NASA TECHNICAL NOTE

NASA TN D-3382

NASA TN D-3382

EFFECT OF REDUCED COMPUTER PRECISION ON A MIDCOURSE NAVIGATION AND GUIDANCE SYSTEM USING OPTIMAL FILTERING AND LINEAR PREDICTION

by Leonard A. McGee

Ames Research Center

Moffett Field, Calif.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • APRIL 1966

NASA TN D-3382

EFFECT OF REDUCED COMPUTER PRECISION ON A MIDCOURSE NAVIGATION AND GUIDANCE SYSTEM USING OPTIMAL FILTERING AND LINEAR PREDICTION

By Leonard A. McGee

Ames Research Center
Moffett Field, Calif.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – Price $0.60

EFFECT OF REDUCED COMPUTER PRECISION ON A MIDCOURSE NAVIGATION AND GUIDANCE SYSTEM USING OPTIMAL FILTERING AND LINEAR PREDICTION

By Leonard A. McGee
Ames Research Center

SUMMARY

Reduced precision, defined as the use of computer words whose precision is less than the maximum single precision length of 27 binary bits, is applied to the lunar midcourse navigation and guidance scheme previously reported in NASA publications TR R-135 and TN D-1208. With the same computer program, the evaluation in this report was based on the outbound leg of an earth-moon trajectory. A fixed schedule of 45 observations and 3 velocity corrections was executed each time the trajectory was traversed with the precision reduced in one or more of the following areas: (1) the estimated trajectory, (2) the optimal filter, (3) the space-angle computations, (4) the transition matrix, or (5) the guidance equations. This procedure was followed down to the level of precision at which the estimated trajectory could no longer be integrated numerically.

The mathematical operations embodied in the optimal filter are shown to be essentially unaffected by the reduced precision over the range studied. This invariance, as determined by the behavior of the covariance matrix of the guidance errors in the estimated state, P, does not imply that P is a correct representation of the existing uncertainties. In fact, as the precision is reduced, the uncertainties indicated by the P matrix become an increasingly conservative representation of those that actually exist. This is due to the failure of the error propagation model to account for the increased errors in the estimated state that occur during the transition period between observations and velocity corrections. When no corrections are made to account for the increased errors, as in this study, it is shown that the optimal filter will give too low a weight to the information gained from space-angle observations. The resulting errors in the estimate of the vehicle's state may, for large precision reductions, accumulate until they are excessively large. Velocity corrections are more poorly done as the precision is reduced because they are dependent on the estimate of the vehicle's state and on a prediction matrix which projects the errors at any time into errors at the final time. This is shown to be due, at least in part, to the degradation of the prediction matrix which, in the lower precision cases, was affected to the extent that the determinant of one of its submatrices became negative near the end of the trajectory.

Author

INTRODUCTION

References 1 and 2 present the results of a digital computer simulation of the midcourse phase of a circumlunar navigation study. The Kalman-Schmidt optimal filter (refs. 2, 3, and 4) is used to process simulated pilot-observed data (derived from the space geometry angles relating the vehicle to the earth and the moon) in order to produce an optimal estimate of the vehicle's position and velocity. With this information and a transition matrix relating end-point deviations to present deviations and knowledge of the state of a reference trajectory, the system computes and makes velocity corrections in the attempt to guide to a fixed end point at a given time.

In certain situations it is desirable to simulate all or part of the midcourse navigation on a digital computer which uses words of less precision than the 27 binary bits used in the studies reported in references 1 and 2. For an intelligent choice of word length, it is necessary to know the degradation of performance due to decreased accuracy. It is the aim of this report to provide this information and at the same time to determine whether the original studies were done with words of sufficient precision.

SYMBOLS

a	column matrix of deviations in observed angles
A	prediction matrix: a transition matrix relating deviations at the end point to present deviations
A_1, A_2, A_3, A_4	3×3 submatrices of A
F	matrix of partial derivatives of the equations of motion with respect to the state variables
H	matrix relating observed angles and vehicle state
I	identity matrix
K	weighting matrix
$N[\mu,\sigma^2]$	normally distributed with mean μ and variance σ^2
P	covariance matrix of estimation error vector, $E[\tilde{x}\tilde{x}^T]$
P'	value of P after an observation
P_1, P_2, P_3, P_4	3×3 submatrices of P

