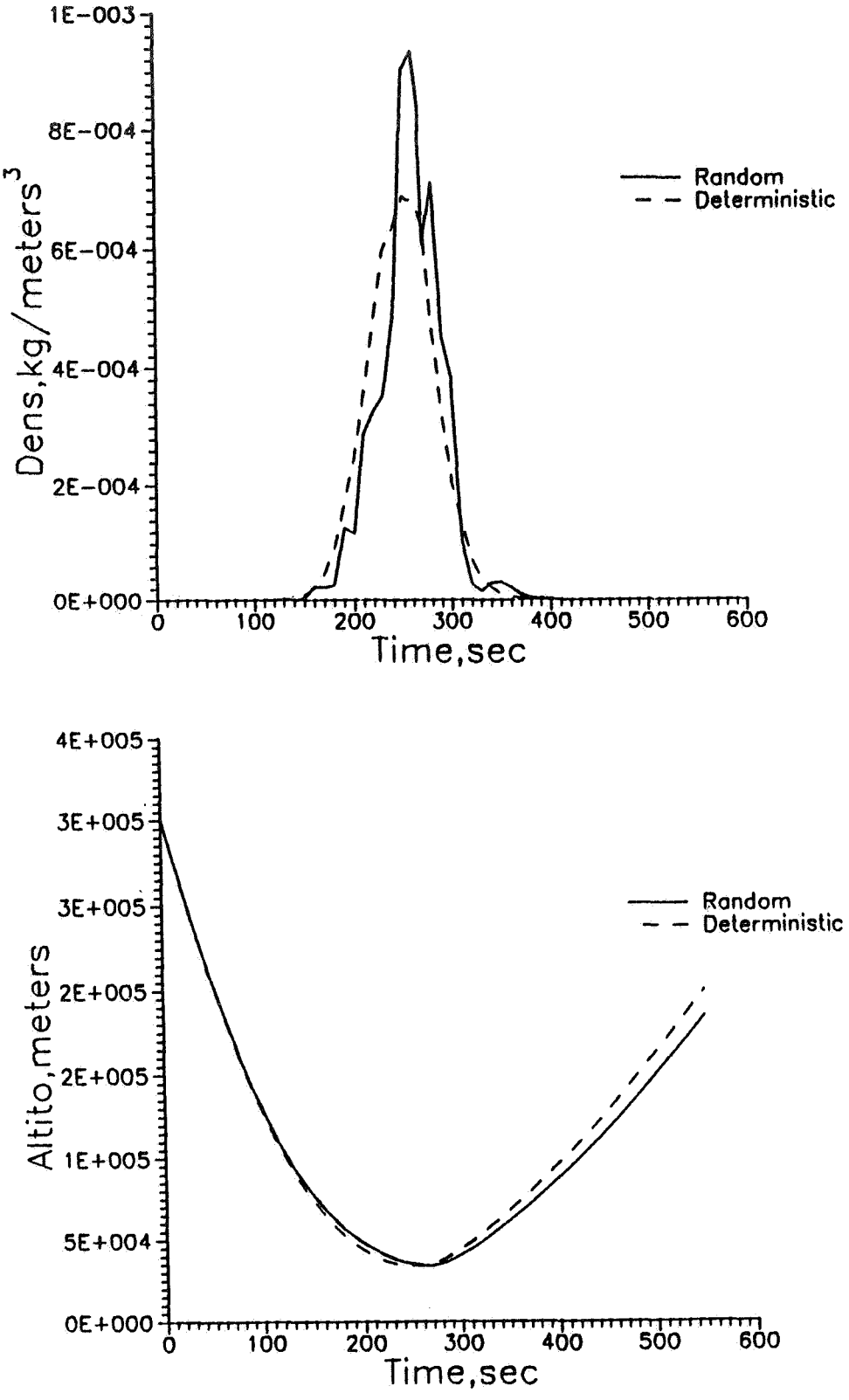
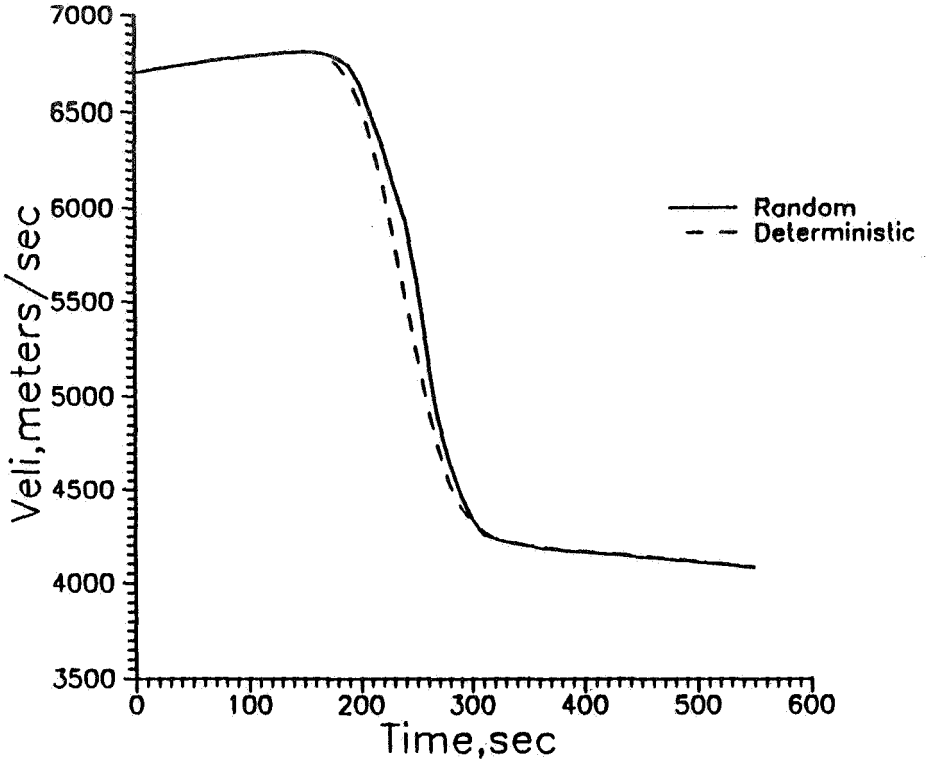


Figure 6. State and Atmosphere Time Histories for a Martian Aerocapture using Adaptive Guidance 2, Flying through a Random Atmosphere with Maximum Density Variations of 50 percent and for a Deterministic Atmosphere.



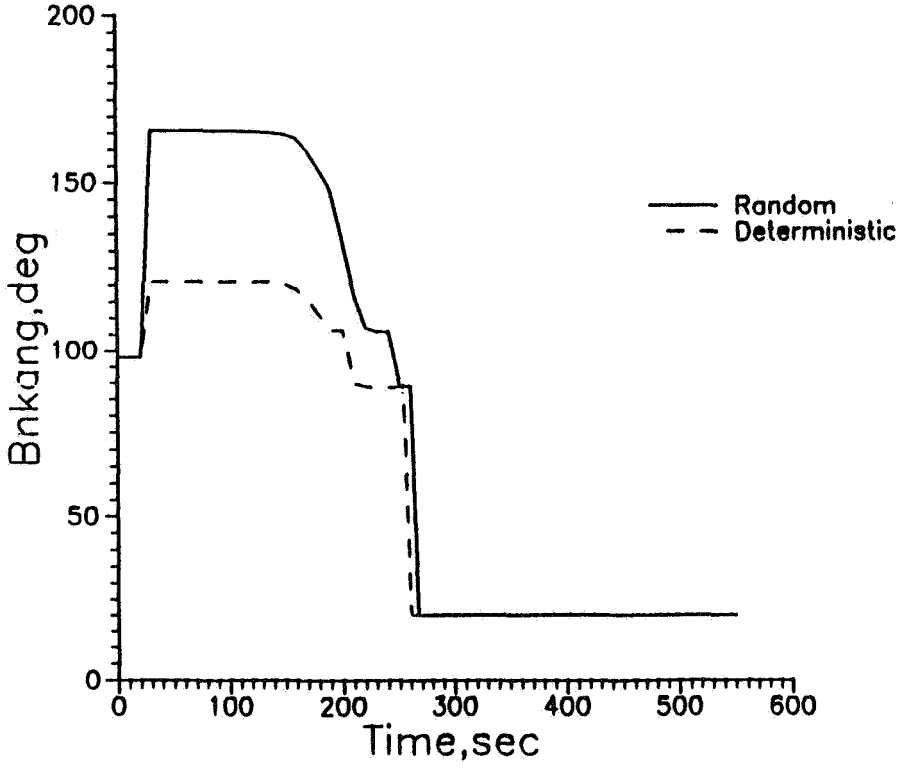


Figure 6. Continued.

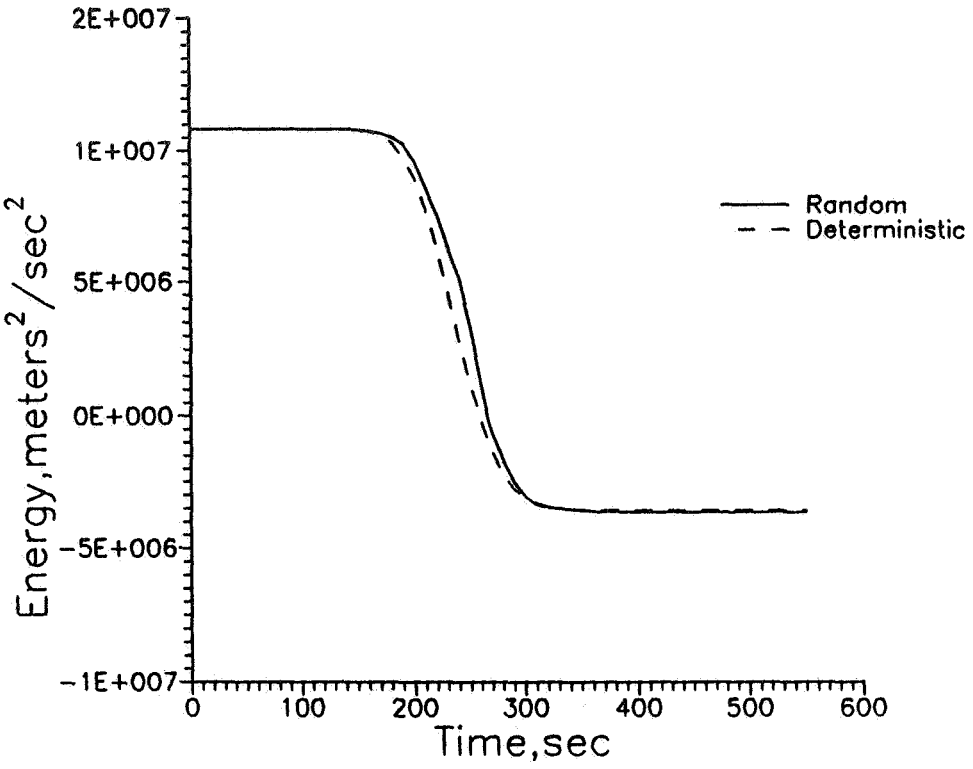


Figure 6. Concluded.

UPPER STAGE IN-FLIGHT RETARGETING TO ENHANCE GEOSYNCHRONOUS SATELLITE OPERATIONS

Otto W. K. Lee *

General Dynamics Space Systems Division
San Diego, CA

ABSTRACT

Real time utilization of propellant reserves that are not needed is available with the implementation of the in-flight retargeting capability for the Centaur Upper Stage. Application to a performance critical, geosynchronous mission is discussed. The operational duration of the satellite may be increased by selectively choosing the appropriate final orbit injection conditions. During ascent Centaur evaluates the amount of propellant excess available and adjusts the final orbit target to consume the excess. Typical satellite mission requirements are introduced to illustrate the mission analysis process to determine the pre-flight nominal target and the in-flight retarget function.

* Engineering Specialist, Trajectory and Performance Dept.

INTRODUCTION

Satellite operational characteristics may be enhanced by innovative targeting of the upper stage. Until now expendable launch systems did not have very sophisticated and adaptive processes for targeting to their desired final orbits. Flight software programs loaded on board contain fixed orbits the vehicle must fly to. Hundreds of pounds of launch vehicle capability may be set aside as reserves to ensure a high probability of a successful mission in the event of dispersed conditions. If the launch system provides nominal or above nominal performance this additional capability is not used in any way to benefit the spacecraft mission objective. There is no flexibility in the flight algorithms to improve beyond the nominally targeted mission.

An in-flight retargeting capability has been developed for the General Dynamics Centaur Upper Stage which can utilize excess propellant reserves to achieve an enhanced final orbit, i.e. an orbit which upgrades the *performance index of the satellite*. The performance index is typically related to the operational lifetime of a satellite. Centaur detects how well the lower stages have performed and releases the appropriate amount of reserves for use. When the final orbit of the upper stage is a geosynchronous transfer orbit, the retargeting algorithm employed is strictly and solely to reduce the orbit's inclination. The spacecraft performs the injection into its operational orbit, at geosynchronous altitude.

When an upper stage is used to deliver the spacecraft into its operational geosynchronous orbit a more elegant retargeting strategy is desired, one which will increase the spacecraft's lifetime. The operational lifetime of a satellite is often dependent on the amount of spacecraft propellants required to trim out any injection errors and to maintain desired orbital characteristics. For example, the satellite's mission may require it to spend its time on station at the low inclinations. The oscillatory motion of a geosynchronous orbit plane is manifested in the propagation of its inclination and RAAN (Right Ascension of the Ascending Node) periodically about the values of 7.55 degrees inclination and 0.0 degrees RAAN. The period is approximately fifty to fifty-five years for injection inclinations less than 10 degrees. The period and the variation of the elements are a function of the initial injection inclination and RAAN.

We shall investigate a sample mission that requires the satellite to maximize its dwell time at inclinations below 6 degrees. However, the nominal performance of the launch vehicle delivers it to 10 degrees inclination. Mission analysis will identify the optimal combinations of targeted/injection inclinations and RAANs which will enhance spacecraft objectives and satisfy constraints based on the expected propellants available on the upper stage.

RETARGETING STRATEGY

The Centaur is a high energy, liquid upper stage with a three burn capability. After a sub-orbital separation from the booster lower stage the first burn inserts Centaur into a low altitude park orbit. For typical geosynchronous orbit missions, the second burn provides insertion into a geosynchronous transfer orbit. The final Centaur burn circularizes the orbit at geosynchronus altitude and completes the inclination change. The objective of this capability is to evaluate actual booster performance and re-target the Centaur second and/or third burn in accordance with mission specifications. The Centaur first burn is not altered, allowing the nominal park orbit to be achieved. The intent is only to compensate for booster phase dispersions, not upper stage dispersions. This decision was made partly based on range safety implications. Other constraining factors considered in the design were: 1) simplicity, 2) minimal flight computer memory use, 3) credibility, 4) reliability, and 5) accuracy.

The re-targeting process is divided into several segments. The first segment is the in-flight detection phase; some detection criteria are imposed to determine if the vehicle is within the envelope of acceptable nominal performance. The predicted time-to-go to Main

Engine Cut-Off (MECO) calculated by the flight computer for the first Centaur burn is used to evaluate the performance of the lower stage(s). It is based on the state vector at booster/Centaur separation. The trajectory is integrated forward, assuming a nominal Centaur acceleration, until the nominal park orbit is achieved. Steering laws are incorporated into the integration process so that the time-to-go prediction is a realistic indication of the actual burn duration under these conditions. The use of nominal Centaur acceleration precludes a low thrust (long burn time) Centaur phase from influencing the re-targeting decision. Before activating the detection criterion, it is required that time-to-go converge such that the current time-to-go is within some tolerance of the value from each of the two previous guidance computation cycles, i.e.:

$$| T_{4i} - T_{4i-1} + \Delta t | < \varepsilon$$

and (1)

$$| T_{4i} - T_{4i-2} + 2\Delta t | < \varepsilon$$

Δt = elapsed time between successive guidance computation cycles
T_{4i} = time-to-go to MECO at the ith guidance cycle
ε = convergence tolerance, set at 1 sec.

When both of these are satisfied, the difference between the in-flight predicted burn duration and the nominal burn duration (loaded pre-flight) is computed:

$$\Delta T_B = T_{B\text{-}NOM} - (T - T_{MES1} + T_{4i}) \quad (2)$$

$T_{B\text{-}NOM}$ = nominal Centaur first burn duration
T_{MES1} = Main Engine Start-1 (MES1) time from go-inertial
T = time from go-inertial

The re-targeting sequence proceeds only if the absolute difference is above some threshold. Additionally, some time limit will be set after which actual Centaur accelerations are admitted (for proper steering corrections) and the re-targeting option is disengaged. Since the predicted time-to-go is a standard guidance calculation there is minimal impact to the flight software and computer memory. The long successful history of Centaur flights has proven the reliability and accuracy of the computations. In particular the NASA Atlas/Centaur-53 HEAO-C mission had similar detection software loaded on board.

A performance index is needed to quantitatively relate the performance excess (or shortfall) to revised final orbit parameters. At General Dynamics Space Systems Division the conventional measurement of performance is the Propellant Excess (PE). This is the total burnable propellants remaining in the final stage after the last main engine burn thrust decay. For this study it is baselined that the nominal mission results in a PE equal to zero. Having computed the difference in burn time from (2), the change in the amount of propellants at MECO1 (due to off-nominal performance of the booster stage) is now calculated based on a nominal mass flow rate:

$$\Delta W_{P1} = \Delta T_B \bullet \dot{w} \quad (3)$$

The change in performance which is supplied by the subsequent Centaur burns can be measured by the variation in propellant excess (ΔPE). The partial derivative based on the mass fraction across the nominal second and third burns is used to calculate ΔPE:

$$\Delta PE = (\Delta W_{P1} \bullet \partial PE/\partial W_{P1}) \quad (4)$$

The final phase of the algorithm re-targets the Centaur burns to an adjusted final orbit which will improve spacecraft capability. The variation of the targeted final orbit inclination and RAAN as a function of available PE is examined in the next sections. This functional relationship will be loaded into the flight computer. Pre-flight analysis determines the function constants for the set of desired final orbit parameters. The algorithm is structured to be able to deploy different functions dependent on the value of the PE. A complete description of the algorithm can be found in Reference 1.

PROPAGATION OF GEOSYNCHRONOUS ORBITS

Earth oblateness, gravitational attractions of the Sun and Moon, and solar radiation pressure affect the long-term orbital variations of geosynchronous satellites. Algorithms that propagate the oscillation of the orbit characteristics have been developed to estimate the stationkeeping requirements for satellites. An understanding of the time history of the satellite orbit precession provides a strategy to adjust the injection targeting for Centaur.

A semi-analytical method which predicts inclination and RAAN given the initial inclination and RAAN values has been developed by Allan and Cook (Ref. 2). This procedure is based on analysis which used an averaged disturbing function in the equations of motion. Expression by vectorial elements (in the satellite orbit frame) reduced the resultant Lagrange equations to a compact and tractable form. Chao and Baker (Ref. 3) generated a program that numerically integrates the equations. Comparison of the semi-analytical approach with the numerical method (Ref. 4) yielded favorable results for several ten-year inclination and RAAN histories. Also noteworthy is that for synchronous, circular orbits an approximate solution to the Lagrange equations yields a "stable" orbit which is at 7.55 degrees inclination and 0 degrees RAAN.

Allan and Cook provide the expression for the period of oscillation of the orbit under these influences by inverting and integrating the equation of motion along one axis:

$$T = 4 \bullet (\lambda_3 - \lambda_2)^{-1/2} \bullet (\lambda_0 - \lambda_1)^{-1/2} \bullet K(k) \quad (5)$$

where λ_1 are eigenvalues of the set of "Euler" equations and λ_0 is an integration constant
K(k) is the complete elliptic integral of the first kind and

$$k = (\lambda_2/(\lambda_3 - \lambda_2))^{1/2} \bullet \tan\Phi \quad (6)$$

Φ is the dihedral angle between the stable orbit and the orbit of interest

$$\Phi = \arccos(\cos\alpha_0 \cos i_0 + \sin\alpha_0 \sin i_0 \cos RAAN_0) \quad (7)$$

$\alpha_0 = 7.55$ deg.