q	observational error vector
Q	covariance matrix of the observational errors, $E[qq^T]$
r	position deviation from reference
$\tilde{r}$	error in estimate of r
$\hat{r}$	estimate of r
rms	root mean square
$\overline{R}$	position vector (subscripts indicate origin and end)
R	magnitude of $\overline{R}$
R_E	earth radius
R_M	moon radius
t	general time arguments
t_e	end-point time
t_k	time of the kth observation
v	velocity deviation from reference
$\hat{v}$	estimate of v
$\tilde{v}$	error in estimate of v
V_{rms}	rms uncertainty of indicated velocity correction
ΔV_G	velocity-vector increment to be gained
$\lvert\Delta V_G\rvert$	magnitude of ΔV_G
x	six vector of position and velocity deviation from a reference trajectory
$\hat{x}$	estimate of x
$\tilde{x}$	error in estimate of x, $x-\hat{x}$
X,Y,Z	space position coordinates
y	observation of space angles, $Hx + q$
$\hat{y}$	estimate of y; $H\hat{x}$

α	declination of observed body
β	right ascension of observed body
γ	one-half subtended angle of observed body
Δ	increment
σ	standard deviation of subscript random variable
$\Phi(t_2;t_1)$	transition matrix relating state at t_2 to state at t_1

Notation Conventions

$(\dot{})$	first-time derivative of ( )
$( )^T$	transpose of matrix ( )
$( )^{-1}$	inverse of matrix ( )
E[ ]	expected value of [ ]
trace [ ]	sum of the diagonal elements of [ ]

Subscripts

e	end point
E	earth
k	at the kth observation
M	moon
S	sun
V	vehicle

COMPUTER PROGRAM FOR SIMULATION

General Description

The digital computer program used in this study is a modified version of the FORTRAN II program for the midcourse navigation studies reported in references 1 and 2.

The reference mission assumed for this study is the outbound leg of an earth-moon trajectory requiring about 78 hours for completion. Imperfect

injection conditions for the space vehicle are assumed, and it is the function of the navigation and guidance system to cause the space vehicle to arrive at the reference end point with a minimum of position error. To determine position and velocity the navigation system will process observations of the earth and moon made from the vehicle at preselected times. As a result of the knowledge gained from the observations, the system will generate an estimate of the state (position and velocity). On the basis of this estimate, the guidance system will make imperfect corrections to the vehicle velocity in an attempt to cause the vehicle to arrive at the reference end point (perilune).

Twenty-seven second-order ordinary differential equations comprising three trajectories and six sets of three perturbation equations are integrated simultaneously. The integration subprogram utilizes a fourth-order Runge-Kutta method for starting and a Cowell "second-sum" method based on sixth differences for continuation. Double precision is used internally to control round-off error.

The three trajectories are (1) the previously determined reference trajectory, (2) the actual trajectory giving the true state of the space vehicle, and (3) the estimated trajectory giving the current estimate of the vehicle state. Comparison of these three trajectories allows evaluation of changes in the system performance resulting from the reduction of precision in certain specified areas of computation.

Program Areas in Variable Precision

In order to study the effects of reduced precision in the various portions of the computations, five masks were developed which allowed independent control of the precision of each of the portions.

The five levels of precision maintained with the masking technique are in addition to the normal 27-bit precision of the IBM 7094. Mask 1 is used in conjunction with the optimal filter and its weighting matrix. Mask 2 is used to control the precision of the estimated trajectory. Mask 3 controls the precision of the space angles computed from the estimated state and the space-angle geometry. Mask 4 controls the precision of the matrix F, the perturbation equations, and the transition matrix, Φ. Mask 5 controls the precision with which the prediction matrix, A, is updated, the precision of the guidance equation computations, and the computations involved in the velocity correction model used to implement a velocity correction. Full 27-bit precision is used on those calculations that provide information about the performance of the system and on calculations involved with the actual and reference trajectories. In addition, the numerical integration subroutine used to integrate the equations of motion and the perturbation equations carried out internal computations in 54 binary bit precision (double precision) as mentioned before. This was true regardless of the effective mantissa length of the second derivatives computed from the equations of motion and the perturbation equations or the initial conditions. A discussion of the masking operations used to reduce the effective word length is given in appendix A.

Processing an Observation

The digital simulation of the complete system is represented in figure 1, which also shows in dashed boxes those areas which are maintained at the precisions specified by masks 1, 2, 3, 4, and 5. A description of the operation of the system and the equations used is given in the remainder of the section. The circled numbers in the text refer to corresponding numbers in figure 1 and are used to show which of the boxes is being discussed at the moment.

Updating the trajectory.- When an observation has been made at time t_k and is to be processed, numerical integration of the equations of motion is initiated (1) for the actual, estimated, and reference trajectories simultaneously. Initial conditions for the actual and reference trajectories are merely the states which existed at the time of the last observation, while the estimated trajectory uses the new estimated state derived from processing the last observation. Also, simultaneously with (1) the sixfold perturbation equations (2) are integrated numerically from the time of the last observation, t_{k-1}, to produce a transition matrix (3). This process continues until the computer time equals the time of the observation, t_k. The integration is then stopped and the covariance matrices are updated.

Updating of the covariance matrices $P(t_k)$ and $Q(t_k)$.- The updating operation on $P(t_k)$ is accomplished through the use of linear prediction. This method requires the assumption that small deviations of state at time t_k can be obtained from a linear combination of the deviations at time t_{k-1}. In particular, when dealing with the deviation $\tilde{x}(t_{k-1})$, the relationship may be written in matrix form as follows:

$$\tilde{x}(t_k) = \Phi(t_k;t_{k-1})\tilde{x}(t_{k-1}) \tag{1}$$

where $\Phi(t_k;t_{k-1})$ is the transition matrix from time t_{k-1} to time t_k. Substituting equation (1) into the definition

$$P(t_k) = E[\tilde{x}(t_k)\tilde{x}^T(t_k)] \tag{2}$$

yields the following updating equation for $P(t_{k-1})$:

$$P(t_k) = \Phi(t_k;t_{k-1})P(t_{k-1})\Phi^T(t_k;t_{k-1}) \tag{3}$$

The covariance matrix of the observational errors, $Q(t_k)$, is given by

$$Q(t_k) = \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & 0 \\ 0 & 0 & \sigma^2 \end{bmatrix} \tag{4}$$

where $\sigma^2 = 100 + (0.001\gamma)^2$ seconds of arc and γ is 1/2 the subtended angle of the observed body.

Thus the updating operations at ④ consist of computing equation (3) and forming $Q(t)$.

Determination of the weighting matrix $K(t_k)$.- The next step is to compute the H matrix followed by the weighting matrix $K(t_k)$. The matrix $H(t_k)$ is computed ⑤ from the reference state values in the form

$$H(t_k) = \begin{bmatrix} \frac{\partial\alpha}{\partial X} & \frac{\partial\alpha}{\partial Y} & \frac{\partial\alpha}{\partial Z} \\ \frac{\partial\beta}{\partial X} & \frac{\partial\beta}{\partial Y} & \frac{\partial\beta}{\partial Z} \\ \frac{\partial\gamma}{\partial X} & \frac{\partial\gamma}{\partial Y} & \frac{\partial\gamma}{\partial Z} \end{bmatrix} \tag{5}$$

where the partial derivatives for the body being observed are given in table I. The matrix $K(t_k)$ is then computed ⑥ from

$$K(t_k) = P(t_k)H^T(t_k)[H(t_k)P(t_k)H^T(t_k) + \dot{Q}(t_k)]^{-1} \tag{6}$$

With $K(t_k)$ known, the $P'(t_k)$ matrix is then computed ⑦ from

$$P'(t_k) = P(t_k) - K(t_k)H(t_k)P(t_k) \tag{7}$$

reflecting the change in P due to the observation just processed;[1] $P'(t_k)$ is then stored. The delay unit ⑧ represents the storage of $P'(t_k)$ until the time of the next observation.

Two quantities are derived from P' and are used to assess the system's estimation performance. These two quantities, $\tilde{r}_{rms}$ and $\tilde{v}_{rms}$, are the standard deviations of the error in estimating the magnitude of the position and velocity deviations from the reference. They are defined by

$$\tilde{r}_{rms} = [\text{trace } (P_1')]^{1/2} \tag{8}$$

and

$$\tilde{v}_{rms} = [\text{trace } (P_4')]^{1/2} \tag{9}$$

[1]The order of matrix multiplication is from left to right. It has been found that even 27 bits is not sufficient to preserve the non-negative definite property of the P matrix if the order of multiplication is not preserved.

where P_1' and P_4' are 3×3 submatrices of the P' matrix partitioned as follows:

$$P' = \begin{bmatrix} P_1' & P_2' \\ P_3' & P_4' \end{bmatrix} \tag{10}$$

Formation of the new estimated state.- In this system space-angle observations consist of the measurement of three angles related to the body (earth or moon) under observation. These angles are α, the right ascension; β, the declination; and γ, 1/2 of the angle subtended by the body.

The estimated angles are computed (9) from the space geometry equations (table II) using the state of the estimated trajectory. From the same space geometry equations, and using the state of the actual trajectory, the actual angles are computed (10). The difference between these two sets of angles is found and noise, scaled according to $Q(t_k)$, is added. The result (11) multiplied by $K(t_k)$ (12) is added to the estimated state (13) to produce a new estimated state. This completes the estimation process.