Figure 1 illustrates the variation in the period of precession of the orbit based on the initial inclination and RAAN. The variation of the precessional period is about one year at the low inclinations (less than 5 degrees) but may be up to seven years at inclinations around 25 degrees. Figure 2 presents the RAAN and inclination time history for three different injection states: inclination of 8 degrees and RAAN of 270, 300, and 330 degrees. Although the difference in period is only ten months among the

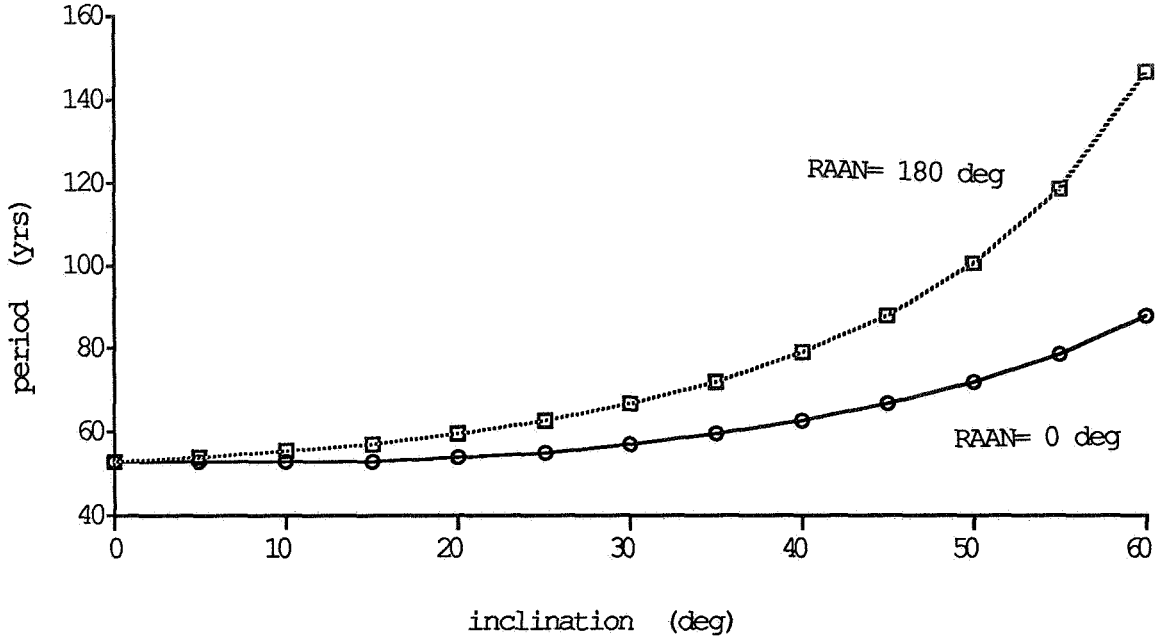


Figure 1 Variation of Precession Period with Inclination and RAAN

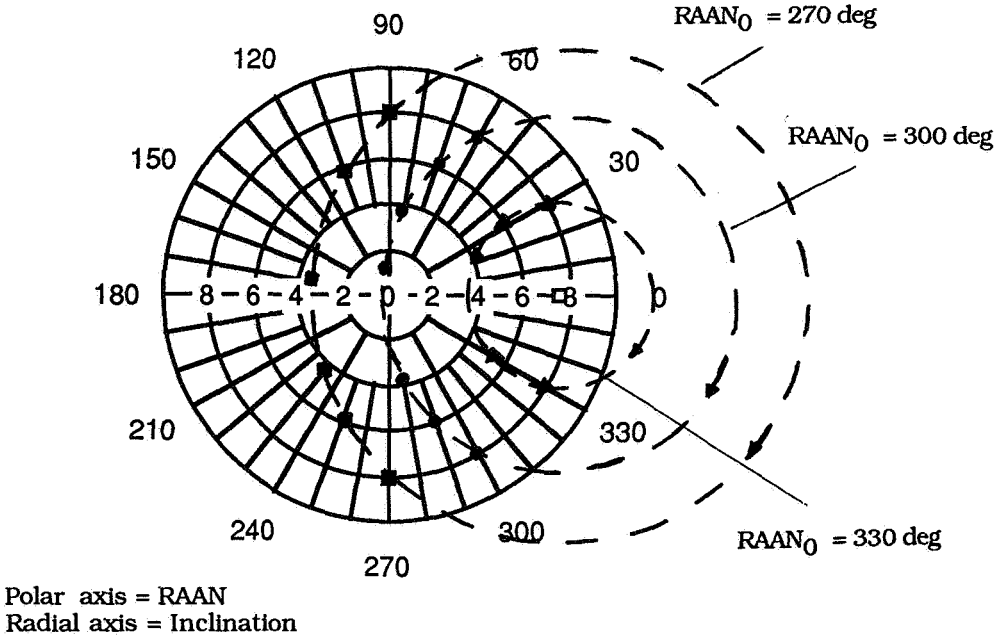


Figure 2 RAAN and Inclination History

three cases, the time spent at a particular inclination band varies significantly among them. For instance the central angle of the arc length of intersection between the 270 degree RAAN trace and the region of inclination less than 8 degrees is 91 degrees. From equation (7), a satellite drifting along this trajectory will experience a dwell time of 13.4 years at inclinations of 8 degrees or less.

$$\text{Dtime} = T \bullet (\text{Central Angle}/360) \qquad (8)$$

T = period of oscillation

Similarly, the 300 degree RAAN trajectory results in a dwell time of 15.9 years and the 330 degree RAAN trajectory yields a dwell time of 23.4 years.

UPPER STAGE MISSION DESIGN

The nominal mission design for the Centaur results in targeting to final orbit parameters which can be achieved by the launch system with a 99.87 percent probability. To guarantee this confidence a certain amount of propellants (vehicle performance) is held in reserve to compensate for dispersions in flight. With Inflight Retargeting, part of this reserve is released and used to achieve a different orbit which enhances the satellite objectives.

Given particular mission requirements such as satellite weight, the Centaur is capable of achieving a geosynchronous orbit with a certain inclination for its nominal level of performance. As reported earlier, for that inclination there is a corresponding RAAN which maximizes the time spent at certain inclination bands. This selection is done pre-flight to generate the complete nominal mission final orbit target parameters which are loaded onboard the flight computer. By adjusting the time of launch each day, achieving the desired RAAN does not incur any performance penalty.

To determine the targeted RAAN simply requires solving for the intersection of two non-concentric circles. Referring to the coordinate system used in Figure 2 the polar equations of an inclination band and the precession trace are expressed:

$$r = i \qquad (9)$$

$$r^2 - 2 \bullet s \bullet r \bullet \cos(\text{RAAN}) + (s^2 - P^2) = 0 \qquad (10)$$

i = inclination
s = inclination of "stable" point = 7.55 deg.
P = radius of the precession trace based on the initial injection inclination, i_0, and Right Ascension of the Ascending Node, RAAN_0

$$P^2 = i_0^2 - 2 \bullet i_0 \bullet s \bullet \cos(\text{RAAN}_0) + s^2 \qquad (11)$$

Equation (10) is solved by substituting eq. (9) and (11):

$$\text{RAAN} = \arccos\{\pm [(i^2 + s^2 - P^2) / 2 \bullet s \bullet i]\} \qquad (12)$$

The two RAAN values indicate the RAAN along the precession trace where the satellite enters the inclination region of interest and the RAAN at which it exits. These angles provide the arc length of travel along the precession trajectory. The central angle, β, of this arc length is calculated by the law of cosines from the triangle shown in Figure 3.

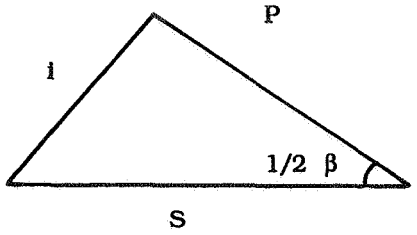


Figure 3 Geometry of Intersection Central Angle

$$\beta = 2 \arccos [(P^2 + s^2 - i^2) / (2 \bullet s \bullet P)] \quad (13)$$

Equation (8) is then used to calculate the dwell time.

Let us examine the application to a sample geosynchronous mission. The satellite requirements are: 1.) maximize the time spent at inclinations of 6 degrees or less with a minimum of 10 years and maximum of 12 years; 2.) assume no spacecraft ΔV has been allocated for out-of-plane injection usage; 3.) the total mission duration will not exceed 15 years. If the nominal launch vehicle can only deliver the spacecraft to 10 degrees inclination, this becomes a performance critical mission. A retargeting function, executed in-flight, may help meet the requirements by utilizing propellant reserves which are not needed.

In developing the retargeting strategy, the nominal target must first be established. Obviously the nominal targeted inclination is 10 degrees but the targeted RAAN is still to be determined. Since any RAAN may be achieved by selecting the appropriate time of day to launch without incurring a performance degradation, one can allow the satellite requirements to drive its selection. Figure 4 shows the dwell time, the drift time, and the total mission duration time for a range of RAANs at 10 degrees inclination. It can be seen that a RAAN of 290 degrees will satisfy the mission requirements: the dwell time at inclinations of 6 degrees and less is 10.5 years, the time to drift into that region is 4.5 years, and thus the total mission duration is 15.0 years. The nominal target is 10 degrees inclination and 290 degrees RAAN. The propellant reserves carried on board will ensure this orbit will be achieved even with three-sigma "low performance" launch system dispersions.

During flight Centaur will detect and evaluate the actual performance of the booster lower stage; if the vehicle is operating nominally, or at any level above the low three-sigma threshold, there are propellant reserves which may be utilized. This propellant excess (PE) can be used to reduce the inclination from 10 degrees. Figure 5 illustrates the effect on the dwell time, drift time, and total mission duration by reducing the injection inclination (RAAN remains the same). For this mission the maximum dwell time of 12 years is satisfied at an injection inclination of 8.4 degrees. Additional PE can then be used to reduce the drift time of 2.8 years at the injection RAAN of 290 degrees and injection inclination of 8.4 degrees. To accomplish this and to maintain the requirement of a 12 year dwell time dictates adjusting the RAAN and inclination according to the precession path (see Figure 2). To determine the RAAN and inclination pairs along this path recall equation (10) and (11).

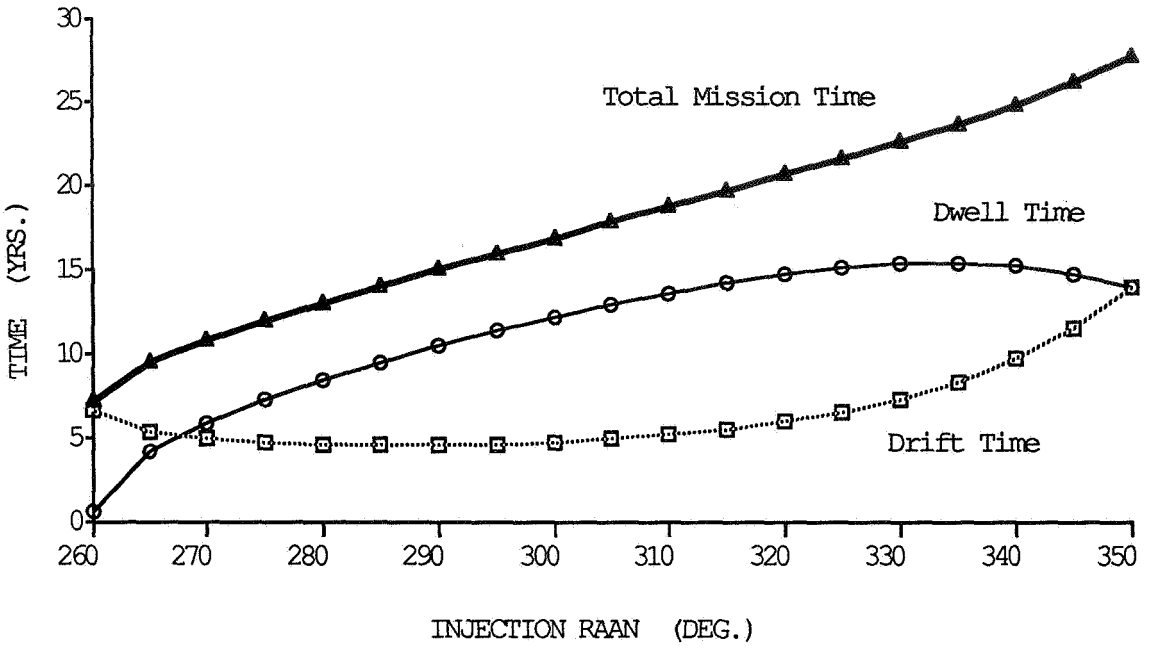


Figure 4 Dwell Time, Drift Time, and Total Mission Duration Varaition with RAAN for inclination = 10 deg.

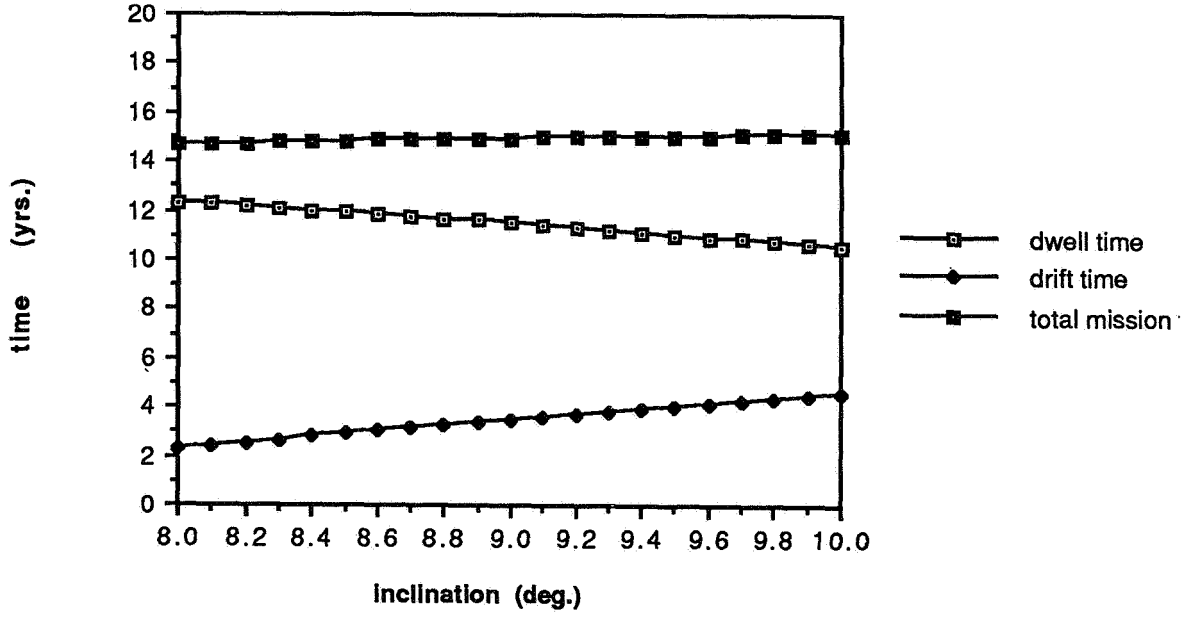


Figure 5 Dwell Time, Drift Time, and Total Mission Duration Variation with Inclination for RAAN = 290 deg.

$$i^2 - 2 \cdot s \cdot i \cdot \cos(RAAN) + (s^2 - P^2) = 0 \quad (14)$$

$$P^2 = i_0^2 - 2 \cdot i_0 \cdot s \cdot \cos(RAAN_0) + s^2 \quad (15)$$

i_0 = 8.4 degrees
$RAAN_0$ = 290 degrees
s = inclination of "stable" point = 7.55 deg.

Table 1 contains the retargeted values of inclination and RAAN and the associated change in drift times.

Table 1 Reduction in Drift Time from Retargeted RAAN and Inclination

RAAN (DEG)	INC (DEG)	DWELL TIME (YRS)	DRIFT TIME (YRS)
290	8.4	12.0	2.77
289	8.22	12.0	2.56
288	8.04	12.0	2.35
287	7.87	12.0	2.14
286	7.69	12.0	1.94
285	7.52	12.0	1.74
284	7.35	12.0	1.54
283	7.18	12.0	1.34
282	7.01	12.0	1.15
281	6.84	12.0	0.96
280	6.69	12.0	0.78
279	6.53	12.0	0.60
278	6.37	12.0	0.42
277	6.21	12.0	0.24
276	6.07	12.0	0.07

To illustrate the final form of the retargeting function that would be loaded on the Centaur flight computer for this hypothetical mission we will use typical values of propellants required to change inclination and RAAN associated with a Centaur geosynchronous orbit.

$$\delta PE / \delta i = 34.02 \text{ kg } (75 \text{ lbs}) / \text{ deg. inclination} \quad (16)$$

$$\delta PE / \delta RAAN = 9.07 \text{ kg } (20 \text{ lbs}) / \text{ deg. RAAN} \quad (17)$$

The retargeted values of inclination and RAAN from the nominal pre-flight set as a function of the propellant excess (PE) as determined in flight is shown in Figure 6.

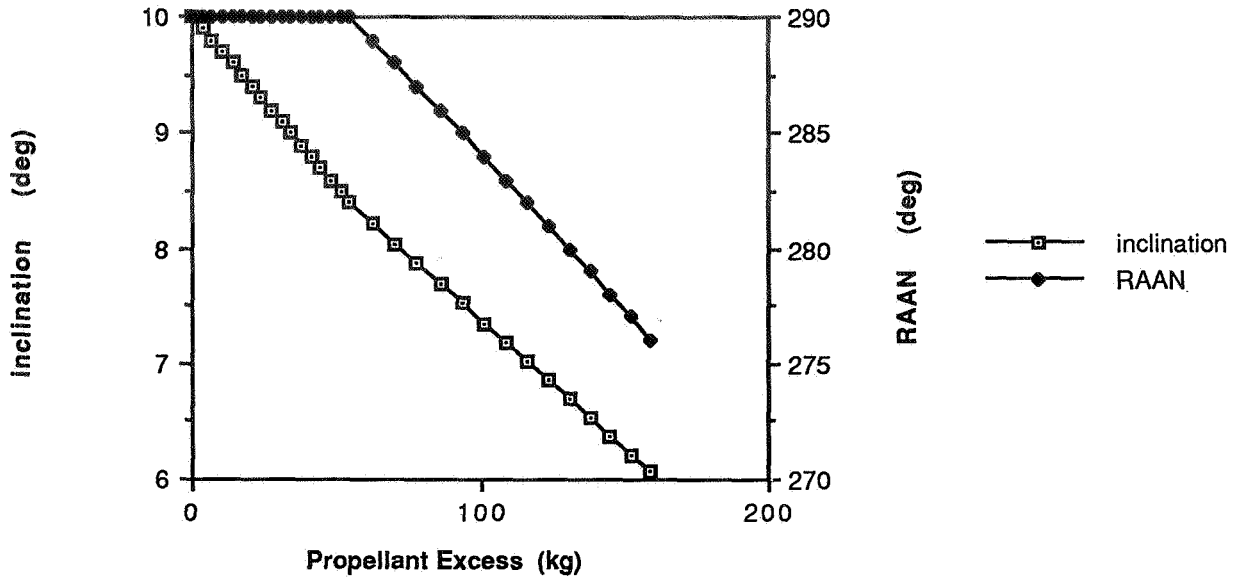


Figure 6 The Retargeting Function

CONCLUSION

Upper stage in-flight retargeting may enhance geosynchronous satellite mission objectives by using excess propellant reserves to achieve different final orbit injection conditions. By analysis of the precessional motion at geosynchronous altitudes a strategy is formulated to adjust the final orbit injection inclination and RAAN which will meet dwell time, drift time, and total mission duration requirements. The variation in inclination and RAAN as a function of the additional propellants detected to be available during flight is generated for loading onto the flight computer.

REFERENCES

1. Lee, O.W.K., "In-Flight Re-Targeting Concepts for the Atlas/Centaur", Paper AAS 89-429, Presented at AAS/AIAA Astrodynamics Specialist Conference, Aug. 7-10, 1989.

2. Allan, R.R., and Cook G.E., "The Long-period Motion of the Plane of a Distant Circular Orbit," Proceedings of the Royal Society (London), Vol. 280, No. 1380, July 7, 1964.

3. Chao, C.C., and Baker J.M., "On the Propagation and Control of Geosynchronous Orbits," J. of the Astronautical Sciences, Vol. XXXI, No. 1, Jan.-Mar. 1983.

4. Hovden, R.E., "Geosynchronous Orbit Inclination and RAAN Prediction," ATM No. 89(9975)-14, Technical Memorandum, The Aerospace Corp., Jan. 4, 1989.

CONSTANT PROPELLANT USE RENDEZVOUS SCENARIO ACROSS A LAUNCH WINDOW FOR REFUELING MISSIONS*

M. E. Hametz and R. Whittier
Computer Sciences Corporation

ABSTRACT

This study investigates active rendezvous of an unmanned spacecraft with the Space Transportation System (STS) Shuttle for refueling missions. The paper first presents the operational constraints facing both the maneuvering spacecraft and the Shuttle during a rendezvous sequence. For example, the user spacecraft must arrive in the generic Shuttle control box at a specified time after Shuttle launch. In addition, the spacecraft must be able to initiate the transfer sequence from any point in its orbit. The standard four-burn rendezvous sequence, consisting of two Hohmann transfers and an intermediate phasing orbit, is presented as a low-energy solution for rendezvous and retrieval missions. However, for refueling missions, the Shuttle must completely refuel the spacecraft and return to Earth with no excess fuel. This additional constraint is not satisfied by the standard four-burn sequence. Therefore, a variation of the four-burn rendezvous, the constant delta-V (ΔV) scenario, has been developed to satisfy the added requirement.

*This work was supported by the National Aeronautics and Space Administration (NASA)/Goddard Space Flight Center (GSFC), Greenbelt, Maryland, Contract NAS 5-31500.

1. INTRODUCTION

This report presents the results of an investigation into analysis and mission-planning techniques for unmanned user spacecraft involved in active rendezvous with the Space Transportation System (STS) Shuttle for refueling purposes. The requirements for an active rendezvous are (1) the maneuver sequence must possess a 360-degree phasing capability (i.e., the two spacecraft must rendezvous from any initial orientation) and (2) the rendezvous must be completed in a fixed amount of time. A standard four-burn rendezvous sequence has been developed for retrieval missions. In this sequence, the amount of fuel used during the rendezvous varies with the initial angular phasing between the two spacecraft, as shown in Figure 1 for a 3-day rendezvous. For refueling missions, an additional rendezvous requirement is that maneuver planning and analysis for premission planning must determine the exact amount of fuel the user spaccecraft will expend during the rendezvous sequence. This allows the Shuttle to transport only the fuel necessary to refuel the spacecraft's tanks. However, since the initial phase angle varies during a Shuttle launch window, the delta-V (ΔV) required during the rendezvous is not fixed. Therefore, the standard four-burn sequence is not appropriate for a refueling mission. Consequently, a constant ΔV rendezvous scenario was developed to meet all requirements.