Computation of the indicated velocity correction.- In order to provide in-flight velocity correction capability, a prediction matrix, $A(t_e;t_k)$, which predicts the end-point state from the present state is computed (14) from

$$A(t_e;t_k) = A(t_e;t_{k-1})\Phi(t_k;t_{k-1})^{-1} \tag{11}$$

Since the initial prediction matrix, $A(t_e;t_o)$, is an input quantity, it is only necessary to update the matrix at each observation. This requires finding $\Phi(t_k;t_{k-1})^{-1}$ from $\Phi(t_k;t_{k-1})$.

Because $\Phi(t_k;t_{k-1})$ is a symplectic matrix (ref. 5), it is invertable merely by rearranging terms and changing signs. Thus, if $\Phi(t_k;t_{k-1})$ is partitioned into 3×3 submatrices, then

$$\Phi(t_k;t_{k-1}) = \begin{bmatrix} \Phi_1 & \Phi_2 \\ \Phi_3 & \Phi_4 \end{bmatrix} \tag{12}$$

and

$$\Phi(t_k;t_{k-1})^{-1} = \begin{bmatrix} \Phi_4^T & -\Phi_2^T \\ -\Phi_3^T & \Phi_1^T \end{bmatrix} \tag{13}$$

Therefore, there are no errors in this inversion process.

With the difference between the estimated and reference states, $\hat{r}$ and $\hat{v}$, (15) and the updated prediction matrix from equation (11), the indicated velocity correction, ΔV_G, is computed (16) from

$$\Delta V_G = - \left[A_2^{-1} A_1 \;\vdots\; I \right] [\hat{x}] \tag{14}$$

The magnitude of ΔV_G is

$$|\Delta V_G| = (\Delta V_G{}^T \Delta V_G)^{1/2} \tag{15}$$

This concludes the computation cycle done at each observation and the computer is now ready to process the next observation unless a velocity correction is to be executed. In this case the correction is made after a time delay in order to simulate the time required to orient the vehicle along the desired thrust vector. To accomplish this delay, processes (1), (2), and (3) are repeated until the computer time equals the time at which the velocity correction is to be made. At this time processes (4), (14), (15), and (16) are repeated and computation control is passed to a separate program section for the performance of the velocity correction maneuver.

Performing the Velocity Correction Maneuver

The flow lines representing velocity correction increments are shown dashed in figure 1 to emphasize the occasional nature of the velocity corrections. Also, these corrections are assumed to occur instantaneously so that no position or time change need be accounted for during the maneuver.

Correction of the actual and estimated states.- The inputs to the velocity correction model are ΔV_G and the random execution errors which consist of cutoff error and pointing error. The cutoff error for this study is assumed to be random with normal distribution $N[0,(\sigma_{\Delta V})^2]$ where $\sigma_{\Delta V}$ = 1 percent. The pointing error is resolved into errors in right ascension and declination. Both of these angles are assumed random with distribution $N[0,\sigma^2]$ where $\sigma = 1^o$. The output (17) of the velocity correction model is the correction applied to the actual state. The correction to the estimated state is the measured value of that applied to actual state. The errors in the measuring device are assumed to be random with distribution $N[0,(1 \text{ cm/sec})^2]$ in each axis. These errors are added to the actual values at (18) to simulate the estimated values. The covariance matrix of these errors in measuring the correction is added to the P matrix (19) as follows:

$$P_4 = P_4' + \begin{bmatrix} (1 \text{ cm/sec})^2 & 0 & 0 \\ 0 & (1 \text{ cm/sec})^2 & 0 \\ 0 & 0 & (1 \text{ cm/sec})^2 \end{bmatrix} \tag{16}$$

where P_4' is the value of P_4 before the velocity is corrected. This completes the velocity correction maneuver. The computer now continues on to process the next space-angle observation.

Generation of Initial Prediction Matrices

To provide the initial prediction matrix $A(t_e;t_k)$ in equation (11) the system was, prior to the actual run, commanded to run from injection to the end point without observations or velocity corrections. This resulted in a transition matrix $\Phi(t_e;t_o)$, which was the initial value of $A(t_e;t_o)$ desired as an input for the actual run. A different $A(t_e;t_o)$ was generated in this way for each word length to be used.

RESULTS AND DISCUSSION

In this section the results of the effects of reduced computational precision are presented and discussed. The areas to be considered are:

1. The effect of word length on earth-moon trajectories.
2. The effect of word length on the optimal filter.
3. The effect of word length on the error propagation model.
4. The effect of word length on state estimation.
5. The effect of word length on guidance.

Effect of Word Length on Earth-Moon Trajectories

In order to determine the effect of word length on trajectories from the earth to the moon, the aperture in mask 2 was set to the desired bit length while all other masks were set to 27 bits. The equations of motion for all three trajectories were then numerically integrated simultaneously from injection to the end point at reference perilune without making any observations.

When the estimated and reference trajectories were compared at the time of reference perilune, t_e, the range magnitude between the two end points, shown in figure 2, gave information showing how position error increased when the precision was reduced. This curve shows that, as precision is reduced below about 25 bits, the error that accumulates over the entire trajectory rises sharply. This tends to limit acceptable word lengths to the range from 25 to 27 bits. However, reference to figure 2 shows that a 25-bit reference trajectory, for example, would be in error by 18 km which would be unsatisfactory in most cases. This indicates that 26 or perhaps 27 bits would be required for reference trajectories. On the other hand, estimated trajectories accumulate errors involving the numerical integration of the equations

of motion only between observations and velocity corrections, and this study has shown that satisfactory results may be expected with precisions as low as 21 bits.

Effect of Word Length on the Optimal Filter

The behavior of the covariance matrix P as a function of word length is used here as a means of assessing the degradation in the performance of the filter as characterized by equations (3) through (7). Errors in those calculations involved in producing the updated P matrix as the result of making an observation, that is, operations ③, ④, ⑤, ⑥, ⑦, and ⑧ in figure 1, will be detected by $\tilde{r}_{rms}$ and $\tilde{v}_{rms}$ defined in equations (8) and (9). Error detection will consist of two steps: first, the effect of the transition matrix on the updating of $P(t_k)$ from time t_{k-1} to t_k in equation (3), and second, the effect on $P'(t_k)$ in equation (7) will be shown.

Effect of transition matrix on the updating of $P(t_k)$.- In table III the quantities $\tilde{r}_{rms}$ for both 16 and 27 bits are given just prior to each sequence of observations and after each velocity correction. Relatively long periods of time elapse between injection and the beginning of the first sequence of observations and between each velocity correction and the succeeding observations sequence. If the transition matrices for these time periods were erroneous for 16 bits it would be expected that $\tilde{r}_{rms}$ and $\tilde{v}_{rms}$ for 16 bits would differ considerably from its 27-bit counterpart at the beginning of the next observation sequence. Referring to table III and comparing $\tilde{r}_{rms}$ and $\tilde{v}_{rms}$ for 16 and 27 bits in each row of the table before a velocity correction and at the beginning of the next observation sequence shows clearly that transition matrix effects are quite small.

Effect of precision on $P'(t_k)$.- With the computation of $P'(t_k)$ in equation (7) the optimal filter calculations are complete. If there were significant errors due to reduced precision in the calculations leading up to computation of $P'(t_k)$, it would be reasonable to expect that a series of observations taken close together would lead to serious errors in $P'(t_k)$. The results of four sequences of observations in which the observations were taken close together is given in table III. However, when $\tilde{r}_{rms}$ and $\tilde{v}_{rms}$ are compared at the beginning of an observation sequence and again at the velocity correction following that sequence, it is seen that these quantities are not adversely affected by the reduction of precision from 27 bits to 16 bits. It is, therefore, concluded that precisions as low as 16 bits would provide adequate results in the filter computations.

Effect of Word Length on the Error Propagation Model

The performance of the system in estimating the errors in state is heavily dependent on the ability to correctly propagate errors at one time into corresponding errors at a later time. This operation is expressed

mathematically in equation (1). Since equation (3) is derived directly from equation (1), it can be seen that the system will not "know" the true error in estimate if $\Phi(t_k;t_{k-1})$ does not correctly represent the way in which errors propagate from one time to the next.

Figure 2 shows that large deviations of the estimated trajectory from the reference trajectory develop for the larger reductions in precisions. Yet, from table III it was shown that the P matrix was virtually invariant with precision. Therefore, $\Phi(t_k;t_{k-1})$ is at fault and does not represent the true way in which errors propagate. Since

$$\Phi(t_k;t_{k-1}) = \int_{t_{k-1}}^{t_k} F(t)\Phi(t_k;t_k)dt \tag{17}$$

and since the numerical integration subroutine has high internal precision it is obvious that any difficulties with $\Phi(t_k;t_{k-1})$ stems from $F(t)$. The end result of using the transition matrix derived from equation (17) is shown in figure 3. Here, $\tilde{r}$, the true error in estimate is compared to $\tilde{r}_{rms}$ for the case where all the masks had a 16-bit aperture. It is seen that $\tilde{r}_{rms}$ and, therefore, P, is much too conservative to adequately represent in a statistical sense the errors in estimate which actually existed. As a result, the system believes that its knowledge of the estimated state is better than it really is. This causes the optimal filter to pay less attention to observational information than it should.