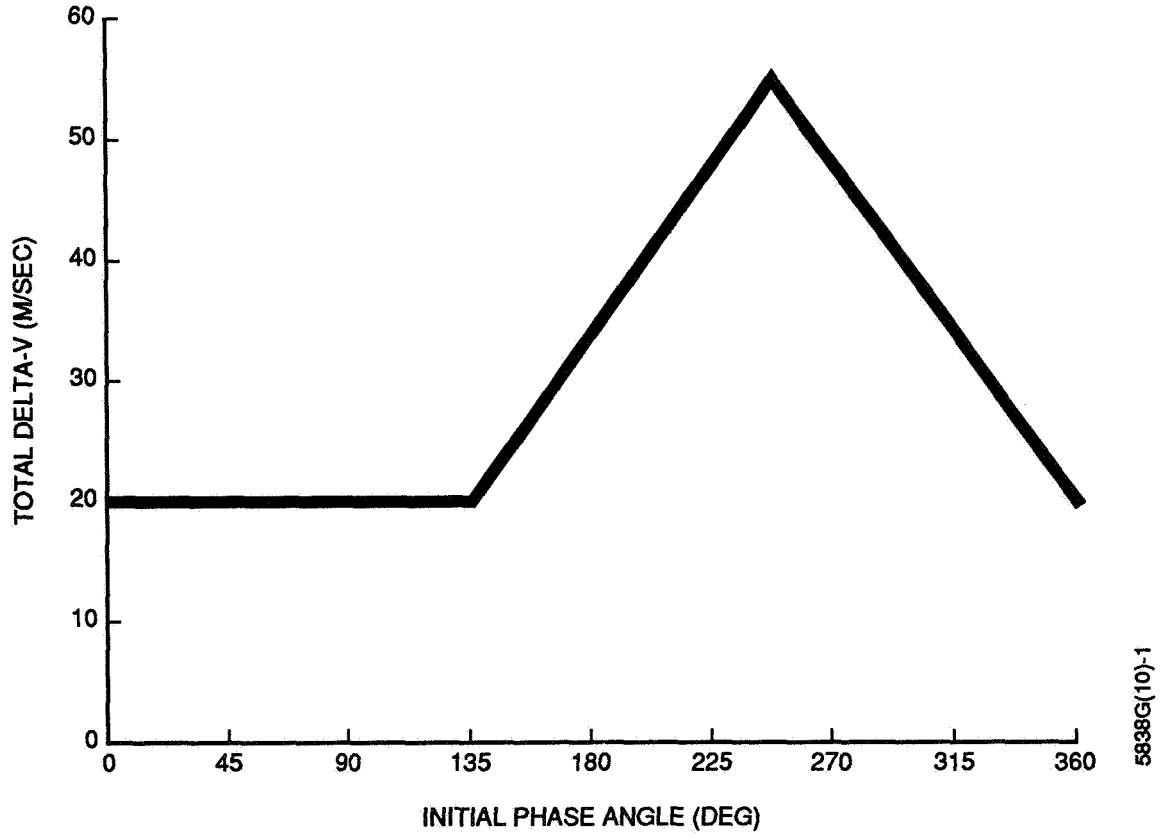


Figure 1. Total ΔV for a 3-Day Rendezvous, Standard Four-Burn Sequence

Section 2 presents background information on the derivation of the standard four-burn sequence from the standard Shuttle rendezvous policies. This sequence, which consists of

two Hohmann transfers and a coast period in a phase orbit, is shown to minimize fuel costs for any set of initial conditions. Section 3 then discusses the concepts of the constant ΔV scenario and explains how it is derived from the standard four-burn sequence. Section 4 details the equations used to solve the constant ΔV case and presents the results. Section 5 presents a summary of the conclusions reached in the report.

2. BACKGROUND: STANDARD FOUR-BURN RENDEZVOUS SEQUENCE

This section presents an overview of the rendezvous sequence designed for retrieval missions. This is essential, since the retrieval sequence is the basis for the constant ΔV case. The section first presents the requirements imposed by the Shuttle on an actively rendezvousing user spacecraft. Then, the four-burn rendezvous sequence is derived as the optimum sequence that satisfies all the Shuttle requirements while minimizing ΔV requirements.

2.1 STS SHUTTLE RENDEZVOUS REQUIREMENTS

The rendezvous sequence is initiated when mission controllers at Johnson Space Center (JSC) issue the "Go for descent" declaration. This is done after the Shuttle has achieved orbit and a systems check has determined that the rendezvous sequence may proceed. Nominally, this occurs at 5 hours mission-elapsed time (MET), or 5 hours after launch.

Upon receiving the "Go for descent" declaration, the unmanned user spacecraft (chase spacecraft) must complete its rendezvous with the Shuttle (target spacecraft) at a predetermined time, currently given as 53 hours MET. JSC refers to this rendezvous completion time as the Control Box Start Time (CBST). The rendezvous is considered complete when the maneuvering spacecraft has achieved the Shuttle control box (Figure 2) and has ceased all translational maneuvering. As illustrated, the control box is a region above and ahead of the Shuttle with its origin at the Shuttle. The horizontal component measures angular separation along the Shuttle orbit, while the vertical component measures radial distance from the Shuttle.

Upon achieving the CBST at the completion of the rendezvous, the user spacecraft must satisfy a semimajor axis and eccentricity requirement limiting the difference in apogee and perigee altitudes to 14.8 kilometers (km). In addition, a maximum angular separation of 0.03 degree (deg) in the orbital planes of the spacecraft is required. The user spacecraft must be capable of absorbing up to approximately 0.1 deg of launch dispersion error in the orbit plane of the Shuttle. Finally, the user spacecraft must be capable of handling Shuttle launch slips of up to 1 hour. This, combined with the possibility of 24-hour Shuttle launch delays, requires that the user spacecraft be capable of completing rendezvous with the Shuttle from any initial orientation (or phasing) with the Shuttle. Stated differently, the user spacecraft must possess a 360-deg phasing capability with the Shuttle.

2.2 THE STANDARD FOUR-BURN RENDEZVOUS SEQUENCE

This section describes the four-burn rendezvous sequence, which is well suited to the operational environment. This method satisfies all the Shuttle requirements while

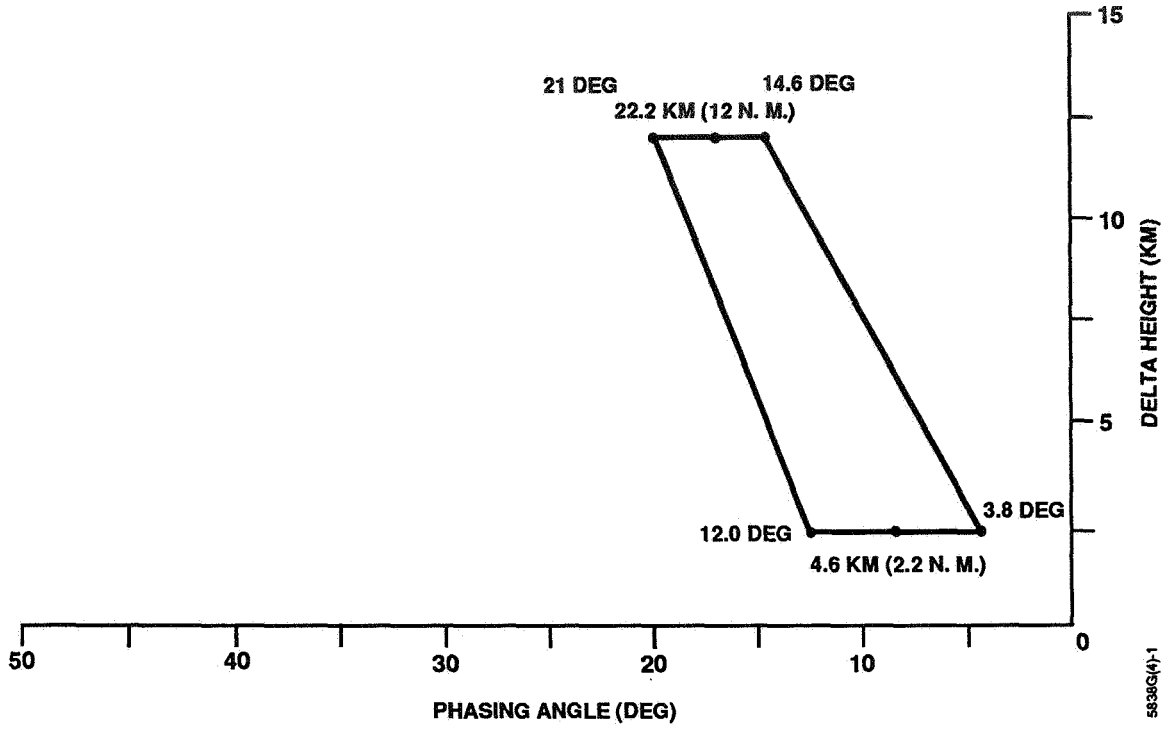


Figure 2. STS Shuttle Generic Control Box; Orbit Normal Out of Page

minimizing ΔV. The section begins with a discussion of the characteristics of the Hohmann transfer and proceeds to describe a rendezvous sequence consisting of a series of Hohmann transfers with an intermediate phase orbit. The rendezvous technique does not require any specific orbital conditions. However, to simplify the current discussion, it is assumed that the user spacecraft begins in a higher orbit than the Shuttle.

A Hohmann transfer is well known as the optimum maneuver sequence for transferring between two circular coplanar orbits. The first burn of such a maneuver places the chase spacecraft in an elliptic transfer orbit with perigee at the same altitude as the target orbit. The second burn occurs 180 deg after the first and makes the transfer orbit circular, leaving the chase spacecraft in the same orbit as the target vehicle.

If the chase and target orbits are not coplanar, a plane change must be done at some point in the maneuver sequence. This could be accomplished by executing the entire plane change in either the initial or the final orbit, independently of the altitude change to be performed. However, the transfer ΔV is optimized by a simultaneous execution of the plane-change and orbital-change maneuvers. Efficiency is further improved by distributing the plane changes between the two burns. In the examples discussed in this paper, however, rendezvous will be assumed to be coplanar. For a more detailed discussion on the Hohmann transfer as it pertains to rendezvous, see References 1 and 2.

If two spacecraft are to rendezvous using a Hohmann transfer, the correct angular separation, or phasing, must exist between the spacecraft at the initiation of the transfer. This

angle is referred to as the Hohmann phase angle (HPA). The relative periods of the two orbits determine the value of the HPA.

The synodic period represents the length of time required for spacecraft in different orbits to return to the same orientation with respect to each other. This is the time between successive occurrences of the HPA. If the synodic period is greater than the amount of time allotted for a particular rendezvous scenario, the required HPA may not be achievable for all initial orientations. For a 2-day rendezvous, the synodic period is longer than the rendezvous duration if the initial user spacecraft altitude is less than 145 km above the nominal Shuttle altitude of 315 km. For a spacecraft such as the Gamma Ray Observatory (GRO), which is nominally only 35 km above the Shuttle at the start of the rendezvous sequence, additional measures must be taken.

The required 360-deg phasing capability can be achieved while maintaining the ΔV advantages inherent in the Hohmann transfer by employing a series of Hohmann transfers. Such a sequence, the four-burn rendezvous sequence (Figure 3), consists of two Hohmann transfers. The first transfer places the chase spacecraft in an intermediate orbit called the phase orbit. The second transfer maneuvers the chase vehicle to the target spacecraft. The phase orbit is computed such that the HPA between the phase and target orbits is reached at the time of the final transfer. By varying the altitude of the phase orbit, the user spacecraft can achieve rendezvous with the Shuttle from any initial relative orientation.

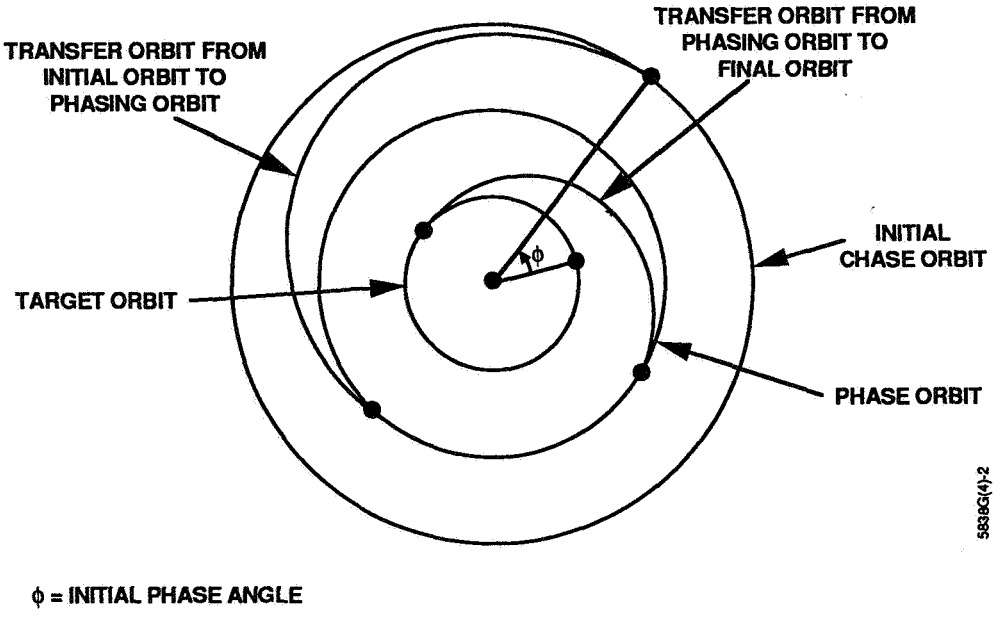


Figure 3. Four-Burn Transfer Scenario

The concept of linking in- and out-of-plane corrections to save ΔV is as applicable to the four-burn scenario as it is to the case of a direct Hohmann transfer. To combine plane

changes and altitude changes, each of the four burns must occur along the relative node defined by the intersection of the user spacecraft and Shuttle orbit planes at the termination of the rendezvous sequence. However, to simplify the cases examined in this paper, rendezvous will be assumed to be coplanar.

To apply the four-burn sequence, it is necessary to accurately compute the semimajor axis of the phase orbit, given a set of initial conditions. This is done using the following equation:

$$0 = \left[\left[\frac{\mu}{a_t^3}\right]^{1/2} - \left[\frac{\mu}{a_p^3}\right]^{1/2}\right]T - \phi - 2\pi - \frac{\pi}{\sqrt{8}}\left[\left[\frac{a_p + a_c}{a_p}\right]^{3/2} + \left[\frac{a_p + a_t}{a_p}\right]^{3/2}\right] \quad (2\text{-}1)$$

where

μ = Earth's gravitational constant (398,600.64 km3/sec2)

a_t = target spacecraft semimajor axis

a_c = chase spacecraft semimajor axis

a_p = phase orbit semimajor axis

ϕ = initial phase angle

T = rendezvous duration

Equation (2-1) is solved iteratively until a value for a_p is found that makes the right-hand side of the equation arbitrarily close to zero.

Figure 4 shows phase orbit altitude as a function of phase angle, ϕ, for a 3-day transfer from 350 to 315 km. The figure demonstrates that two phase orbit solutions exist for each initial phase angle: one above the target spacecraft and the other below. The solid portions of the curves show the phase orbit solutions having the lower ΔV requirement for each specific initial phase angle. The crossover point from the upper to the lower solution occurs when both solutions require equivalent ΔV expenditure.

Further examination of variations in phase orbit altitude with rendezvous time and initial spacecraft altitudes suggests several noteworthy trends. The phase orbit semimajor axis is essentially a linear function of phase angle, with the upper and lower solutions being nearly parallel. Furthermore, the y-intercept of the upper phase orbit altitude/phase angle function is the target spacecraft semimajor axis, and its slope varies inversely with T, the rendezvous duration. With these relationships in mind, it is possible to write the following three analytical equations, which accurately predict the phase orbit altitudes and the

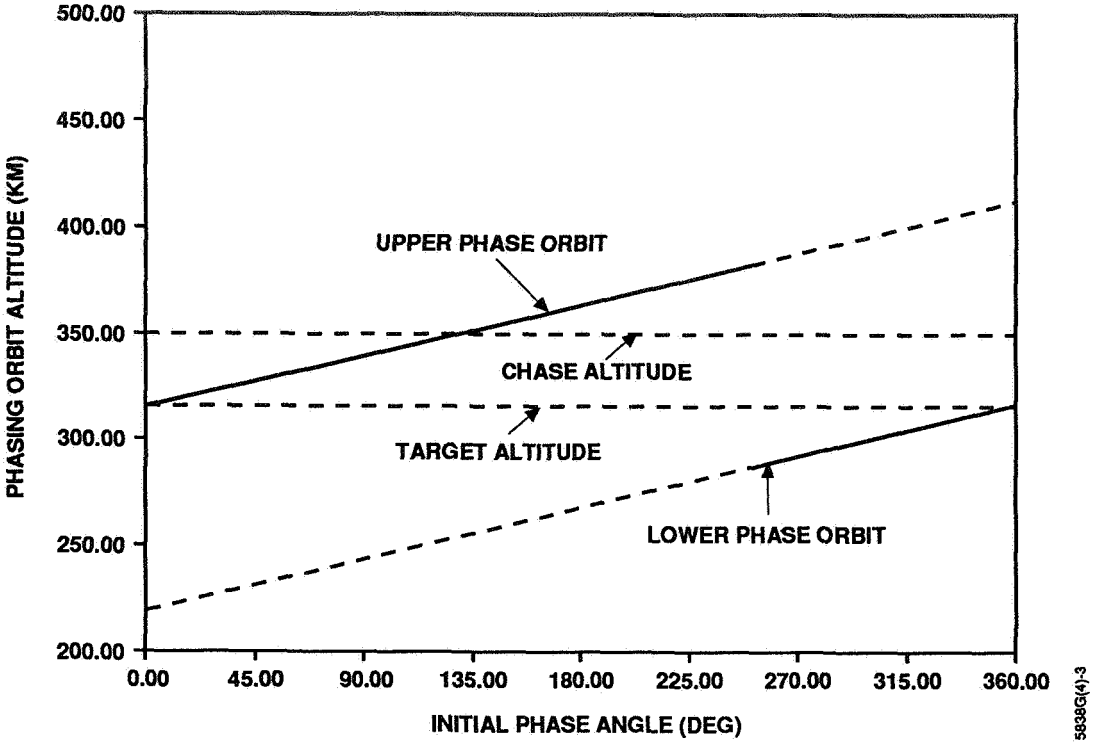


Figure 4. Phase Orbit Altitude as a Function of Initial Spacecraft Phase Angle for a 3-Day, 350- to 315-km Scenario

crossover point over the ranges of Shuttle altitudes (300 to 350 km), user spacecraft altitudes (300 to 500 km), and rendezvous durations (2 to 5 days) under consideration:

$$a_{pu}(\phi) = \frac{k_u}{T}\phi + a_t \tag{2-2}$$

$$a_{pl}(\phi) = \frac{k_l}{T}\phi + \left[a_t - 2\pi\frac{k_l}{T}\right] \tag{2-3}$$

$$\phi_c = \frac{T}{k_l + k_u}\left[a_c - a_t + \frac{2\pi k_l}{T}\right] \tag{2-4}$$

where

a_{pu} = semimajor axis of the upper phase orbit

a_{pl} = semimajor axis of the lower phase orbit

ϕ_c = phase angle at which crossover occurs

k_u, k_l = constants

The expressions for k_u and k_l were derived by taking a Taylor series expansion of an expression for phase orbit altitude based on spacecraft angular rates and assuming only linear terms to be significant. Numerical analysis can be performed to demonstrate that, in agreement with the initial simplifying assumption of a linear relationship between phase orbit altitude and ϕ, k_u and k_l do remain essentially constant over the ranges under consideration. The derivation of k_l and k_u and the associated numerical analysis can be found in Reference 3.

3. CONCEPTS OF THE CONSTANT ΔV RENDEZVOUS

The purpose of this section is to introduce the constant ΔV rendezvous scenario. A constant ΔV rendezvous means that for the same two spacecraft and rendezvous duration, a rendezvous requires the same ΔV for every initial angular orientation (phasing). Such a rendezvous is required if the Shuttle is to refuel a user spacecraft and return to Earth without excess fuel. Calculation of the fuel the user spacecraft requires during the rendezvous allows the Shuttle to transport only the fuel necessary to refuel the spacecraft's tanks.

In the standard four-burn rendezvous sequence, the fuel cost of the rendezvous varies with the initial phase angle of the two spacecraft and with the rendezvous duration. Since the possibility of launch slips and delays makes it impossible to predict the initial phase angle of the two spacecraft and the fuel cost of the rendezvous, a variation of the four-burn sequence must be developed that ensures that a constant ΔV rendezvous occurs.

The ΔV required for a four-burn sequence can be computed for specific GRO and Shuttle altitudes and rendezvous duration. For a low-Earth orbit, the ΔV of a maneuver is a linear function of the altitude change. However, since the chase and target altitudes are fixed in the case of a four-burn scenario, the total change in altitude (ΔA) and ΔV are functions of the altitude of the phase orbit. Figures 5 and 6 show the phase orbit altitudes and corresponding total ΔV costs associated with all possible initial phase angles for a rendezvous occurring between 350 km and 315 km altitude. Two phase orbit solutions exist for each phasing, one above the target orbit and the other below the target orbit. Figure 5 shows the lower ΔV cost solutions as solid lines. When the phase orbit lies between the initial and final orbits, the total altitude change and, therefore, ΔV cost remain constant. For phase orbit altitudes above the initial chase orbit, the ΔV has a maximum value when the lower phase orbit fuel cost equals the upper orbit cost. The lower solution is then employed, and the total altitude change and ΔV costs decrease.

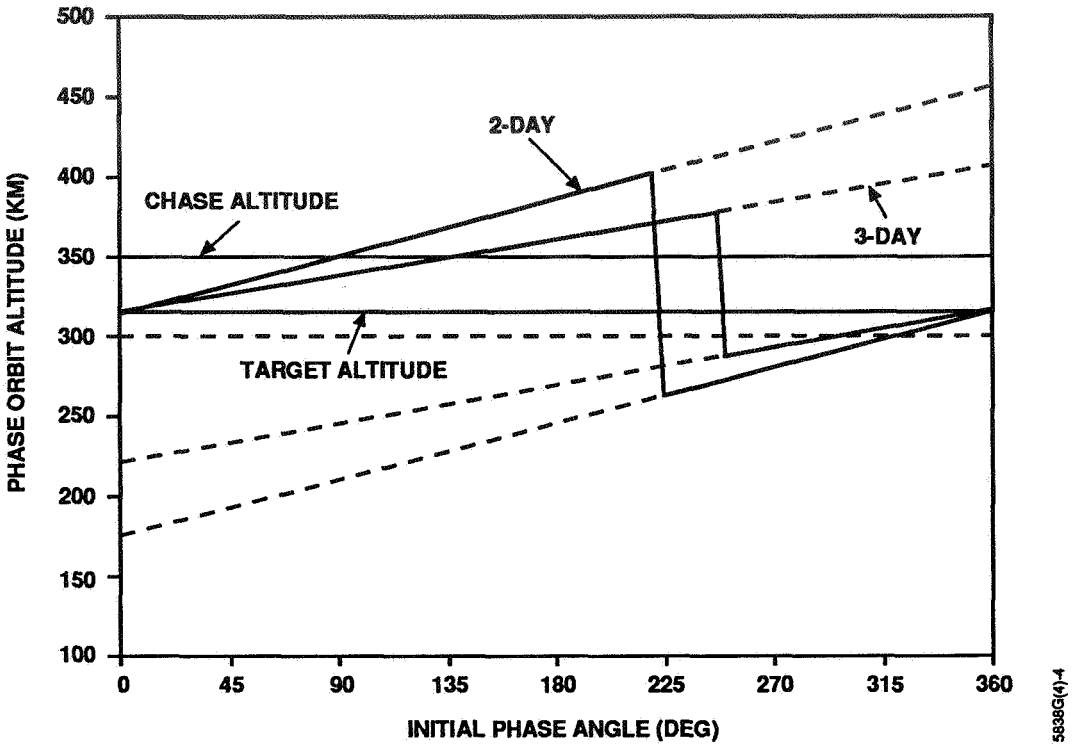


Figure 5. Phase Orbit Altitude as a Function of Initial Phase Angle (2- and 3-Day Rendezvous)

The standard four-burn scenario, composed of two Hohmann transfers and an intermediate coast period in a phase orbit, yields the optimum fuel cost solution for the rendezvous of two spacecraft. The phase orbit altitudes shown in Figure 5 result in the minimum altitude change ΔA and, thus, ΔV for each possible phasing for a rendezvous occurring between 350 and 315 km. For the 3-day curve in Figure 5, the 247-deg phase angle requires the largest total altitude change during the four-burn sequence. These phase orbit altitudes are 380 km or 285 km, respectively, and both altitudes result in a total ΔV of 55 meters/second (m/sec). Since no method exists that can change the phase orbit altitude such that ΔA decreases, a constant ΔV rendezvous is not possible below 55 m/sec. For example, the graph in Figure 6 show that a constant ΔV of 40 m/sec for a 3-day rendezvous is not possible for phase angles between 200 deg and 295 deg, since 40 m/sec is below the minimum cost profile. Instead, 55 m/sec would be the minimum constant ΔV value possible. Therefore, in order to achieve a constant ΔV of 55 m/sec, a variation of the standard four-burn sequence must be designed that requires all phase angles to use phase orbit altitudes of 380 km or 285 km. This would ensure a constant altitude change; thus, a constant ΔV scenario would exist.

Altering the phase orbit altitude is accomplished by incorporating an initial coast period into the rendezvous before burn 1. As the spacecraft coast freely in their initial orbits, the phasing between the spacecraft changes due to the differences in the angular rates of

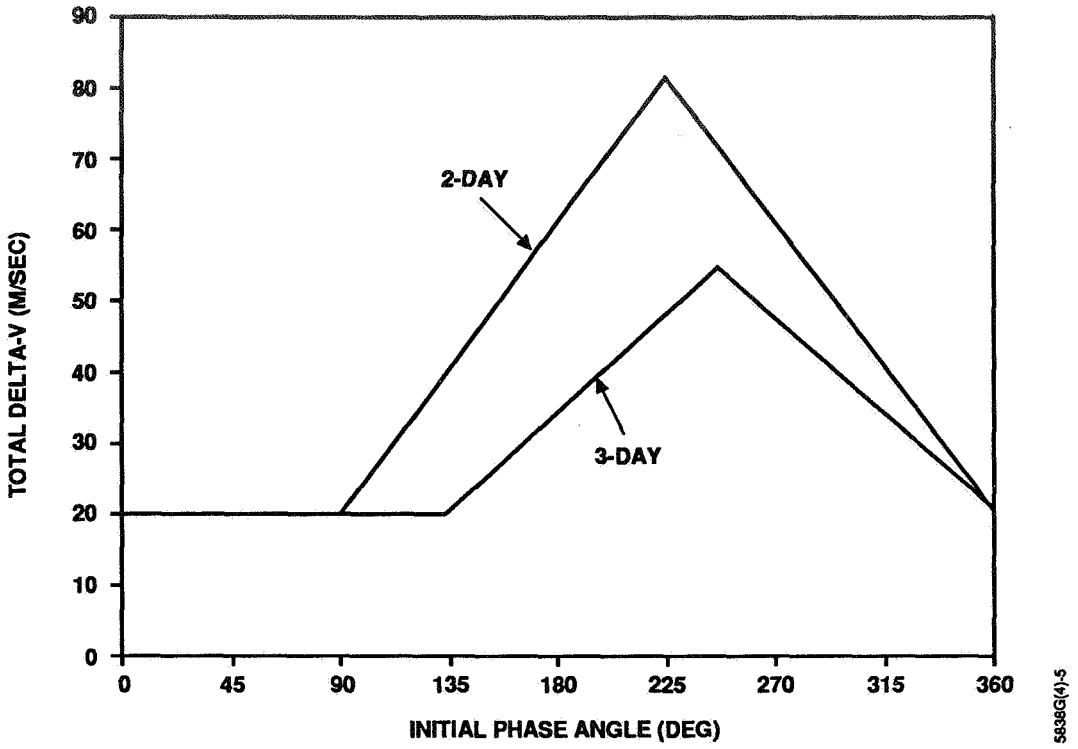


Figure 6. Total ΔV as a Function of Initial Phase Angle (2- and 3-Day Rendezvous)

the spacecraft. In addition, an initial coast time reduces the effective rendezvous duration; e.g., a 3-day rendezvous duration with an initial coast period of 12 hours has only 2-1/2 days to execute the four-burn sequence. Both the change in phasing and the reduced time interval in which to execute the rendezvous burn sequence combine to move the phase orbit farther away from the chase and target orbits, which increases the total altitude change.

Figures 7 and 8 demonstrate the effects of an initial coast on phase orbit altitude and ΔV. Initial coast durations of 12 and 24 hours are plotted along with the standard four-burn case (no initial coast). As the coast period increases, the slope of the phase orbit altitude plot increases, either raising the phase orbit farther above the initial orbit or lowering the phase orbit farther below the final orbit. As shown in Figure 8, the additional altitude change increases the rendezvous costs. For the 3-day constant ΔV example, the phase orbit must be raised or lowered to altitudes of 380 km or 285 km, resulting in a constant ΔV rendezvous of 55 m/sec. A 12-hour initial coast period will satisfy this constraint for phase angles of 225 and 285 deg, whereas a 24-hour coast is required for phasings of 205 and 325 deg. It is evident that for each initial phasing, a unique coast time is required that will enable the phase orbit to be altered to the constant ΔV altitude.

Figure 9 illustrates the constant ΔV technique for an initial phase angle of 180 deg. For a 3-day rendezvous, the target phase orbit altitude is 380 km, yielding a ΔV of 55 m/sec.

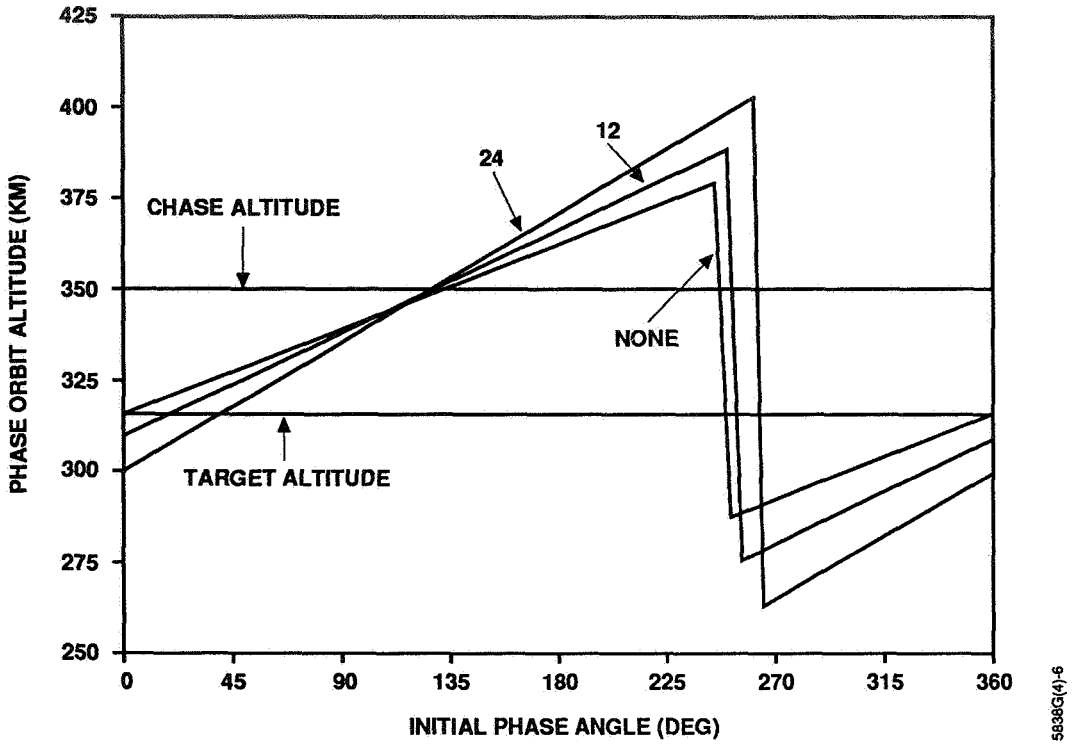


Figure 7. Phase Orbit Altitude as a Function of Initial Phase Angle (No Initial Coast Period and 12- and 24-Hour Initial Coast Periods; 3-Day Rendezvous)

If a standard four-burn rendezvous is employed, the phase orbit altitude required is 363 km, which requires a ΔV of 34.5 m/sec. Therefore, an initial coast must be performed to increase the phase orbit altitude. Figure 9 shows the progression of the initial coast period for both phase angle and phase orbit altitude. After the spacecraft has coasted for 41 hours and 18 minutes, execution of a four-burn rendezvous sequence, given the new phase angle and remaining rendezvous time, will require a phase orbit altitude of 380 km. If such a procedure is executed at each initial phase angle, the ΔV profile will be horizontal at 55 m/sec, as shown in Figure 10. Furthermore, all rendezvous requirements will be met and ΔV expenditures will be minimum.

4. COAST TIME EQUATIONS

A set of analytic formulas that describe the coast time necessary for a coplanar constant ΔV rendezvous was derived based on linear approximations for computing the phase orbit altitude, and it was determined that the relation between ΔV and altitude changes. The primary approximation is that ΔV is linearly proportional to the total altitude change,

$$\Delta V = m(\Delta A) \tag{4-1}$$

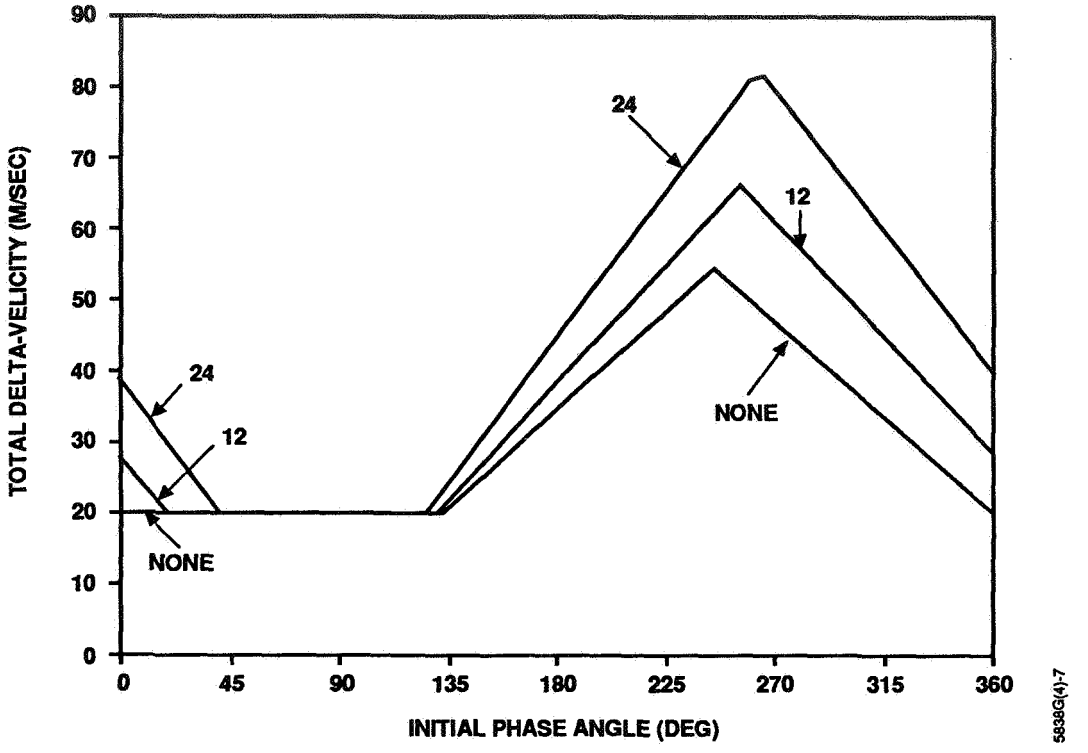


Figure 8. Total ΔV as a Function of Initial Phase Angle (No Initial Coast Period and 12- and 24-Hour Initial Coast Periods; 3-Day Rendezvous)

The slope, m, can be determined by examining a single Hohmann transfer. First, the equation for the ΔV of the transfer is modified to be a function of altitude change and final altitude. Then, the partial derivative of this function with respect to ΔA results in an equation for the slope, m. This equation has been tested numerically, and the slope has been shown to be nearly constant (approximately 0.0005 $m/sec^{-1}/km^{-1}$) for transfers below 500 km.

The total altitude change in a four-burn sequence is defined by the following equation:

$$\Delta A = |A_p - A_c| + |A_p - A_t| \qquad (4\text{-}2)$$

where

A_c = initial chase (user spacecraft) orbit altitude

A_p = phase orbit altitude

A_t = target (final) orbit altitude

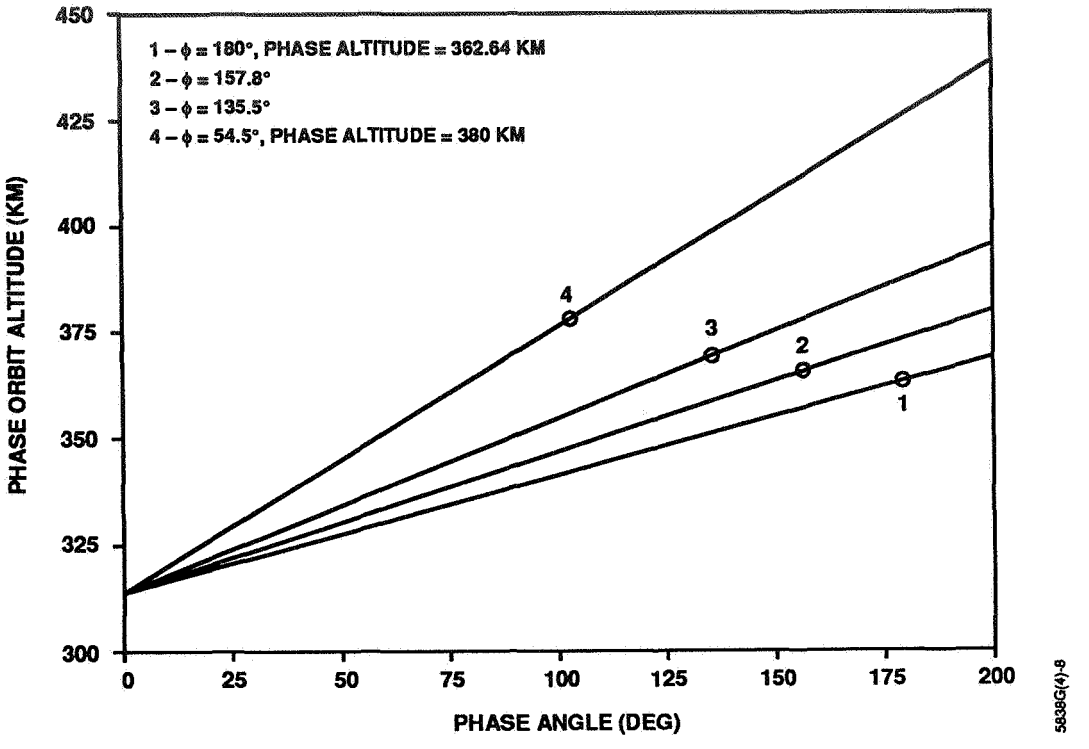


Figure 9. Changes in Phase Altitude and Angle as Initial Coast Period Increases

Absolute values are required, since the position of the phase orbit may lie between the chase and target orbits, above the chase orbit, or below the target orbit.

The phase orbit altitude can be computed from the analytic equations discussed in Section 2.2. These equations describe the upper and lower phase orbit altitudes for a standard four-burn rendezvous as functions of the rendezvous duration and the initial phase angle.

For a constant ΔV rendezvous, the equations must be modified to include the initial coast period. The modified equations include the changes in the rendezvous time and the phase angle.

$$A_{pu} = k_u \frac{(w_c - w_t)\, t + \Phi_o}{(T - t)} + A_t \tag{4-3}$$

$$A_{pl} = k_l \frac{(w_c - w_t)\, t + \Phi_o}{(T - t)} + A_t - \frac{2\pi k_l}{(T - t)} \tag{4-4}$$

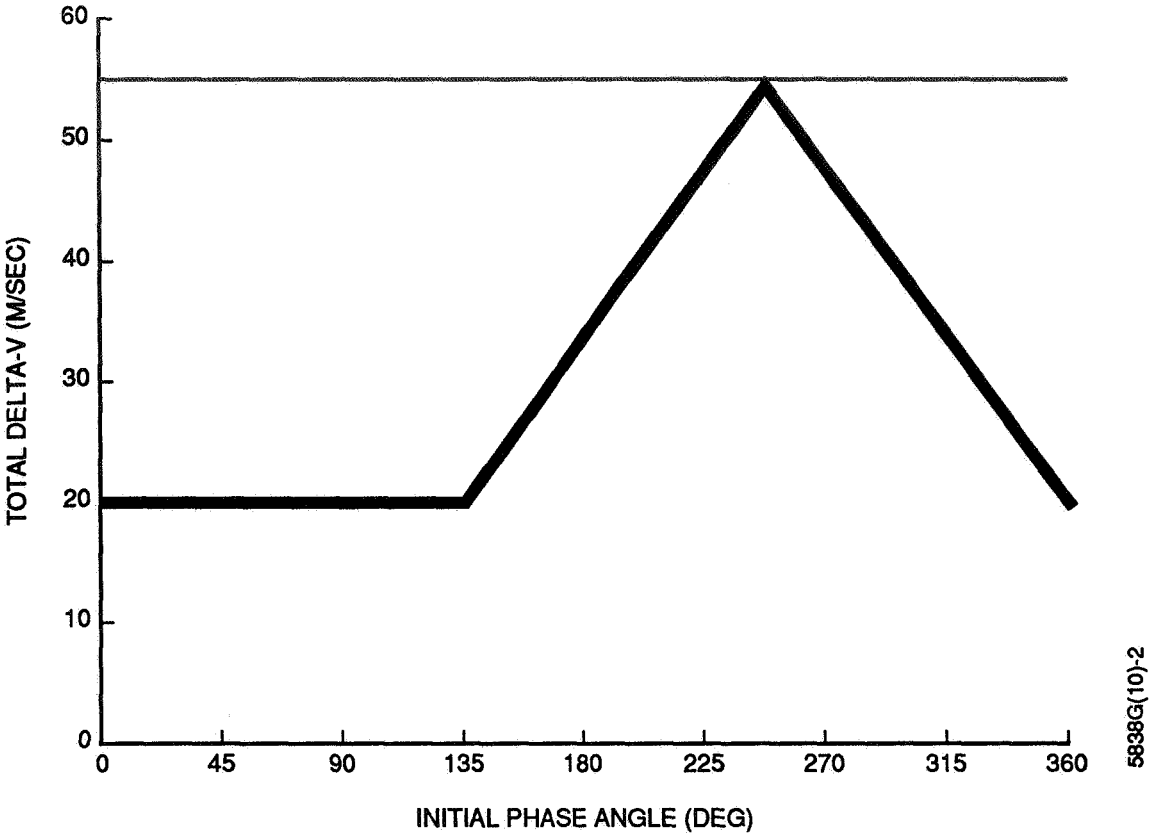


Figure 10. Total ΔV for a 3-Day Rendezvous, Standard Four-Burn Sequence Versus Constant ΔV Sequence

where

A_{pu} = phase orbit solution above the target orbit

A_{pl} = phase orbit solution below the target orbit

T = total initial rendezvous duration

t = initial coast time

w_c = angular rate of the chase vehicle in its initial orbit

w_t = angular rate of the target vehicle in its orbit

Φ_o = initial phase angle

k_u = equation constant, 45.76 km day

k_l = equation constant, 44.18 km day

When Equations (4-3) and (4-4) are then substituted into Equation (4-2), three different solutions result for ΔA as a function of phase angle, corresponding to the three possible positions of the phase orbit relative to the initial chase and target orbits. Specifically,

these solutions are (1) above the chase orbit, (2) below the target orbit, and (3) between the chase and target orbits. The equations for the total altitude change are as follows:

$$\Delta A_u = A_t - A_c + \frac{2 * k_u}{(T - t)} [(w_c - w_t)\, t + \Phi_o] \tag{4-5}$$

$$\Delta A_l = A_c - A_t - \frac{2 * k_l}{(T - t)} [(w_c - w_t)\, t + \Phi_o - 2\pi] \tag{4-6}$$

$$\Delta A_b = A_c - A_t \tag{4-7}$$

where

ΔA_u = total altitude change for a phase orbit above the chase orbit

ΔA_l = total altitude change for a phase orbit below the target orbit

ΔA_b = total altitude change for a phase orbit between the chase and target orbits

Multiplying the above equations by the slope constant, m, in Equation (4-1) yields equations for ΔV as a function of initial coast time, t.

$$\Delta V_u = m_u \left[A_t - A_c + \frac{2k_u[(w_c - w_t)\, t + \Phi_o]}{(T - t)} \right] \tag{4-8}$$

$$\Delta V_l = m_l \left[A_c - A_t + \frac{2k_l[(w_c - w_t)\, t + \Phi_o - 2\pi]}{(T - t)} \right] \tag{4-9}$$

$$\Delta V_b = m_b\, (A_c - A_t) \tag{4-10}$$

Solving for t in these equations yields

$$t_u = \frac{[T\, (\Delta V/m - A_t + A_c) - 2k_u\, \Phi_o]}{[2k_u\, (w_c - w_t) + \Delta V/m - A_t + A_c]} \tag{4-11}$$

$$t_l = \frac{[T\,(A_c - A_t - \Delta V/m) + 4\pi\,k_l - 2k_l\,\Phi_o]}{[2k_l\,(w_c - w_t) + A_c - A_t - \Delta V/m]} \qquad (4\text{-}12)$$

$$t_b = \text{undefined} \qquad (4\text{-}13)$$

where ΔV = maximum ΔV case with no initial coast.