Appendix B gives a method for overcoming the difficulties with $F(t)$, but its practical value is limited by the requirement of detailed knowledge of the way in which the individual errors, caused by reduced precision, propagate with time. For this reason it was not used in this study but is presented as a method that might be applicable to a less complex problem or in the event future study should reveal a simple model of the propagation of the reduced precision errors.

The question of how $F(t)$ is affected by reduced precision is answered by comparing $\Phi(t_e;t_o)$ for 16 bits with the transition matrix for 27 bits. This comparison shows that some degradation accumulates when t_e and t_o are separated in time by as much as 12 hours, but for shorter periods of an hour or less $\Phi(t_e;t_o)$ is essentially invariant with precision. Therefore, if long transition periods are to be encountered, precisions higher than 16 bits are required in the calculation of $F(t)$.

Effect of Word Length on State Estimation Through Space-Angle Observations

Navigation with the system being investigated requires observations of one or more space angles which are related to the vehicle state through the space-angle geometry. When an observer measures these space angles with an imperfect instrument, the quality of the knowledge gained will depend on the combined instrument and observer accuracies. These actual space angles must

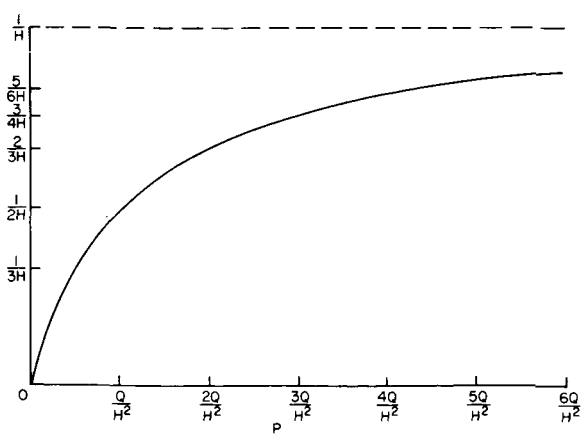
then be fed into a computer. The precision with which the angles are accepted by the computer is dependent on its word length. In this simulation, the input precision is controlled by mask 3.

Effect of the space angles.- Two computer runs were made in which mask 3 was 27 bits in one and 16 bits in the other. All other masks remained fixed at 16 bits. The similarity of these runs was such that the maximum position deviation between the two estimated trajectories was approximately 11 km. This indicates that bits of lower significance than 16 in the space angles tend to be ignored in the updating of the estimated state. This implies that little or no improvement in the estimated state can be expected from increasing the precision of the space angles unless the precision of $K(t_k)$ is also increased. This loss of the lower order bits results in a considerable loss of system sensitivity to position errors since the quantization error in the 16th bit is quite large in terms of seconds of arc.

Effect of the P matrix.- The effect of a conservative $P(t_k)$ matrix on the correction of the estimated state at (13) in figure 1 may be seen from the following simplified example: Suppose the state vector is a 1×1 and only one variable is being observed. Then all the matrices are 1×1. Suppose also that the output vector at (11) in figure 1 is the column matrix a. Then the correction to the estimated state at (13) is $K(t_k)a$ which, from equation (6) using 1×1 matrices and dropping the subscript, t_k, is

$$Ka = \frac{PH}{PH^2 + Q}\, a = \frac{1}{H}\,\frac{P}{P + Q/H^2}\, a \tag{18}$$

At time t_k both H and Q are constants. Sketch (a) shows how K, and therefore, the correction to the estimated state increases with increasing P. Therefore, a P matrix which is too conservative will always result in too small a correction to the estimated state even if the actual space angles are sufficiently accurate.



Sketch (a)

Errors in correcting the estimated state.- When the P matrix is a very conservative estimator of the existing errors and the K matrix causes large numbers of the least significant bits from the space angles to be ignored, errors in the estimated state may accumulate rapidly with each additional observation. This situation is shown in figure 4 where very large peaks develop in the midcourse region where the rate of change of the space angles is low.

Figure 5 is a plot of $\tilde{r}$ for the 16-bit case shown in figure 4. Included are points before and after each observation rather than only after each observation as in previous plots. The growth of errors between observation sequences is also shown. Inspection of figure 5 shows that the effect of some of the observations is to actually make the estimated state worse. When observations are switched from one body to another in this region, the space-angle differences are large because of the large deviation between the actual and estimated state. The first observation of the moon is this sequence (observation 20) was made with a most dramatic drop in $\tilde{r}$. To examine why a moon observation was so effective, an error ellipsoid is used.