If the optimum solution (before adding the coast) for a given phase angle is the upper solution (phase orbit above the chase orbit), then Equation (4-11) describes the coast time required before burn 1 for a constant ΔV rendezvous. If it is originally the lower solution (phase orbit below the target orbit), then Equation (4-12) should be used.

If the original phase orbit is between the initial and final orbits, the altitude change is constant. Therefore, since no equation exists as a function of t for the constant portion of the curve in Figure 6, no formula for an initial coast time can be extracted. This occurs because an initial coast will decrease the phase angle that still requires a phase orbit between the chase and target orbits, and the total altitude change remains the same. However, as the coast time increases, the phase angle reaches 0 deg, or 360 deg. At these angles, Equation (4-11) may be used to solve for the additional coast time required to achieve a constant ΔV. In summary, the coast time required from the initial phase angle would equal the coast time from the initial phasing to a phasing of 0 deg plus the coast time generated from Equation (4-10) for a phase angle of 360 deg. A simpler approach to the problem is to apply Equation (4-11) and adjust the initial phase angle by adding 360 deg. From this angle on the lower solution, an initial coast time may be found directly. Therefore, the following equation solves for t as a function of Φ for initial phase angles that require a phase orbit between the chase and target orbits:

$$t_l = \frac{[T\,(A_c - A_t - \Delta V/m) + 4\pi\,k_l - 2k_l\,(\Phi_o + 2\pi)]}{[2k_l\,(w_c - w_t) + A_c - A_t - \Delta V/m]} \qquad (4\text{-}14)$$

Equations (4-11), (4-12), and (4-14) are the only equations needed to compute the coast time before burn 1 that will result in a constant ΔV rendezvous.

Figure 11 presents the coast time solutions for a 2- and 3-day GRO/STS rendezvous. The maximum coast time equals the rendezvous duration and occurs when the low-energy phase orbit altitude (without an initial coast) equals the initial altitude. The minimum point on each curve equals zero. This occurs at the maximum ΔV case in Figure 3, which is the constant ΔV value chosen.

The coast time results from the above equations are presented in Figure 6 for 2- and 3-day rendezvous. These values were tested in rendezvous cases using the current rendezvous software, RENDEV. For the majority of the phase angles, the ΔV costs for a 2-day

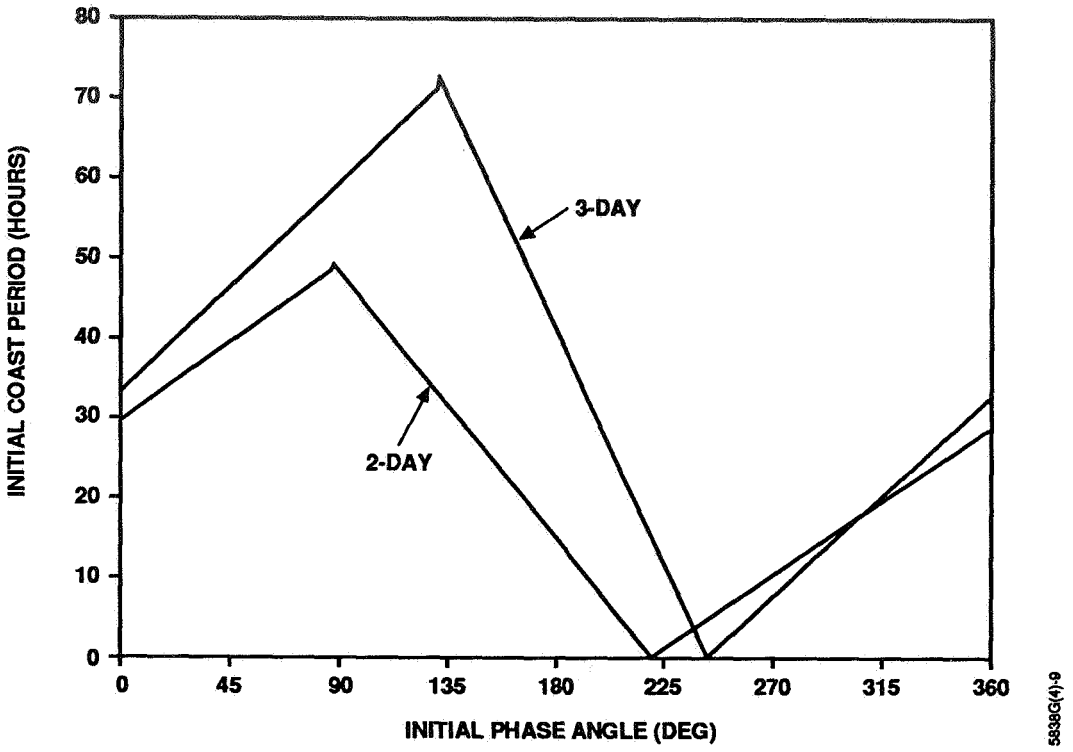


Figure 11. Initial Coast Period as a Function of Initial Phase Angle (2- and 3-Day Rendezvous)

rendezvous were within 6.5 percent of the maximum ΔV value computed by RENDEV. For the 3-day case, errors were under 5 percent for most initial phase angles. These percentage errors can be attributed mainly to the phase orbit altitude approximation formulas. The limitation of these formulas is that they do not include the time and phasing changes in the transfer orbits. This approximation can offset the phase orbit altitude by as much as 2 km. Still, the overall coast time results are good approximations for most cases. The phase angles for which the coast times are not accurate occur when the coast time is within 12 hours of the total rendezvous duration. The accuracy of the linear approximations used in the derivations declines rapidly in the 12 hours or less available for the four-burn rendezvous sequence.

5. CONCLUSIONS

This paper has considered active rendezvous between a low-Earth-orbit user spacecraft and the STS Shuttle for refueling missions. A four-burn rendezvous sequence consisting of a series of Hohmann transfers, which was derived in a previous study, is presented as an optimal solution for rendezvous and retrieval missions. However, this sequence does not readily satisfy the mission constraints for refueling scenarios. Therefore, a variation of the standard four-burn sequence is derived as a method that satisfies all constraints for refueling missions while optimizing ΔV costs.

The characteristics of the constant ΔV rendezvous scenario are described in detail. In addition, a number of analytic equations are derived that solve for the initial coast period used in the rendezvous solution. These equations were tested with current software, RENDEV, which models the four-burn sequence after the initial coast period. The coast time equations were found to be good approximations for the majority of initial phase angles. However, for a small range of phasings, the solutions are not accurate, since the approximations made in the analytic equations for phase orbit altitudes are not valid. Therefore, accurate solutions for all phasings require iterative solutions.

REFERENCES

1. Computer Sciences Corporation, *Active Rendezvous Between a Low-Earth Orbit User Spacecraft and the Space Transportation System (STS) Shuttle*, H. L. Hooper and J. R. Herrnstein, March 1989

2. ––, CSC/TM-90/6034, *Spacecraft Rendezvous and Retrieval Missions Handbook for Flight Dynamics Analysis, Revision I*, J. P. Carrico, M. E. Hametz, and J. R. Herrnstein, March 1990

3. ––, CSC/TM-87/6013, *Noncoplanar Rendezvous and the Biasing Technique*, J. R. Herrnstein and J. P. Carrico, April 1989

Optimal Trajectories Based on Linear Equations[1]

Thomas E. Carter[2]
Eastern Connecticut State University
Willimantic, Connecticut

Abstract

The Principal results of a recent theory of fuel optimal space trajectories for linear differential equations are presented. Both impulsive and bounded-thrust problems are treated. A new form of the Lawden Primer vector is found that is identical for both problems. For this reason, starting iteratives from the solution of the impulsive problem are highly effective in the solution of the two-point boundary-value problem associated with bounded thrust. These results were applied to the problem of fuel optimal manuvers of a spacecraft near a satellite in circular orbit using the Clohessy-Wiltshire equations. For this case two-point boundary-value problems were solved using a microcomputer, and optimal trajectory shapes displayed. The results of this theory can also be applied if the satellite is in an arbitrary Keplerian orbit through the use of the Tschauner-Hempel equations. A new form of the solution of these equations has been found that is identical for elliptical, parabolic, and hyperbolic orbits except in the way that a certain integral is evaluated. For elliptical orbits this integral is evaluated through the use of the eccentric anomaly. An analogous evaluation is performed for hyperbolic orbits.

[1]This project was partially supported by the 1987 and 1988 NASA/ASEE Summer Faculty Fellowship Programs at the California Institute of Technology and the Jet Propulsion Laboratory, and was performed in the Advanced Projects Group of Section 312 of the Jet Propulsion Laboratory, Pasadena, California. Another part of this work pertaining to bounded thrust trajectories was performed in colaboration with Professor John Brient of the University of Texas at El Paso.

[2]Professor, Department of Mathematics and Computer Science

1 Introduction

Many problems of optimal trajectories, maneuvers, and rendezvous of spacecraft have been investigated using linearized equations of motion. Linear equations describing the relative motion of a spacecraft near a satellite in circular orbit [1-4] have been very useful to aerospace researchers. Some of the published applications are cited here [5-8]. The linear equations describing relative motion near a satellite in an elliptical orbit [9,10] are not quite so well known, but have also been useful in applications [11-16]. These equations also generalize to a description of relative motion near a satellite in an arbitrary Keplerian orbit [17,18]. Another approach to similar problems involves the linearization with respect to the orbital parameters associated with a satellite or spacecraft [19]. Other sets of linearized equations describing the motion of an object near one of the five Lagrange points in the restricted three body problem are well known in celestial mechanics [20].

In this paper we present necessary and sufficient conditions for fuel-optimal trajectories of a spacecraft whenever the equations of motion are linear. This theory encompasses all of the preceeding examples in which the equations of motion of a spacecraft are usefully linearized about an equilibrium condition [1-20].

The work divides naturally into two distinct areas, impulsive problems, and bounded thrust problems, and is based largely on recent investigations [21,22] in these two areas. Early work on these problems was done by Neustadt [23,24] who formulated the linear impulsive spacecraft trajectory problem mathematically as an unbounded thrust problem in a class of more general nonlinear programming problems. He obtained an existence theorem and necessary conditions for solution of this problem. He presented also a precise sense in which the impulsive problem solution is a limit of the bounded thrust problem solution as the bound becomes arbitrarily large, and showed that the necessary conditions for the impulsive problem are obtained from the necessary conditions for the bounded thrust problem by passing to the limit.

Our approach differs significantly from Neustadt's, especially in the impulsive problem. We formulate this problem through the use of a finite number of velocity increments as independent variables instead of thrust functions having unbounded range. This results in a simpler problem that can be solved without the use of mathematical control theory or advanced mathematics. The bounded thrust problem is then solved by the well known Principle of Pontryagin. It is found that the two problem solutions are closely related, and insight into either problem is sometimes obtained from the other. Probably the most useful relationship is the fact that both problems are found to contain the identical new form of the primer vector function originally defined by Lawden [25].

It is shown that a class of two-point boundary-value problems associated with bounded thrust can be readily solved from starting iteratives that are obtained from a primer vector function associated with the impulsive problem.

Section two demonstrates necessary and sufficient conditions for solution of the impulsive minimum fuel problem with linear equations of motion. At the end of the section this material is applied to the case of a spacecraft near a circular orbit using the Clohessy-Wiltshire [1-4] equations of motion. Simulations of impulsive spacecraft trajectories are presented. In section three, necessary and sufficient conditions are also revealed for solution of the related bounded thrust minimum fuel problem. Simulations of optimal bounded thrust spacecraft trajectories near circular orbit are also presented for this problem using the same equations of motion. The final section indicates how this material can be applied for a spacecraft near a non circular Keplerian orbit. The Tschauner-Hempel equations of motion [9,10] are applicable to this problem. The work of Tschauner and Hempel and Weiss [15] has demonstrated that solutions of these equations can be used to define a fundamental matrix solution that applies to either circular or elliptical orbits. This work was found useful in constructing two-impulse solutions to rendezvous problems involving objects in elliptical orbits of high eccentricity [16] and can also be used to apply the theory presented here for rendezvous of a spacecraft with objects in circular or elliptical orbits. We complement this work by presenting a new form [18] of the solution of the Tschauner-Hempel equations that is valid for elliptical, parabolic, or hyperbolic orbits, and that avoids the removable singularities found in some earlier papers. This form of solution is then used to construct a fundamental matrix solution. Based on this fundamental matrix solution, the theory of sections two and three can be applied to the problem of fuel optimal rendezvous of a spacecraft with an object near a satellite in a general Keplerian orbit.

2 The Impulsive Minimization Problem

2.1 The Impulsive Problem Formulation

We let m and n be positive integers, and x_0, v_0, x_f, v_f, Δv_i $(i = 1, ..., n)$ are all elements of the m-dimensional Euclidean space $\Re^m$. The real numbers θ_0 and θ_f define a bounded interval $\Theta = [\theta_0, \theta_f]$. We denote the generalized instantaneous position and velocity of a spacecraft respectively by $x(\theta)$ and $v(\theta)$ in $\Re^m$ where $\theta \in \Theta$ represents a convenient independent variable such as the normalized time in flight, applications the true anomaly of a satellite in Keplerian orbit. [10,17] A prime is used to indicate differentiation with respect to θ. The norm or magnitude of any vector will be represented by the symbol $|\ .\ |$. The symbol $\times$ will be used with reference to rows and columns of a matrix or will denote the Cartesian Product of sets. The superscript T refers to the transpose of a matrix or vector. For each $\theta \in \Theta$ the vector $y(\theta) \in \Re^{2m}$ refers to the transpose of the pair $(x(\theta)^T, v(\theta)^T)$. Similarly the vectors y_0 *and* y_f *in* $\Re^{2m}$ denote respectively the transposes of (x_0^T, v_0^T) and (x_f^T, v_f^T). We assume that a velocity increment Δv_i is added to the velocity $v(\theta_i)$ of the spacecraft at a point θ_i $(i = 1, ..., n)$. This assumption represents an idealization of the effect of the application of n short duration thrusts to the spacecraft, resulting in instantaneous changes in velocity without change in position. This type of idealization is in common practice by mission planners and can

be useful in providing approximate data for iterative schemes to solve problems having more accurate models. The n points of application of the velocity increments can be specified a priori by the mission planner. Another possibility is that the locations of these velocity increments be determined through the optimization process. Our approach is flexible enough to include either of these possibilities or a combination of the two.

Specifically we let p be a positive integer that represents the number of elements in Θ that are specified a priori, and $N_p = (\theta_{j_1}, ..., \theta_{j_p}) \in \Theta^p$ is this ordered set of specified points in Θ where velocity increments are to be applied. For technical reasons we shall require that both end points θ_0 and θ_f are in N_p so that $p \geq 2$. We also specify an integer $r \geq 0$, and we shall seek an ordered set $N_r = (\theta_{k_1}, ..., \theta_{k_r}) \in \Theta^r$ of r elements of Θ that are to be determined through the optimization process, and that represent optimum locations for the application of velocity increments. If $r = 0$ then N_r is empty. We now set $n = p + r$, order the components of $(N_p, N_r) \in \Theta^n$ to form the ordered set $(\theta_1, ..., \theta_n) \in \Theta^n$ where $\theta_i \leq \theta_{i+1}$ $(i = 1, ..., n-1)$. We observe that it is possible to have $\theta_i = \theta_j$ for $i \neq j$. We associate a velocity increment Δv_i with each θ_i $(i = 1, ..., n)$ even though the optimization process may show some of these velocity increments to be zero. Although we associate velocity increments with the end points θ_0 and θ_f, there are problems where one or both of these velocity increments become zero. We now define the impulsive optimal spacecraft trajectory problem for linear equations of motion as follows:

2.1.1 Statement of the Linear Impulsive Minimization Problem

Having specified N_p and the integer r, find $N_r = (\theta_{k_1}, ..., \theta_{k_r}) \in \Theta^r$ and the velocity increments $\Delta v_i \in \Re^m$ $(i = 1, ..., n)$ which are added to the velocity vectors $v(\theta_i) \in \Re^m$ $(i = 1, ..., n)$ in order to minimize the total characteristic velocity

$$J = \sum_{i=1}^{n} |\Delta v_i| \tag{1}$$

of a spacecraft, whose motion is defined by the linear differential equation

$$y'(\theta) = A(\theta) y(\theta) \tag{2}$$

for each $\theta \in \Theta$ except at θ_i $(i = 1, ..., n)$, where A is a $2m \times 2m$ real matrix-valued function continuous on Θ, and whose initial and terminal conditions are given by

$$y(\theta_0) = y_0, \; y(\theta_f) = y_f. \tag{3}$$

2.1.2 Fundamental Matrix Solution

It is known from the theory of linear differential equations that a $2m \times 2m$ fundamental matrix solution which we shall denote by $\Phi(\theta)$ is associated with the linear differential equation (2), its elements are continuously differentiable, its inverse $\Phi(\theta)^{-1}$ exists and its elements are also continuously differentiable, for each $\theta \in \Theta$. As we shall show, the $2m \times m$ matrix which consists of the last m columns of $\Phi(\theta)^{-1}$ (i.e., the right hand half of $\Phi(\theta)^{-1}$) plays an important role in the determination of the optimal impulses and will be denoted by $R(\theta)$. Because of the continuity of $A(\theta)$ the $2m \times m$ matrix $R(\theta)$ of real-valued functions is continuously differentiable on Θ. Moreover if the elements of $A(\theta)$ are analytic on Θ, then the elements of $R(\theta)$ are known to be analytic on Θ also. Our principal result, as presented in the next section, is stated in terms of the matrix $R(\theta)$, and the first step in solving actual problems is to determine the matrix $R(\theta)$.

2.2 Necessary and Sufficient Conditions

As we shall see, the effects of the initial and terminal positions and velocities (3) in defining an optimal solution are determined exclusively through the vector $b \in \Re^{2m}$ which is defined as follows:

$$b = \Phi(\theta_f)^{-1} y_f - \Phi(\theta_0)^{-1} y_0. \tag{4}$$

A family of initial and final positions and velocities having the same value of the vector b is associated with a family of problems having identical optimal impulsive solutions.

In the following we define $\alpha_i = \mid \Delta v_i \mid$ $(i = 1, ..., n)$ and for each i such that $\alpha_i \neq 0$ we define $u_i = \Delta v_i / \alpha_i$. These are simply the magnitudes and the unit direction vectors of the nonzero velocity increments. We now state our principal result as follows:

Theorem 2.1 *Suppose that b is nonzero, $A(\theta)$ exists and is continuous on Θ, and N_p contains both of the end points θ_0 and θ_f of Θ. Then for each $r = 0, 1, 2, ...$ and $n = p + r$, the element $(\theta_{k_1}, ..., \theta_{k_r}, \Delta v_1, ..., \Delta v_n)$ solves the minimization problem defined by (1 − 3) subject to the imposition of the n velocity increments, over the set $\Theta^r \times \Re^{mn}$ if and only if there exists a nonzero vector $\lambda \in \Re^{2m}$ such that*

$$\alpha_i = 0 \text{ or } u_i = -R(\theta_i)^T \lambda \;\; (i = 1, ..., n) \tag{5}$$

$$\alpha_i = 0 \text{ or } \lambda^T R(\theta_i) R(\theta_i)^T \lambda = 1 \;\; (i = 1, ..., n) \tag{6}$$

$$\alpha_{k_i} = 0, \text{ or } \theta_{k_i} = \theta_0 \text{ or } \theta_{k_i} = \theta_f \text{ or } \lambda^T R'(\theta_{k_i}) R(\theta_{k_i})^T \lambda = 0 \;\; (i = 1, ..., r) \tag{7}$$

$$\sum_{i=1}^{n}[R(\theta_i)R(\theta_i)^T\lambda]\alpha_i = -b \tag{8}$$

$$\alpha_i \geq 0 \ (i = 1, .., n) \tag{9}$$

$$\sum_{i=1}^{n}\alpha_i = -b^T\lambda > 0 \tag{10}$$

$$-b^T\lambda \textit{ is a minimum on the set}\{\lambda \in \Re^{2m} \mid (6-10) \textit{ is valid}\}. \tag{11}$$

A proof of this theorem is presented in [21]. The requirement that b is nonzero in this theorem prevents the problem from becoming trivial. If $b = 0$ then the minimization problem can be solved by setting $\Delta v_i = 0$ $(i = 1, ..., n)$. In applying this theorem to actual problems one specifies the number r of points of application of velocity increments to be determined by the optimization process. Of course, the choice of the number r affects the solution of the minimization problem defined by (1-3) and the resulting minimum value of J. Although the proof introduces some additional mathematical framework and will not be considered here, it can be shown that if the elements of $A(\theta)$ are analytic functions of θ then there is an integer $r*$ such that the solution of the minimization problem for $r > r*$ does not lead to a lower value of J than does the solution for $r*$. One consequence of this fact is that N_p has no effect on the minimum value of J if the integer r is sufficiently large. For this reason, if r is sufficiently large, then there is no loss in generality in making the assumption that N_p contains the elements θ_0 and θ_f. The determination of the smallest such integer $r*$ for a wide class of problems defined by (1-3) is a problem whose solution would reveal to the mission planner the maximum number of thrust impulses needed for a mission. Neustadt's work [23] would suggest that $r* \leq 2m$. We can sometimes keep the value of r general. If in the determination of the locations for velocity increments, we obtain $\theta_i = \theta_j$ for $i \neq j$ (i.e. repeated roots) then the value of r can be lowered without altering the minimum value of J.

2.2.1 The Primer Vector Function

Some of the results of this theorem can be interpreted in terms of the well known primer vector theory, introduced by Lawden [25] and developed for impulsive spacecraft trajectory applications by Lion and Handelsman [26,27]. We shall refer to the function $q : \Theta \longrightarrow \Re^m$ defined by $q(\theta) = R(\theta)^T\lambda$ as the *primer vector* function. In terms of this function, (6) and (7) respectively become

$$\alpha_i = 0 \textit{ or } \mid q(\theta_i) \mid = 1 \ (i = 1, ..., n) \tag{12}$$

$$\alpha_{k_i} = 0, \textit{ or } \theta_{k_i} = \theta_0, \textit{ or } \theta_{k_i} = \theta_f, \textit{ or } \mid q\ (\theta_{k_i}) \mid' = 0 \ (i = 1, ..., r). \tag{13}$$

Geometrically, this says that nonzero optimal impulses must occur where the magnitude of the primer is unity and, if they are at interior points of Θ, where the magnitude of the primer is tangent to a horizontal line one unit above the horizontal axis. This and other related geometric conditions were observed by Lion and Handelsman, even though the class of problems under their investigation was very different from the above.

2.2.2 A Priori Specification of Velocity Increments

As previously indicated, this theorem is presented in such a way that the mission planner can specify some of the locations of the velocity increments a priori. It can be made even more flexible by allowing the mission planner also to specify completely some of the velocity increments a priori, as well. This is accomplished as follows. We assume that an additional s velocity increments are completely specified by the mission planner where s is a positive integer, and N_p is extended to include the locations of these velocity increments. The total number of velocity increments is then $n+s$, and the s completely specified velocity increments will be denoted $\Delta v_{n+1}, ..., \Delta v_{n+s}$ with respective locations $\theta_{n+1}, ..., \theta_{n+s} \in N_p$. The Theorem can be restated with this adjustment and only minor changes in the proof are necessary. The principal change occurs in Lemma 4.1 of the next section where the vector b is adjusted as follows:

$$b = \phi(\theta_f)^{-1} - \phi(\theta_0)^{-1}y_0 - \sum_{j=1}^{s} R(\theta_{n+j})\Delta v_{n+j}$$

This definition of b replaces (4) only in the following adjusted result:

Theorem 2.2 *Suppose that b, as defined above, is nonzero, $A(\theta)$ exists and is continuous on Θ, and that s velocity increments $\Delta v_{n+1}, ..., \Delta v_{n+s} \in \Re^m$ are specified a priori and N_p contains their respective locations $\theta_{n+1}, ..., \theta_{n+s}$ and the end points θ_0 and θ_f of Θ. Then for each $r = 0, 1, 2, ...$ and $n = p - s + r$, the element $(\theta_{k_1}, \cdots, \theta_{k_r}, \Delta v_1, ..., \Delta v_n)$ solves the minimization problem defined by (1-3), subject to the imposition of the $n + s$ velocity increments, over the set $\Theta^r \times \Re^{mn}$ if and only if there exists a nonzero vector $\lambda \in \Re^{2m}$ such that (5-11) are satisfied where b is defined above.*

2.3 Applications

This approach to the solution of the minimization problem defined by (1-3) requires solving the system (5-11). The primary computational problem is in the solution of (6-8) which are quadratic in λ and linear in α_i $(i = 1, ..., n)$. In most application to spacecraft trajectories, the position and velocity vectors of the spacecraft are in three dimensional Euclidean space so that $m = 3$, or if the motion is restricted to an orbital plane, $m = 2$. If the equations of motion of the spacecraft are based on linearization about a satellite in circular or Keplerian orbit, we conjecture that it is not necessary to employ more than four impulses for one orbital period or less $(\theta_f \leq \theta_o + 2\pi)$, so we recommend setting $n = 4$ for these problems. This is in agreement with a similar conjecture which was made for bounded thrust problems. [28,29] Neustadt's results [23] for the general linear unbounded thrust problem show that the maximum number of impulses is 6 for $m = 3$ and 4 for planar problems.

2.3.1 Types of Problems

We list the following types of problems:

1. Planar problem with four impulses with specified locations. Here (6) presents four quadratic equations in four unknowns to determine λ. For some cases these can be solved by hand or by a computer programmed for symbol manipulation. Half of these solutions are eliminated by the test $-b^T\lambda > 0$ of (10). The remaining solutions are used in (8) to solve for $\alpha_1, ..., \alpha_n$ by Gauss-Jordan elimination. All that do not satisfy (9) are thrown out. Usually only one solution remains after this test. If more than one solution remains, the value or values of λ that satisfy (11) are retained, and from this the velocity increment directions are determined through (5). An outstanding feature of this problem is that a change in the initial or terminal conditions does not require a solution of (6) again. The solutions of (6) are a set of vectors like spokes on a wheel, any one of which may point out the correct solution. A change in the boundary conditions defines another vector from the set which satisfies condition (6) and defines the optimal impulses only through the solution of sets of linear equations (8) to determine $\alpha_1, ..., \alpha_n$ and then applies the tests (9-11). An example of this type of problem is presented in the next section.
2. Four impulse planar problem with intermediate locations determined optimally. For this problem (6) and (7) determine a set of six equations in the six unknowns λ, θ_2 and θ_3. After obtaining multiple solutions of these, the tests (9) and (10) and the linear equations (8) can be used as in the preceding. For practical problems it may be preferable computationally only to approximate the solution of θ_2 and θ_3. We guess the locations θ_2 and θ_3 and solve the preceding problem 1). We repeat this process several times with new values of θ_2 and θ_3. We then pick the set which most closely approximates (7) . In this manner (7) can be approximated to any desired accuracy. An alternative is to use gradient methods on θ_2 and θ_3 to minimize J.
3. Four impulse three dimensional problem with specified locations of impulses. In this case (6) and (8) define a set of 10 equations in the 10 unknowns $\lambda, \alpha_1, \alpha_2, \alpha_3, \alpha_4$. Each solution of this system is subjected to the tests (9) and (10) as in the previous problems. In case of multiple solutions, the test (11) is applied.
4. A two impulse solution with impulses at the ends. Rather than using the equations (5-11) for this trivial problem, it is easier to use Lemma 4.1 from a recent work [21] and solve for the two impulses through the linear equations

$$R(\theta_o)\Delta v_o + R(\theta_f)\Delta v_f = b. \tag{14}$$

This method is computationally simpler than the two impulse method of Weiss [15] used for high eccentricity elliptic orbits.

2.3.2 Spacecraft Maneuvers near Circular Orbit

As an application, we consider the problem of determining optimal maneuvers of a spacecraft near a satellite in circular orbit. The linearized equations of motion with reference to a coordinate frame fixed in the satellite were determined independently by Wheelon, Clohessy and Wiltshire, Geyling, and Spradlin [1-4]. With these equations, optimal impulsive solutions were investigated by Prussing [6] in 1969 using primer vector theory. Our approach is to find and invert a fundamental matrix solution, and apply Theorem 2.1. The equations of motion when put in the state vector form $y'(t) = A(t)y(t)$, are determined by the 6×6 matrix

$$A(t) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2\omega & 0 \\ 0 & 3\omega^2 & 0 & -2\omega & 0 & 0 \\ 0 & 0 & -\omega^2 & 0 & 0 & 0 \end{bmatrix}, \tag{15}$$

Table 1: The sixteen solutions of Eq.(6) for a planar circular orbit

		1	2	3	4
1	λ_1	0.065987656	0.106065818	-0.065987656	-0.106065818
	λ_2	-0.621919006	-0.999647	0.621919006	0.999647
	λ_3	-0.651106246	0.2387088	0.651106246	-0.2387088
	λ_4	0	0	0	0
2	λ_1	-0.098201378	-0.098201378	-0.098201378	0.098201378
	λ_2	-1.798123431	0.052928935	1.798123431	-0.052928935
	λ_3	-0.182280857	0.182280857	0.182280857	-0.182280857
	λ_4	0.436298624	0.436298624	-0.436298624	-0.436298624
3	λ_1	0	0	0	0
	λ_2	1	$\frac{-3}{5}$	-1	$\frac{3}{5}$
	λ_3	0	0	0	0
	λ_4	0	$\frac{4}{5}$	0	$\frac{-4}{5}$
4	λ_1	0	0	0	0
	λ_2	0	0	0	0
	λ_3	$\frac{\sqrt{5}}{5}$	$\frac{\sqrt{5}}{5}$	$-\frac{\sqrt{5}}{5}$	$-\frac{\sqrt{5}}{5}$
	λ_4	$\frac{\sqrt{5}}{5}$	$-\frac{\sqrt{5}}{5}$	$-\frac{\sqrt{5}}{5}$	$\frac{\sqrt{5}}{5}$

where the real number ω is the orbital angular speed of the satellite, and the flight time is denoted by t. It is not difficult to find and invert a fundamental matrix $\Phi(t)$ associated with $A(t)$, and obtain

$$\Phi(t)^{-1} = \begin{bmatrix} 1 & -6\theta(t) & 0 & 3\theta(t) & 2 & 0 \\ 0 & -2 & 0 & 1 & 0 & 0 \\ 0 & 3\sin\theta(t) & 0 & -2\sin\theta(t) & -\cos\theta(t) & 0 \\ 0 & -3\cos\theta(t) & 0 & 2\cos\theta(t) & -\sin\theta(t) & 0 \\ 0 & 0 & \sin\theta(t) & 0 & 0 & \cos\theta(t) \\ 0 & 0 & \cos\theta(t) & 0 & 0 & -\sin\theta(t) \end{bmatrix} \tag{16}$$

where $\theta(t) = \omega t$. From (16) $R(t)$ is evident. We drop the argument t and consider θ as the independent variable.

We shall consider only the planar case. For this problem we find the 4×2 matrix

$$R(\theta) = \begin{pmatrix} 3\theta & 2 \\ 1 & 0 \\ -2sin\theta & -cos\theta \\ 2cos\theta & -sin\theta \end{pmatrix}. \tag{17}$$

For simplicity we shall pick $\theta_o = 0, \theta_f = 2\pi$, and we shall select four impulses at $0, \pi/2$, $3\pi/2$, and 2π. Based on earlier studies for optimal trajectories with bounded thrust [8] we suspected that these locations are not far from the optimal locations. After computation of the λ values based on these locations, this suspicion was confirmed by substituting these values of θ_i into (7). The values of the matrix $R(\theta_i)$ $(i = 1, 2, 3, 4)$ are easily obtained so that (6) is a manageable set of four quadratic equations in the four unknown components of $\lambda; \lambda_1, \lambda_2, \lambda_3, \lambda_4$. Solution of these four quadratic equations reveals the following $2^4 = 16$ solutions which we denote by λij $(i, j = 1, 2, 3, 4)$. These solutions are presented in Table 1. We consider three distinct sets of initial and terminal conditions for computation of the optimal impulses using the method outlined in 1) of Section 2.3.1. We find that each of the three cases requires a different value of λ from Table 1, and a different number of impulses for optimality. (We neglect very small impulses, and assume that they are zero because of their negligible effect in comparison with much larger impulses in the same case.)

Case 1. A spacecraft is one unit behind a satellite in the same circular orbit. Its initial relative velocity is zero, and its object is to rendezvous with the satellite having the same velocity at $\theta_f = 2\pi$. The tests (8-11) determine $\lambda 12$ as optimal. For this value of λ, the solution of (8) establishes that $\alpha_2 = \alpha_3 = 0$, a two impulse solution. The optimal impulses are both horizontal and opposite:

$$\Delta v(0) = \begin{bmatrix} \frac{1}{6\pi} \\ 0 \end{bmatrix}, \Delta v(2\pi) = \begin{bmatrix} \frac{-1}{6\pi} \\ 0 \end{bmatrix}.$$

A computer simulation of the optimal trajectory is presented in Figure 1.

Case 2. A spacecraft is one unit above a satellite in circular orbit with the same initial velocity as the satellite. Its object is to rendezvous with the satellite at $\theta_f = 2\pi$ having the same final velocity. For this problem the tests (8-11) determined $\lambda 21$

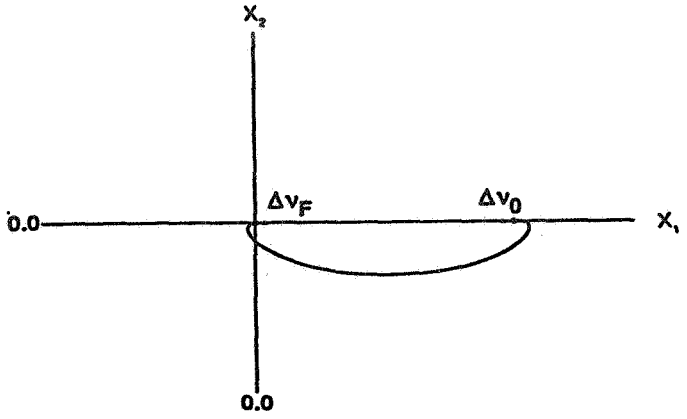


Fig. 1. Optimal trajectory with two impulses.
$x(0)^T = (1,0), v(0)^T = (0,0), x(2\pi)^T = (0,0), v(2\pi)^T = (0,0)$

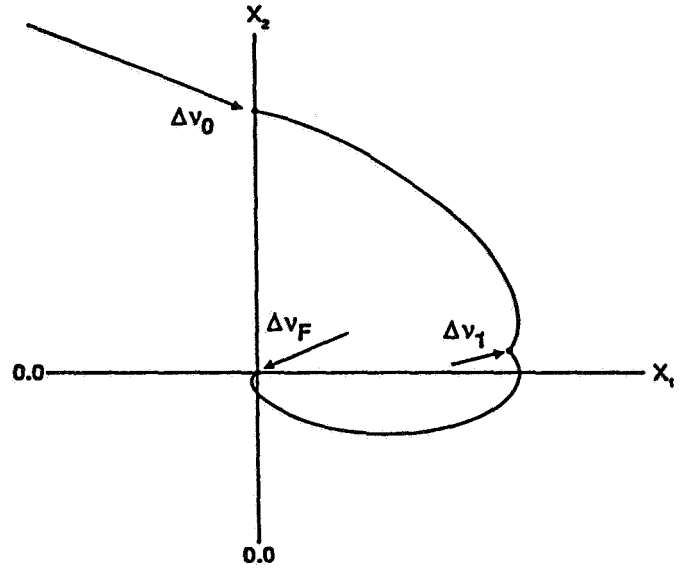


Fig. 2. Optimal trajectory with three impulses.
$x(0)^T = (0,1), v(0)^T = (0,0), x(2\pi)^T = (0,0), v(2\pi)^T = (0,0)$

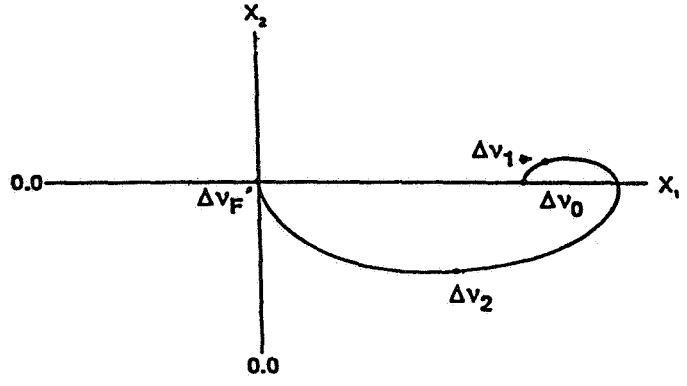


Fig. 3. Optimal trajectory with four impluses.
$x(0)^T = (1,0), v(0)^T = (0,0), x(2\pi)^T = (0,0), v(2\pi)^T = (0,0.427)$

as optimal. For this value λ, eq.(8) defined $\alpha_3 = 0$, a three impulse solution. The optimal impulses are

$$\Delta v(0) = \begin{bmatrix} 1.6294 \\ -.66667 \end{bmatrix}, \Delta v(\frac{\pi}{2}) = \begin{bmatrix} .39010 \\ .09640 \end{bmatrix}, \Delta v(2\pi) = \begin{bmatrix} -.06332 \\ -.02591 \end{bmatrix}.$$

A computer simulation of the optimal trajectory using these impulses is found in Figure 2.

Case 3. A spacecraft is in circular orbit one unit behind the satellite with the same velocity as in Case 1, but in this case its object is to reach the satellite with a positive vertical velocity. This causes quite a different trajectory from Case 1. The optimal λ changes to $\lambda 22$ and a four impulse solution results. The optimal impulses are

$$\Delta v(0) = \begin{bmatrix} -.02729 \\ .03436 \end{bmatrix}, \Delta v(\frac{\pi}{2}) = \begin{bmatrix} .08965 \\ .01194 \end{bmatrix},$$

$$\Delta v(\frac{3\pi}{2}) = \begin{bmatrix} -08965 \\ .01194 \end{bmatrix}, \Delta v(2\pi) = \begin{bmatrix} .02729 \\ .03436 \end{bmatrix}.$$

The optimal trajectory is presented in Figure 3.

Each of the figures can be compared with similar figures in [8] which are based on bounded thrust and continuous velocity. The similarities in the shapes are apparent.

Extension of this work to other Keplerian orbits requires the replacement of the Clohessy-Wiltshire equations with the Tschauner-Hempel equations [9,10] for elliptical orbits or the equations in [18] for more general orbits. For the Tschauner-Hempel equations the fundamental matrix solution has been found and inverted by Weiss [15], so that the matrix $R(\theta)$ can be defined.

3 Bounded Thrust Space Trajectories

3.1 The Bounded Thrust Minimization Problem

We let m be a positive integer and x_o, v_o, x_f, v_f are elements of the Euclidean Space $\Re^m$. The real numbers t_o and t_f define a closed and bounded interval $T = [t_o, t_f]$ where $t_o < t_f$. A dot above a variable will denote differentiation with respect to $t \in T$. The Euclidean norm or magnitude of a vector will be referred to by the symbol $|\ |$. The superscript T denotes the transpose of a vector or a matrix. The elements y_o and y_f in $\Re^{2m}$ are defined by $y_o^T = (x_o^T, v_o^T)$ and $y_f^T = (x_f^T, v_f^T)$. We let U denote the class of *admissible control functions*, which is the set of all Lebesgue measurable functions that map T into the closed unit ball U in $\Re^m$ a.e. on T.

We shall consider the linear nonhomogeneous differential equation

$$\dot{y}(t) = A(t)y(t) + \beta w(t) \tag{18}$$

which is defined a.e. on T subject to the initial condition

$$y(t_o) = y_o, \tag{19}$$

and the terminal condition

$$y(t_f) = y_f, \tag{20}$$

where A is a real $2m \times 2m$ matrix valued function on $T, w(t)^T = (0^T, u(t)^T)$ where 0 is the zero element in $\Re^m, u \in U$, and β is a positive real number. We shall assume throughout that the elements of A are real valued analytic functions on T. A solution y of (18) at $t \in T$ is an element of $\Re^{2m}$, and we shall write $y(t)^T = (x(t)^T, v(t)^T)$ where $x(t) \in \Re^m$ and $v(t) \in \Re^m$ for each $t \in T$. We define the *cost* of any element $u \in U$ to be

$$J[u] = \int_{to}^{t_f} |\ u(t)\ |\ dt \tag{21}$$

which represents the Lebesgue integral of the norm of u over T. Here we are assuming that a spacecraft has constant exhaust velocity, and that its total fuel consumption over the interval T is proportional to $J[u]$. The linear bounded thrust minimization problem can therefore be stated as the problem of minimizing $J[u]$ over U subject to the conditions (18 - 20).

Many spacecraft maneuvers and rendezvous missions can be formulated in terms of this problem. In these cases the equations of motion of a spacecraft are approximated by the linear equations (18) where t represents a flight time or related variable such as the true anomaly of a satellite in Keplerian orbit, and $\beta u(t)$ usually represents an applied acceleration on the spacecraft caused by the engine thrust. The problem is a bounded control problem because of the restriction that $u \in U$ (i.e. $|\ u(t)\ | \leq 1$ a.e. on T). For simplicity, we assume the spacecraft mass is constant over T so that β is a positive constant representing the maximum magnitude of the applied acceleration. We can, however, include the effect of the variation of mass as in previous work [29]. To avoid an additional equation that is nonlinear, we shall not do this. If the matrix $A(t)$ is partitioned into four $m \times m$ matrices so that the top left is zero and the top right is the $m \times m$ identity, then $\dot{x}(t) = v(t)$, and if t represents time in flight then the elements $x(t)$ and $v(t)$ can be interpreted respectively as relative positions and velocities of the spacecraft.

3.2 Necessary and Sufficient Conditions

Before we present our principle result, we shall place a restriction on the matrix $A(t)$ that will simplify the solution of the minimization problem by eliminating singular abnormal solutions. The matrix $A(t)$ can be partitioned into four $m \times m$ matrices as follows:

$$A(t) = \begin{pmatrix} A_{11}(t) & A_{12}(t) \\ A_{21}(t) & A_{22}(t) \end{pmatrix} \quad (t \in T). \tag{22}$$

We shall say that the matrix valued function A is *primer-compatible* if the only solution of the two equations

$$A_{12}(t)^T p(t) = 0, \ \dot{p}(t) = -A_{11}(t)^T p(t) \tag{23}$$

on T is the solution $p(t) = 0$ $(t \in T)$. It is readily seen that A is primer-compatible if $A_{12}(t)$ is nonsingular for each $t \in T$. For this reason we see that in the usual problems in which $\dot{x}(t) = v(t)$ for each $t \in T$ the matrix A is primer- compatible because $A_{12}(t) = I$ the $m \times m$ identity matrix. By differentiating we see also that A is primer-compatible if $\dot{A}_{12}(t) - A_{11}(t)A_{12}(t)$ is nonsingular for each $t \in T$. Other conditions for primer-compatibility can be obtained by differentiating further.

We now present our principal result. The proof can be found in [22].