When the P matrix is partitioned into four 3 × 3 submatrices as in equation (10), the submatrix P_1 may be represented as a three-dimensional error ellipsoid of position and, similarly, P_4 an error ellipsoid of velocity. The relative magnitude and direction of the axes represent the relative magnitudes and directions of the uncertainty at the particular vehicle position in space.

Changes to the error ellipsoid, as the result of an observation, are indicative of the knowledge gained from the observation. In particular, these changes are twofold. First, except when near the observed body, the major axis of the ellipsoid tends to aline itself with the vector to the observed body. Second, a reduction in the ellipsoid volume occurs since, by equation (7), an observation requires that the P matrix must be reduced or left unchanged. Also, a large reduction in the volume of an ellipsoid implies a large correction to the estimated state.

Figure 6 shows the position error ellipsoid both before and after observation 20. Included are vectors indicating directions from the vehicle to both the earth and the moon. Because of the immediately preceding four earth observations, the major axis is only 10° from the vehicle-earth vector and the ellipsoid is quite elongated. As a result of the moon observation, the elongation is considerably reduced and the volume is reduced by a factor of 2.46. The orientation of the ellipsoid is skewed toward the vehicle-moon vector.

In summary, the computation of the space angles with precisions greater than that employed by the K matrix will not improve the ability to correct the estimated state by a significant amount. Furthermore, the effect of a P matrix which is too conservative an estimator of the existing errors is to cause the K matrix to give too low a weight to space-angle information. This results in corrections to the estimated state which are too small at each observation. The conservatism of the P matrix may be remedied by correcting the error propagation model or by augmenting the filter computation as outlined in appendix B.

Effect of Word Length on Guidance

The evaluation of the effect of precision on guidance was made by considering the effects on the indicated velocity correction vector, ΔV_G, due to

the estimated state and the prediction submatrix, A_2. Also considered is the way in which reduced precision affects the actual state through the velocity corrections.

Effect of precision on ΔV_G.- Figure 7 shows the magnitude of ΔV_G at each observation and velocity correction for 16, 18, and 27 bits. The separation of the curves for the three different precisions at the time of the first velocity correction (12 hours) is shown to differ from the 27-bit value. The separation at this time increases rapidly as the precision is reduced and is to be expected because of the dependence of ΔV_G on $\hat{x}$ given in equation (14). The separation of the three curves at subsequent velocity corrections should be expected to increase because immediately after the first velocity correction, if it were perfect, ΔV_G would be zero. The execution of ΔV_G was, however, not perfect but at least the pointing and cutoff errors in each case were the same except for the number of least significant bits discarded by the masking. Therefore, the growth in the magnitude of ΔV_G in the time period between a velocity correction and the next observation is due to the errors in determining $\hat{x}$ and the errors in executing the velocity correction.

It is interesting to note that the erratic behavior of $|\Delta V_G|$ is rapidly reduced with time until about 74.5 hours in the 18-bit case. At this point, a sharp dip began and was followed a short time later by a second even sharper dip. Since the computation of ΔV_G depends on the behavior of A_2^{-1}, a plot of the determinant of A_2 was made and is shown in figure 8. Clearly, the peculiar behavior of the 18-bit case with an 18-bit prediction matrix was associated with a large dip during which the determinant of A_2 was negative for a short while. Also shown is the unusual behavior of the 16-bit case when a 16-bit prediction matrix was used. Here, the curve becomes negative sometime after 36 hours and remains so throughout the remainder of the run.

The peculiar behavior of the determinant of A_2 may be explained as follows: If infinite precision were used, the product of all the individual transition matrices between observations and velocity corrections would be given by

$$\prod_{i=1}^{e} \Phi(t_i;t_{i-1})\Phi t_o;t_o) = \Phi(t_e;t_o) = A(t_e;t_o) \tag{19}$$

where $\prod_{i=1}^{e} \Phi(t_i;t_{i-1})\Phi(t_o;t_o)$ means

$$\Phi(t_e;t_{e-1}) \ . \ . \ . \ \Phi(t_2;t_1)\Phi(t_1;t_o)\Phi(t_o;t_o) \tag{20}$$

and

$$\Phi(t_o;t_o) = I$$

It is obvious that at the end point the prediction matrix under these ideal conditions is given by

$$A(t_e;t_e) = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} = I \qquad (21)$$