Theorem 3.1 *Suppose that the elements of the $2m \times 2m$ matrix $A(t)$ are analytic on T, and that A is primer-compatible. Then u is a minimum of $J[u] = \int_{t_o}^{t_f} \mid u(t) \mid dt$ over U subject to the conditions (18 - 20) if and only if there exists a real number $\ell_o \geq 0$ and $\lambda \in \Re^{2m}$ such that $q(t) = R(t)^T\lambda$ for each $t \in T$ and*

$$y(t) = \Phi(t)\left[\Phi(t_o)^{-1}y_o + \beta\int_{t_o}^{t} R(\tau)u(\tau)d\tau\right] \quad (t \in T) \tag{24}$$

where $\Phi(t)$ is any fundamental matrix associated with $A(t)$ for each $t \in T$, its inverse exists and is partitioned so that $\Phi(t)^{-1} = (L(t), R(t))$ where $L(t)$ and $R(t)$ are $2m \times m$ matrices for each $t \in T$, and either
i) λ is the zero element in $\Re^{2m}, u(t) = 0$ a.e. on T, and $\Phi(t_f)\Phi(t_o)^{-1}y_o = y_f$, or else
ii) λ is nonzero, the equation $\mid q(t) \mid = 0$ has at most finitely many solutions on $T, y(t_f) = y_f$, and

$$u(t) = -\frac{q(t)}{\mid q(t) \mid}f(t) \tag{25}$$

a.e. on T where either the equation $\mid q(t) \mid = \ell_o/\beta$ has at most finitely many solutions on T and

$$f(t) = \begin{cases} 0, & \mid q(t) \mid < \ \ell_o/\beta \\ 1, & \mid q(t) \mid > \ \ell_o/\beta \end{cases} \tag{26}$$

a.e. on T, or else it is satisfied identically on T and $0 \leq f(t) \leq 1$ a.e. on T. Moreover $\lambda^T(\Phi(t)^{-1}y(t) - \Phi(t_o)^{-1}y_o) = -\beta\int_{t_0}^{t} \mid q(\tau) \mid f(\tau)d\tau \leq 0 \quad (t \in T)$. In particular, $\lambda^T(\Phi(t_f)^{-1}y_f - \Phi(t_o)^{-1}y_o) \leq 0$ and equality holds if and only if $f(t) = 0$ a.e. on T.

Remark 1: The final condition establishes a geometric restriction on the vector $\lambda \in \Re^{2m}$. If we define the function $z : T \longrightarrow \Re^{2m}$ by $z(t) = \Phi(t)^{-1}y(t) - \Phi(t_o)y_o$ then λ is restricted by the boundary condition $\lambda^T z(t_f) \leq 0$. Moreover λ is restricted by the whole trajectory, i.e. $\lambda^T z(t) \leq 0$ $(t \in T)$ and, upon differentiating, we obtain $\lambda^T\dot{z}(t) = - \mid q(t) \mid f(t)$ a.e. on T. This shows that λ is further restricted by the tangent vector to $z(t)$ i.e. $\lambda^T\dot{z}(t) \leq 0$ a.e. on T. Finally we observe that $\lambda^T z(t)$ is monotone decreasing on T. All of these conditions place restrictions on the shape and location of an optimal trajectory.

Remark 2: If the real number ℓ_o whose existence is asserted by this theorem is zero, then *i)* cannot hold; therefore (25, 26) of *ii)* must hold where (26) requires that $f(t) = 1$ a.e. on T. A solution where $\ell_o = 0$ is called an *abnormal* solution.

On the other hand, a solution in which $\ell_o > 0$ is called a *normal* solution. For normal solutions the constant l_o/β can be absorbed by the function q [22]. Effectively, this is equivalent to setting $\ell_o = \beta$ in the preceeding theorem. For normal solutions the equation $\mid q(t) \mid = 1$ either has at most finitely many solutions on T and (26) is replaced by

$$f(t) = \begin{cases} 0, & \mid q(t) \mid < \ 1 \\ 1, & \mid q(t) \mid > \ 1 \end{cases} \tag{27}$$

a.e. on T, or else

$$\mid q(t) \mid = 1 \quad (t \in T) \tag{28}$$

and f is any real valued measurable function satisfying

$$0 \leq f(t) \leq 1 \tag{29}$$

a.e. on T such that (24, 25, 20) are satisfied.

3.3 Singular Solutions

We shall say that a normal solution of the problem defined by (18 - 21) is *singular* if (28) holds, otherwise we say that it is *nonsingular*. Since (28) cannot hold for abnormal solutions, we shall also call them nonsingular. It follows from the preceeding theorem that (27) must hold for normal nonsingular solutions or else $u(t) = 0$ a.e. on T. For this reason and the continuity of q we can say that a normal nonsingular solution consists of *thrusting* intervals and *coasting* intervals separated by points on T called switches. It is sometimes convenient to refer to the *switching function* as the real valued function s on T defined by

$$s(t) = \mid q(t) \mid -1. \tag{30}$$

The thrusting intervals are determined by the condition that the switching function is positive; similarly a negative value determines a coasting interval. A *switch* is a zero of the function s such that every open interval containing it also contains points where s is positive and points where s is negative. The preceeding theorem asserts that there are at most finitely many switches. The number of switches for a given interval T is not known, even for the linearized problem of a spacecraft near a circular orbit [28]. Since Theorem 3.1 shows that it is impossible to have $s(t) = 0$ on a subset of T of positive measure and also $s(t) \neq 0$ on another subset of T of positive measure, we do not need to define singular solutions on a subset of T. This property is lost, however, if the differential equation (18) is replaced by certain nonlinear differential equations such as those representing spacecraft trajectories in a Newtonian gravitational field, as discovered by Robbins [30] who found singular and nonsingular regimes on the same interval. If a solution is singular then (29) is valid a.e. on T, but the form of the function f is not specified. For this reason one might suspect that singular solutions to a boundary value problem are not unique. We shall show that this is the case for a certain large class of singular solutions. A singular solution is called an *intermediate thrust solution* if there exists a measurable subset S of T of positive measure such that

$$0 < f(t) < 1 \tag{31}$$

for each $t \in S$. We shall show that the intermediate thrust solutions are degenerate in the sense of the following theorem. This type of degeneracy was first discovered by La Salle [31] in the problem of time optimal control of linear systems. The theorem generalizes previous results based on linearized equations about a satellite in circular orbit [32]. An earlier observation of the degeneracy of singular solutions near a circular orbit was made by Marec [19].

Theorem 3.2 *Suppose that the assumptions of Theorem 3.1 are satisfied. If the problem of minimizing* $J[u] = \int_{t_o}^{t_f} \mid u(t) \mid dt$ *subject to (18 - 20) has an intermediate thrust solution, then it has infinitely many intermediate thrust solutions.*
The proof can be found in [22].

3.4 The Two Point Boundary Value Problem

In order to apply Theorem 3.1 one must solve a two-point boundary-value problem. A value of the element $\lambda \in \Re^{2m}$ must be found such that the trajectory (24) satisfies the terminal condition (20). Specifically λ defines the primer vector function q, which for normal nonsingular solutions defines the control function u through (25,27), establishing the trajectory through (24). We see therefore that the terminus $y(t_f)$ of the trajectory is a function of $\lambda \in \Re^{2m}$. The terminal condition (20) can therefore be viewed as the solution of $2m$ nonlinear equations in the $2m$ components of λ.

Newton's method is a well known computational method for the solution of several nonlinear equations in several variables, but reliable numerical convergence to a root requires knowledge of the approximate location of the root. We use a method of approximating the location of the root through the solution of the related impulsive problem [21]. The related impulsive problem provides a very good approximation if the number β in (18) is sufficiently large. The mathematical justification that the unbounded thrust problem can be viewed as the limit of the bounded thrust problem as β tends to infinity is due to Neustadt [24]. The remarkable accuracy of the root for large values of β is also partly due to the fact that the form of the primer vector function is found to be identical for both problems! Compare q in Theorem 3.1 with its corresponding definition in section 2.2.1 Having solved the boundary-value problem for a large value of β, one can then successively lower the value of β, solving the boundary value problem at each step, until the solution is reached for the designated value of β. The whole sequence can usually be performed with only a microcomputer. The effectiveness of approximating a bounded thrust problem by an impulsive one and then lowering the bound in steps to solve the boundary value problem was demonstrated by Handelsman in 1966 [33]. Starting iteratives for solution of the linear bounded thrust minimization problem can be obtained from solution of the related linear impulsive problem. A very attractive short cut is available, however, for many of the most useful problems in rendezvous and orbital maneuvers of spacecraft. For many problems in which $t_f - t_o, \mid v_o \mid$, and $\mid v_f \mid$ are not too large, nonsingular bounded thrust solutions consist of at most two short duration thrusting intervals separated by a relatively large coasting interval. Starting iteratives are obtained for these problems from the related impulsive problem having at most two impulses at the end points, an initial increment $\Delta v_o \in \Re^m$ at t_o and a terminal increment $\Delta v_f \in \Re^m$ at t_f. Computation of Δv_o and Δv_f requires only the solution of the following set of $2m$ linear equations as determined by (14) and in [21]

$$R(t_o)\Delta v_o + R(t_f)\Delta v_f = \Phi(t_f)^{-1} y_f - \Phi(t_o)^{-1} y_o. \tag{32}$$

3.4.1 Algorithm for Computation of Optimal Trajectories Having Two Short Thrusting Intervals

We now outline an algorithm for computation of certain two-point boundary-value problems associated with the solution of the optimal linear bounded thrust problem. The method is applicable to problems in which y_o, y_f, and T satisfying the following:

1. A *nonsingular* solution over the interval T exists connecting the points y_o and y_f and these points are not sufficiently close to points in which the only solutions are singular. It has been shown that, in some situations, the only solutions connecting y_o and y_f are singular [32], and in these situations the algorithm is not applicable. Since the definition of q in Section 2.2.1 and (28) show that singular solutions are characterized by

$$\lambda^T R(t) R(t)^T \lambda = 1 \quad (t \in T) \tag{33}$$

   it is frequently possible to determine which elements $\lambda \in \Re^{2m}$ determine singular solutions and avoid those cases.

2. The magnitudes $\mid v_o \mid, \mid v_f \mid$, and $t_f - t_o$ are small enough that nonsingular solutions consist of relatively short initial and terminal thrusting intervals separated by a relatively long coasting interval. It has been shown for the linear problem of optimal maneuvers near circular orbit that given any nontrivial interval T there are boundary values where more than two thrusting intervals occur; also more can occur for large values of $t_f - t_o$ [8]. It is conjectured for that problem that no more than four thrusting intervals can occur during one period of the circular orbit [28]. For problems such as these the algorithm is not applicable, and the boundary-value problem is more difficult.

   The algorithm is outlined below. The process should generally begin using a much larger value of the number β than is given in the problem.

Step 1. The linear system (32) is solved for Δv_o and Δv_f. In some cases the $2m \times 2m$ matrix $(R(t_o), R(t_f))$ is singular. In these cases the value of t_f should be changed by a small amount so that the resulting matrix is nonsingular.

Step 2. The length of the initial and terminal thrusting intervals Δt_o and Δt_f are approximated by

$$\Delta t_o = \mid \Delta v_o \mid / \beta, \quad \Delta t_f = \mid \Delta v_f \mid / \beta.$$

The switches t_{so} and t_{sf} that define the coasting interval are

$$t_{so} = t_o + \Delta t_o, \quad t_{sf} = t_f - \Delta t_f.$$

Step 3. The values t_o and t_f are replaced in (32) respectively by t_{so} and t_{sf}, and steps 1 and 2 are repeated. If the magnitude of the difference in sucessive values of t_{so} and t_{sf} satisfies a specified tolerance we proceed to step 4, otherwise steps 1 and 2 are repeated. If the tolerance is not satisfied after a specified number of loops, the value of β can be increased and the process can begin again at step 1.

Step 4. The current values of Δv_o and Δv_f are used to calculate the primer vector at the current values of t_{so} and t_{sf} respectively.

$$q(t_{so}) = -\Delta v_o / \mid \Delta v_o \mid, \quad q(t_{sf}) = -\Delta v_f / \mid \Delta v_f \mid .$$

These equations hold for the related impulsive problem defined on the interval $[t_{so}, t_{sf}]$ [21].

Step 5. Utilizing the definition of q in Section 2.2.1 we can find the element λ by solving the linear equations

$$(R(t_{so}), R(t_{sf}))^T \lambda = \begin{pmatrix} q(t_{so}) \\ q(t_{sf}) \end{pmatrix}.$$

If the matrix of coefficients is singular a very small adjustment in t_{sf} is made, and the system is then solved for λ.

Step 6. Knowing λ, the primer vector and its derivative are determined. The control function u can be calculated from (25) where $f(t)$ is zero on the interval $[t_{so}, t_{sf}]$ and $f(t) = 1$ otherwise. The vector $y(t_f)$ is obtained through (11). If numerical integration of the differential equations describing y and q is preferred, then the initial values of the primer and its derivative are obtained from

$$q(t_o) = R(t_o)^T \lambda, \quad q'(t_o) = R'(t_o)^T \lambda.$$

Step 7. The above yields $y(t_f)$. If this is sufficiently close to a root of the equation $y(t_f) - y_f = 0$, then Newton's method with iteration on λ can be applied to determine a root. Standard commercial packages are available using shooting methods for numerical determination of the root, as well as performing numerical integration of the differential equations for q and y if this approach is preferred, and solving the systems of linear equations used herein [34]. After determination of the root, the value of β is lowered, and a root is again found using the preceeding solution for starting iteratives. This process is repeated until β is lowered to the correct value.

3.4.2 Spacecraft Maneuvers Near Circular Orbit

We present some of the results for a 90 minute circular satellite orbit ($\omega = 2\pi/3600$ rad./sec.), a spacecraft mass of $3400kg$, and a flight time of 600 sec. ($t_o = 0, t_f = 600$). Although the maximum thrust of the spacecraft is $267N$, we assumed higher values initially, solved the boundary value problem, and successively lowered the thrust to its actual value.

A typical example of the use of this algorithm is presented in figures 1 to 8. The transition from near impulsive to a low bounded thrust is shown graphically on the switching function curves where the portion above zero corresponds to thrusting and the portion of the curve below zero corresponds to coasting in figures 1 to 4. The trajectories required for fuel optimal rendezvous corresponding to these different thrust values are presented in figures 5 to 8. This particular calculation can be started at either $4000N$ thrust where the maneuver requires thrusting for only 3% of the time which closely approximates the impulsive solution, or at $1000N$ where the maneuver requires thrusting for 11% of the time. In either case the estimates of the starting values of Lawden's primer vectors are close enough to obtain a converged solution. The thrust is then systematically lowered after each converged solution until a thrust of $267N$ is reached that requires thrusting for 65% of the maneuvering time.

In general the starting primer vectors change very slowly as the maximum thrust is changed as long as the magnitude of thrust is large enough for the thrusting time to be less than 10% of the flight time. When the impulsive calculation predicts thrusting times greater than 20% of the flight time, then the changes in the magnitude of thrust between converged solutions produce rapidly increasing changes in the starting primer vectors. As the thrusting time approaches 50% of the flight time, progress becomes difficult unless the thrust magnitude is limited to small changes between converged solutions.

The calculations were made on an IBM PS/2 Model 50 computer which is on a Token Ring Network with IBM Personal Computer DOS Version 3.30. The programs are written and compiled with Microsoft(R) Fortran Optimizing Compiler Version 4.10. The calculations are in single precision.

4 A Fundamental Matrix Solution for Spacecraft Maneuvers Near a General Keplerian Orbit

4.1 Lawden's Integral and Some Previous Work for Noncircular Keplerian Orbits

As early as 1954 Lawden [35] introduced a change of the variable from time to the true anamaly in describing the equations of motion of a rocket under the action of an inverse square law of force, and in order to evaluate his *primer vector* during a null thrust (Keplerian) interval, he introduced the integral

$$I(\theta) = \int_{\theta_0}^{\theta} \frac{d\xi}{(1 + e\cos\xi)^2 \sin^2 \xi} \tag{34}$$

where θ denotes the true anamaly at time t and θ_0 denotes its initial value at time t_0. In 1963 this integral appears again in his book [25] in connection with the solution of his equations (5.30) - (5.32) which determine the transformed primer vector along a Keplerian interval.

These same equations, with only minor modifications, appeared at approximately the same time in the work of De Vries [9] to describe the relative motion of two nearby points in elliptical orbits. Except for changes in the coordinate systems, a nonhomogeneous version of these equations was used by Tschauner and Hempel [10] to describe the rendezvous of a spacecraft with a target in elliptical orbit. These equations were used again by Tschauner [14] and were investigated through a change of variable from the true anamaly to the eccentric anamaly. Restricted or generalized forms of these equations were employed by Shulman and Scott [11], Euler and Shulman [13], and Euler [12] for rendezvous of a spacecraft with an object in elliptical orbit. The approach of Tschauner and Hempel, as used by Weiss [15] in 1981 was found to be effective in constructing two-impulse solutions to rendezvous problems involving objects in elliptical orbits of high eccentricity [16]. In all these studies, the solution of Lawden's equations was not investigated through the use of the integral $I(\theta)$ of Eq. (34).

Eckel's paper [36] of 1982 returns to the integral $I(\theta)$ with Lawden's equations to determine the primer vector and solve the problem of optimal impulsive transfer between noncoplanar elliptical orbits. Recently Carter and Humi [17], and Carter [29] used this integral in solving Lawden's equations to determine both the primer vector and the structure of an optimal rendezvous of a spacecraft with an object near a point in general Keplerian orbit.

The previously mentioned studies demonstrate a variety of applications and interpretations of Lawden's equations. They also indicate considerable variation in the form of the solution of these equations.

We prefer those forms which include the integral $I(\theta)$ because it appears naturally in the most straightforward approaches to solving Lawden's differential equation. This integral is singular, however, at values where the true anamaly is a multiple of π. Even though these are removable singularities, they may lead to computational instabilities in the solution. Especially bothersome is the fact that computational problems can occur where the true anamaly is near zero. These problems are avoided in the work of Tschauner and Hempel [10,14] and others [11-13, 15] who use a form of solution that does not involve $I(\theta)$, but their work is confined to elliptical orbits.

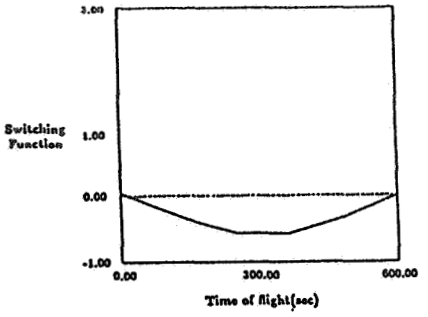


Figure 1. Switching Function vs. Flight Time for Thrust of 4000N, $X_1(0) = 0, X_2(0) = 1000m, X_1(600) = 0, X_3(600) = 0.$

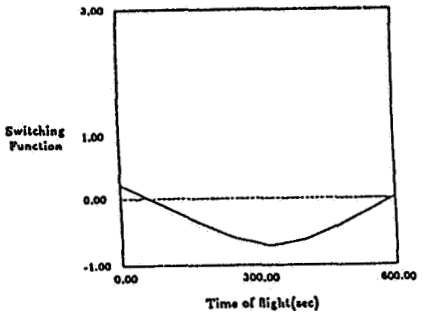


Figure 2. Switching Function vs. Flight Time for Thrust of 1000N, $X_1(0) = 0, X_2(0) = 1000m, X_1(600) = 0, X_3(600) = 0.$

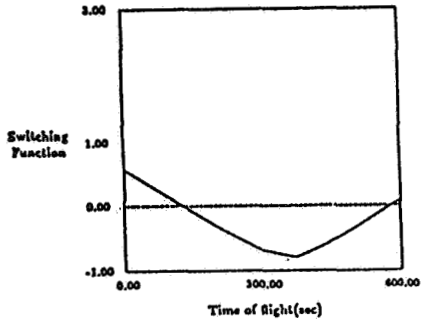


Figure 3. Switching Function vs. Flight Time for Thrust of 500N, $X_1(0) = 0, X_2(0) = 1000m, X_1(600) = 0, X_3(600) = 0.$

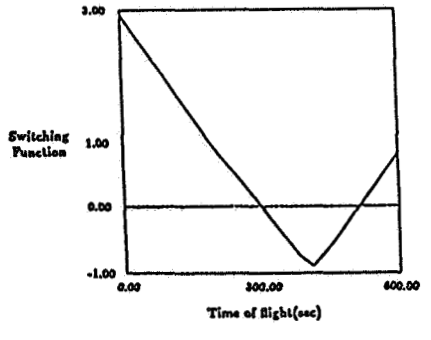


Figure 4. Switching Function vs. Flight Time for Thrust of 267N, $X_1(0) = 0, X_2(0) = 1000m, X_1(600) = 0, X_3(600) = 0.$

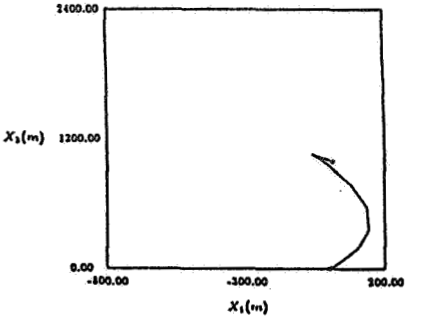


Figure 5. Spacecraft Flight Path for Thrust of 4000N, $X_1(0) = 0, X_2(0) = 1000m, X_1(600) = 0, X_3(600) = 0.$

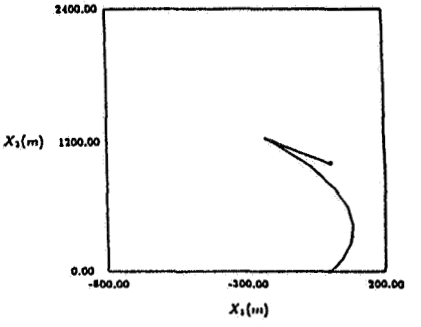


Figure 6. Spacecraft Flight Path for Thrust of 1000N, $X_1(0) = 0, X_2(0) = 1000m, X_1(600) = 0, X_3(600) = 0.$

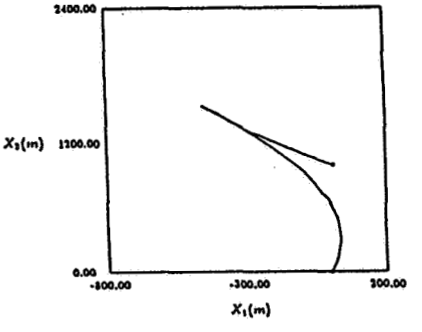


Figure 7. Spacecraft Flight Path for Thrust of 500N, $X_1(0) = 0, X_2(0) = 1000m, X_1(600) = 0, X_3(600) = 0.$

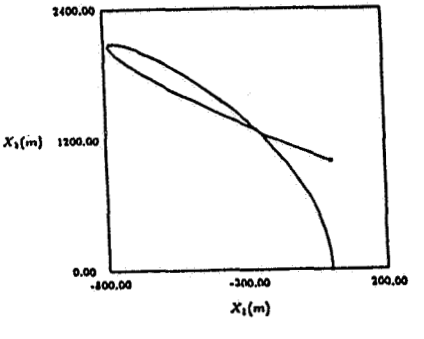


Figure 8. Spacecraft Flight Path for Thrust of 267N, $X_1(0) = 0, X_2(0) = 1000m, X_1(600) = 0, X_3(600) = 0.$

The approach presented in this paper is to modify the original form by replacing $I(\theta)$ by a related integral $J(\theta)$, thereby removing all singularities and computational instabilities. The resulting solution, in terms of $J(\theta)$, is identical for hyperbolic, parabolic, or noncircular elliptic orbits, but the particular case determines the nature of the closed-form evaluation of $J(\theta)$.

4.2 Transformed Equations of a Spacecraft near Keplerian Orbit

We consider a rotating coordinate frame centered at a point moving in a Keplerian orbit about a central attractive body. The positive x_2 axis is directed away from the central body, the positive x_1 axis is perpendicular to it and opposes the direction of the motion, and the x_3 axis completes a right handed coordinate system. We consider the equations of motion of a point mass spacecraft relative to this coordinate system in which the Newtonian gravitational force function has been linearized about the point in Keplerian orbit. The independent variable is the true anamaly θ defined on the closed interval $\theta_0 \leq \theta \leq \theta_f$ which we denote by Θ. All vectors are assumed to be elements of three dimensional Euclidean space. The position vector $x(\theta) = (x_1(\theta), x_2(\theta), x_3(\theta))$ of the spacecraft in this coordinate system is transformed to the vector $z(\theta) = (z_1(\theta), z_2(\theta), z_3(\theta))$ by the equation

$$z(\theta) = r(\theta)x(\theta) \tag{35}$$

where

$$r(\theta) = 1 + e\cos(\theta) \tag{36}$$

and e denotes the eccentricity of the Keplerian orbit.

This development which is presented in detail in previous work [17] results in the following transformed equations of the spacecraft

$$z_1''(\theta) = 2z_2'(\theta) + a_1(\theta)$$

$$z_2''(\theta) = \frac{3}{1+e\cos\theta} z_2(\theta) - 2z_1'(\theta) + a_2(\theta)$$

$$z_3''(\theta) = -z_3(\theta) + a_3(\theta) \tag{37}$$

where the prime indicates differentiation with respect to θ, and the vector $a(\theta) = (a_1(\theta), a_2(\theta), a_3(\theta))$ is given by

$$a(\theta) = \beta(\theta)u(\theta) \tag{38}$$

where

$$\beta(\theta) = k/r(\theta)^3. \tag{39}$$

In this expression the positive constant k is $L^6 T_m/\mu^4 m$ where L is the magnitude of the constant angular momentum of the object in Keplerian orbit divided by its mass, T_m is the maximum magnitude of the thrust of the spacecraft, μ is the product of the universal gravitational constant and the mass of the central body of attraction, and m is the mass of the spacecraft. The vector $u(\theta) = (u_1(\theta), u_2(\theta), u_3(\theta))$ represents the normalized thrust of the spacecraft and is subject to the constraint

$$| u(\theta) | \leq 1. \tag{40}$$

Since this investigation is restricted to linear equations, we shall assume that the mass m is constant. Previous investigations, however, have taken into account the change in mass of the spacecraft as propellant is consumed [29,18].

Equations (37) are essentially the equations of Tschauner and Hempel [10], and their homogeneous form represents essentially the equations of De Vries [9] and Lawden [25].

Here the class of admissible control functions is the set of all Lebesque measurable vector valued functions that satisfy (40) a.e. on Θ. The optimal rendezvous problem associated with a point in Keplerian orbit is defined as the determination of an admissible control function u that minimizes the cost function

$$J[u] = \int_{\theta_0}^{\theta_f} \frac{| u(\theta) |}{r(\theta)^2} d\theta \tag{41}$$

subject to the conditions (37-40) which are valid a.e. on Θ and the end conditions

$$z(\theta_0) = z_0, z'(\theta_0) = v_0,$$

$$z(\theta_f) = z_f, z'(\theta_f) = v_f \tag{42}$$

where the vectors z_0 and z_f define the initial and terminal values of the transformed position z, and v_0 and v_f define the initial and terminal values of its derivative.

Theorem 3.1 is not directly applicable to this problem because (38) and (4.2) do not conform to (18) and (21) respectively. The theorem can be generalized to include this problem, however, and the results are very similar if the primer vector $q(\theta)$ is replaced by the transformed primer $Q(\theta) = q(\theta)/r(\theta)$ as in previous work. [17,29]

If we define the state vector function $y : \Theta \rightarrow \Re^6$ by $y(\theta) = (z(\theta)^T, z\prime(\theta)^T)^T$, then equations (37-39) can be put in a form similar to (18) where $\beta(\theta)$ replaces β and

$$A(\theta) = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 3/r(\theta) & 0 & -2 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}. \tag{43}$$

Here θ replaces t and Θ replaces T. By defining $y_0 = (z_0^T, v_0^T)^T$ and $y_f = (z_f^T, v_f^T)^T$ the boundary conditions (41) satisfy (19, 20).

In order to obtain a fundamental matrix solution $\Phi(\theta)$ associated with this problem, we solve the homogeneous form of the system (37), that is, we find the complete solution where $a(\theta)$ is identically zero. This solution was found by Lawden [35] in terms of the integral $I(\theta)$. Using a similar form of solution [17], we obtain

$$\begin{aligned} z_1(\theta) &= -b_1 r(\theta)^2 - b_2[r(\theta)^2 I(\theta) + cot\theta] - b_3 sin\theta[1 + r(\theta)] + b_4 \\ z_2(\theta) &= er(\theta)sin\theta[b_1 + b_2 I(\theta)] - b_3 r(\theta)cos\theta \\ z_3(\theta) &= b_5 cos\theta + b_6 sin\theta \end{aligned} \tag{44}$$

where $b_1, b_2, b_3, b_4, b_5, b_6$ are arbitrary constants of integration.

These equations can be used to define the relative motion of two nearby points in noncircular Keplerian orbits by transforming from $z(\theta)$ to $x(\theta)$ using (35), and is a generalization of the work of De Vries. [9].

4.3 New Form of the Rendezvous Equations and the Resulting Fundamental Matrix Solution

The Equation (43) has removable singularities at integer multiples of π which we denote by $n\pi$. These singularities appear in the expressions $I(\theta)$ and $cot\theta$. Computation is troublesome at or near these singularities. These computational instabilities can be removed. If we integrate (34) by parts, it becomes

$$I(\theta) = 2eJ(\theta) - \frac{cot\theta}{r(\theta)^2} + c \tag{45}$$

which holds except at the singularities of $I(\theta)$ where c is a constant of integration and

$$J(\theta) = \int_{\theta_0}^{\theta} \frac{\cos\xi}{r(\xi)^3} d\xi. \tag{46}$$

It is observed that the integral $J(\theta)$ has no singularities. It follows from the continuity of $J(\theta)$ that

$$2eJ(n\pi) + c = \lim_{\theta \rightarrow n\pi} \left[I(\theta) + \frac{\cot\theta}{r(\theta)^2} \right] \tag{47}$$

so that the singularities in (43) are removed. It follows that there are no singularities if solutions of (37) are stated in terms of $J(\theta)$ instead of $I(\theta)$. [18] With these changes the solution (43) of the unpowered equations of motion becomes

$$\begin{aligned} z_1(\theta) &= -r(\theta)^2[b_1 + 2b_2 eJ(\theta)] - b_3[1 + r(\theta)]sin\theta + b_4 \\ z_2(\theta) &= r(\theta)sin\theta[b_1 e + 2b_2 e^2 J(\theta)] - cos\theta[\frac{b_2 e}{r(\theta)} + b_3 r(\theta)] \\ z_3(\theta) &= b_5 cos\theta + b_6 sin\theta. \end{aligned} \tag{48}$$

These equations are valid for all noncircular Keplerian orbits. The integral $J(\theta)$ can be evaluated in closed form, and the particular form is determined from the type of orbit. Equations (47) are more useful than equations (43). If we differentiate (47) we obtain the vector-valued function $z' : \Theta \rightarrow \Re^3$. The pair $(z(\theta)^T, z\prime(\theta)^T)^T$ defines the state vector $y(\theta)$. The equation $z(\theta)$ from (47) and its derivative $z'(\theta)$ together are represented by

$$y(\theta) = \Phi(\theta)b \quad (\theta \epsilon \Theta) \tag{49}$$

where $b = (b_1, b_2, b_3, b_4, b_5, b_6)^T$ is a constant vector in $\Re^6$ and $\Phi(\theta)$ is a fundamental matrix solution associated with $A(\theta)$. It follows from (47) that

$$\Phi(\theta) = \begin{bmatrix} -r(\theta)^2 & -2r(\theta)^2 J(\theta) & -(1+r(\theta))\sin\theta & 1 & 0 & 0 \\ er(\theta)\sin\theta & 2er(\theta)\sin\theta J(\theta) - \cos\theta/r(\theta) & -r(\theta)\cos\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cos\theta & \sin\theta \\ 2er(\theta)\sin\theta & 4er(\theta)\sin\theta J(\theta) - 2\cos\theta/r(\theta) & e - 2r(\theta)\cos\theta & 0 & 0 & 0 \\ er(\theta)\cos\theta - e^2\sin^2\theta & 2e(r(\theta)\cos\theta - e\sin^2\theta)J(\theta) + \frac{\sin\theta(1+2e\cos\theta)}{r(\theta)^2} & (r(\theta)+e\cos\theta)\sin\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\sin\theta & \cos\theta \end{bmatrix}. \tag{50}$$

Our goal has been to determine this matrix function. If the problem is restricted to a plane, we delete the third and sixth row and the fifth and sixth column. The resulting 4×4 matrix function is a fundamental matrix solution for the planer problem. There are several approaches[3] to the problem of inverting a fundamental matrix solution and applying a generalization of the preceeding theory. These should result in new methods of computing either impulsive or bounded thrust trajectories of a spacecraft near Keplerian orbit. Further results in this area are forthcoming.

4.4 Closed-Form Evaluation of The New Integral

We show here that the integral $J(\theta)$ can be evaluated easily if we transform from the true anamaly θ to the eccentric anamaly E for elliptical orbits or its analog H for hyperbolic orbits.

4.4.1 Elliptical Orbits

For orbits in which $0 < e < 1$ we have the relationship between the eccentric anomaly E and true anomaly θ given by

$$\cos\theta = \frac{cosE - e}{1 - e\cos E} \tag{51}$$

where $\sin\theta$ and $\sin E$ always have the same algebraic sign. Changing the variable to E in (45) establishes the much simpler integral

$$J(\theta) = (1-e^2)^{-5/2}\int_{E_0}^{E}(1 - ecos\xi)(\cos\xi - e)d\xi \tag{52}$$

where E_0 is the eccentric anomaly at θ_0. This integral is easily evaluated using elementary methods to obtain

$$J(\theta) = -(1-e^2)^{-5/2}\left[\frac{3e}{2}E - (1+e^2)\sin E + \frac{e}{2}\sin E\cos E + C\right] \tag{53}$$

where C is an arbitrary constant.

4.4.2 Hyperbolic Orbits

In a similar way we can evaluate $J(\theta)$ for orbits in which $e > 1$. We introduce the analog of the eccentric anomaly H by the relationship

$$\cos\theta = \frac{e - \cosh H}{e\cosh H - 1} \tag{54}$$

where $\sin\theta$ and $\sinh H$ always have the same algebraic sign. With this substitution the integral defined by (45) becomes

$$J(\theta) = (e^2-1)^{-5/2}\int_{H_0}^{H}(e\cosh\xi - 1)(e - \cosh\xi)d\xi \tag{55}$$

where H_0 denotes the value of H at θ_0. This integral also is easily evaluated. The result is

$$J(\theta) = -(e^2-1)^{-5/2}\left[\frac{3eH}{2} - (1+e^2)\sinh H + \frac{e\sinh H}{2}\cosh H + C\right] \tag{56}$$

where again C denotes an arbitrary constant.

4.4.3 Parabolic Orbits

For the case in which $e = 1$, the integral $J(\theta)$ can be evaluated directly using the identity $\cos\theta = 2\cos^2\frac{\theta}{2} - 1$. The result is

$$J(\theta) = \frac{1}{4}tan\frac{\theta}{2} - \frac{1}{20}tan^5\frac{\theta}{2} + C \tag{57}$$

where C is again an arbitrary constant. Since this expression is only defined on the region $r(\theta) > 0 (i.e. \cos\theta > -1)$, it has no singularities.

[3]This matrix was inverted by Prof. Mayer Humi of the Mathematial Sciences Department at Worchester Polytechnic Institute using MAXYSMA. The author has also inverted this matrix using the adjoint system to that defined by (42). Numerical inversion is also feasible.

References

[1] Wheelon, A.D.,"Midcourse and Terminal Guidance," *Space Technology*, Wiley, New York, 1959.

[2] Clohessy, W.H., and Wiltshire, R.S., "Terminal Guidance System for Satellite Rendezvous," *Journal of the Aerospace Sciences*, Sept. 1960, pp. 653-658, 674.

[3] Geyling, F.T., "Satellite Perturbations from Extra-Terrestrial Gravitation and Radiation Pressure," *Journal of the Franklin Institute*, Vol. 269, 1960, pp. 375-407.

[4] Spradlin, L.W., "The Long-Time Satellite Rendezvous Trajectory," *Aerospace Engineering*, Vol. 19, 1960, pp. 32-37.

[5] Tschauner, J., and Hempel, P., "Optimale Beschleunigungsprogramme für das Rendezvous-Manöver," *Astronautica Acta*, Vol. 10, pp. 296-307, 1964.

[6] Prussing, J.E., "Optimal Four-Impulse Fixed Time Rendezvous in the Vicinity of a Circular Orbit," *AIAA Journal*, Vol. 7, pp. 928-935, 1969.

[7] Prussing, J.E., "Optimal Two-and Three-Impulse Fixed-Time Rendezvous in the Vicinity of a Circular Orbit," *AIAA Journal*, Vol. 8, pp. 1221-1228, 1970.

[8] Carter, T., "Fuel-Optimal Maneuvers of a Spacecraft Relative to a Point in Circular Orbit," *Journal of Guidance, Control, and Dynamics*, Nov./Dec., 1984, pp. 710-716.

[9] De Vries, J.P., "Elliptic Elements in Terms of Small Increments of Position and Velocity Components", *AIAA Journal*, Vol. 1, pp. 2626-2629, 1963.

[10] Tschauner, J., and Hempel, P., "Rendezvous zu einemin Elliptischer Bahn Umlaufenden Ziel", *Astronautica Acta*, Vol. 11, pp. 104-109, 1965.

[11] Shulman, Y., and Scott, J., "Terminal Rendezvous for Elliptical Orbits", AIAA Paper No. 66-533, AIAA Fourth Aerospace Sciences Meeting, Los Angeles, June, 1966.

[12] Euler, E.A., "Optimal Low-Thrust Rendezvous Control," *AIAA Journal*, Vol. 7, pp. 1140-1144, 1969.

[13] Euler, E., and Shulman, Y., "Second Order Solution to the Elliptical Rendezvous Problem," *AIAA Journal*, Vol. 5, 1967, pp. 1033-1035.

[14] Tschauner, J., "Elliptic Orbit Rendezvous," *AIAA Journal*, Vol. 5, 1967, pp. 1110-1113.

[15] Weiss, J., "Solution of the Equation of Motion for High Elliptic Orbits," TN PRV-5 No. 7/81, ERNO Reumfahrttechnik, Bremen, Nov., 1981.

[16] Wolfsberger, W., Weiss, J., and Rangnitt, D., "Strategies and Schemes for Rendezvous on Geostationary Transfer Orbit," *Acta Astronautica*, Vol. 10, 1983, pp. 527-538.

[17] Carter, T., and Humi, M., "Fuel-Optimal Rendezvous Near a Point in General Keplerian Orbit", *Journal of Guidance, Control, and Dynamics*, Nov./Dec., 1987, pp. 567-573.

[18] Carter, T., "New Form for the Optimal Rendezvous Equations Near a Keplerian Orbit," *Journal of Guidance, Control, and Dynamics*, Jan./Feb., 1990.

[19] Marec, J.P., *Optimal Space Trajectories*, Elsevier, New York, 1979.

[20] Szebehely, V., *Theory of Orbits*, Academic Press, New York, 1967.

[21] Carter, T., "Optimal Impulsive Space Trajectories Based on Linear Equations," Accepted for publication in the *Journal of Optimization Theory and Applications.*

[22] Carter, T. and Brient, J., "Optimal Bounded Thrust Space Trajectories Based on Linear Equations," Submitted for publication.

[23] Neustadt, L.W., "Optimization, A Moment Problem, and Nonlinear Programming," *SIAM Journal of Control*, Vol. 2, 1964, pp. 33-53.

[24] Neustadt, L.W., "A General Theory of Minimum-Fuel Space Trajectories," *SIAM Journal of Control*, Vol. 3, 1965, pp. 317-356.

[25] Lawden, D.F., *Optimal Trajectories for Space Navigation*, Butterworths, London, 1963.

[26] Lion, P.M., "A Primer on the Primer", *STAR Memo No. 1*, Department of Aerospace and Mechanical Sciences, Princeton University, April 21, 1967.

[27] Lion, P.M., and Handelsman, M., "Primer Vector on Fixed-Time Impulsive Trajectories," *AIAA Journal*, Vol. 6, No. 1, Jan. 1968.

[28] Carter, T., "How Many Intersections Can a Helical Curve Have With the Unit Sphere During One Period?" Unsolved Problems Section, *The American Mathematical Monthly*, Vol. 93, pp. 41-44, Jan., 1986.

[29] Carter, T., "Effects of Propellant Mass Loss on Fuel-Optimal Rendezvous Near Keplerian Orbit", *Journal of Guidance, Control, and Dynamics*, Jan./Feb., 1989.

[30] Robbins, H.M., "Optimal Rocket Trajectories with Subarcs of Intermediate Thrust," 17th International Astronautical Congress-Madrid, Oct., 1966.

[31] LaSalle, J.P., "The Time Optimal Control Problem", *Contributions to the Theory of Nonlinear Oscillations*, Vol. 5, pp. 1-24, Princeton University Press, Princeton, N.J., 1960.

[32] Carter, T., "Singular Fuel-Optimal Space Trajectories Based on a Linearization About a Point in Circular Orbit," *Journal of Optimization Theory and Applications*, Vol. 54, No. 3, Sept., 1987, pp. 447-470.

[33] Handelsman, M., "Optimal Free-Space Fixed-Thrust Trajectories Using Impulsive Trajectories as Starting Iteratives", *AIAA Journal*, Vol. 4, June 1966, pp. 1077-1082.

[34] Press, W.H., et al, *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press, 1986.

[35] Lawden, D.F., "Fundamentals of Space Navigation", *Journal of the British Interplanetary Society*, Vol. 13, pp. 87-101, 1954

[36] Eckel, K.G., "Optimal Impulsive Transfer with Time Constraint", *Astronautica Acta*, Vol. 9, pp. 139-146, 1982

NASA National Aeronautics and Space Administration

Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CP-3102		

4. Title and Subtitle	5. Report Date
Flight Mechanics/Estimation Theory Symposium - 1990	December 1990
	6. Performing Organization Code: 554

7. Author(s)	8. Performing Organization Report No.
Thomas Stengle, Editor	91B00018
	10. Work Unit No.

9. Performing Organization Name and Address	11. Contract or Grant No.
Flight Dynamics Division Goddard Space Flight Center Greenbelt, Maryland 20771	
	13. Type of Report and Period Covered: Conference Publication
12. Sponsoring Agency Name and Address	14. Sponsoring Agency Code
National Aeronautics and Space Administration Washington, D. C. 20546-0001	

15. Supplementary Notes

Thomas Stengle is Head, Attitude Analysis Section, Flight Dynamics Analysis Branch at the Goddard Space Flight Center, Greenbelt, Maryland.

16. Abstract

This conference publication includes 32 papers and abstracts presented at the Flight Mechanics/Estimation Theory Symposium on May 22-24, 1990. Sponsored by the Flight Dynamics Division of Goddard Space Flight Center, this symposium features technical papers on a wide range of issues related to orbit-attitude prediction, determination and control; attitude sensor calibration; attitude determination error analysis; attitude dynamics; and orbit decay and maneuver strategy. Government, industry, and the academic community participated in the preparation and presentation of these papers.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Flight Mechanics Estimation Theory Attitude Determination Mission Analysis Spacecraft Dynamics Orbit Determination	Unclassified-Unlimited Subject Category 13

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of pages	22. Price
Unclassified	Unclassified	568	A24

NASA FORM 1626 OCT 86

END DATE APRIL 5, 1990

National Aeronautics and
Space Administration
Code NTT-4

Washington, D.C.
20546-0001

Official Business
Penalty for Private Use, $300

NASA

National Aeronautics and
Space Administration

Washington, D.C.
20546

SPECIAL FOURTH CLASS MAIL
BOOK

Postage an Fees Paid
National Aeronautics and
Space Administration
NASA-451

Official Business
Penalty for Private Use $300

U.S. MAIL

L1 001 CP-3102 91010452090569A
NASA
SCIEN & TECH INFO FACILITY
ACCESSIONING DEPT
P O BOX 8757 BWI ARPRT
BALTIMORE MD 21240

NASA

POSTMASTER: If Undeliverable (Section 158
Postal Manual) Do Not Return