If A_1, A_2, A_3, and A_4 are 3×3 submatrices, then $A_1 = A_4 = I$ and $A_2 = A_3 = 0$. In general, however, the left side of equation (19) does not equal $A(t_e;t_o)$. This is because the product of individual piecewise transition matrices which have been generated from exact initial conditions by numerical integration over successive short-time intervals yields a different final result than a transition matrix integrated from exact initial conditions over a time period which is the sum of the successive short-time intervals. Therefore, at the end point A_1 and A_4 do not equal I and also A_2 and A_3 do not equal zero because of accumulated errors. As the trajectory nears the end point, these errors dominate the situation more and more with the dominance occurring earlier as the precision is reduced. This can be seen by comparison of the 16- and 18-bit cases in figure 8. The effect of substituting an initial A matrix generated with 27-bit precision may also be seen by comparing the 16-bit case using a 27-bit initial prediction matrix and the 27-bit case using a 27-bit A matrix with the 16-bit case results in a curve which is quite similar to the 27-bit case in which a 27-bit starting A matrix is used. (Note that determinant of A_2 is not zero at the end point as would be expected if infinite precision had been used.)

The effect on guidance due to the use of a 16- and 27-bit A matrix is given in table IV. Here it is seen that the results with the 27-bit A matrix are generally better. Furthermore, at the end point the results are markedly different as would be expected because of the negative determinant of A_2 at the time of the third correction.

Effect of guidance errors on the actual state.- Although the actual trajectory was computed in full precision, it is affected by the estimated state through velocity corrections. Table V shows in column 1 the standard 27-bit case. In column 2 is shown the effect of reducing the precision of the estimated trajectory to 16 bits (mask 2) when all other masks were set to 27 bits. Column 3 shows the effect when all masks have 16-bit apertures except for mask 2 which was 27 bits. Columns 4 and 5 show the cases when all masks had 16-bit apertures but different starting prediction (A) matrices. All starting prediction matrices were 27 bit except for the case shown in column 5 where a 16-bit starting prediction matrix was used. In all columns the magnitude of r, $|r|$, is given along with $|\hat{r}|$, the magnitude of $\hat{r}$. The quantity $|\hat{r}|$ is included because $|r|$ depends on the previous value of $\hat{r}$. Referring to equation (14) we see that the results in column 2 show the effect of $\hat{x}$ on the actual state and on the estimated state. Similarly, column 3 shows the effect of reduced precision on the computation of $A_2^{-1}A_1$. Column 4 shows the combined effects of precision on $\hat{x}$ and on $A_2^{-1}A_1$. Column 5 shows, in addition to that shown in column 4, the additional effect on $A_2^{-1}A_1$ of a 16-bit starting prediction matrix.

The end-point results for columns 2, 4, and 5 are obviously unsatisfactory from a guidance point of view. Column 3 is somewhat questionable and in general would be considered unsatisfactory. To improve the results in column 3 would require greater precision in forming $A_2^{-1}A_1$.

In summary, the prediction matrices generated with precisions of 18 bits and lower should not be used. In fact, a 27-bit prediction matrix should always be used if available. Computation of $A_2^{-1}A_1$ in the guidance equation should be done with the aperture in mask 5 no less than 20 bits.

CONCLUSIONS

The results of this study show that satisfactory performance of the navigation and guidance scheme proposed in NASA TN D-1208 and NASA TR R-135 can be accomplished at greatly reduced precision in some general areas and with lesser reductions in others provided that the increased errors are properly accounted for. This should allow similar simulations to be carried out by computers using greater precision only in certain critical areas rather than throughout the entire simulation as would have been required heretofore. The general areas and the corresponding precisions required are as follows:

1. The precision of the estimated trajectory must be about 21 to 27 bits (mask 2).

2. Reference trajectories require a precision of 26 or 27 bits.

3. The computations embodied in the optimal filter may be computed with precisions as low as 16 bits (mask 1).

4. The transition matrix $\Phi(t_k;t_{k-1})$ (controlled by mask 4) may be derived from $F(t)$ where 16-bit precision is used. However, if t_k and t_{k-1} are widely separated in time, the precision should be increased. For $\Phi(t_k;t_{k-1})$ to represent the correct error propagation model requires that $F(t)$ and/or the filter computations be augmented to take into account the increase in uncertainty.

5. Increasing the precision of the space-angle computations (mask 3) over that employed by the weighting matrix, K, was found to produce only a negligible improvement in the estimated state.

6. The effect of a P matrix which is too conservative an estimator of the errors in the estimated state is to cause the weighting matrix, K, to give too little weight to space-angle information. This can contribute to the build-up of very large errors in the estimated state.

7. Prediction matrices generated with precisions of 18 bits and lower should not be used. If available, 27-bit prediction matrices should be used. The computation of $A_2^{-1}A_1$ in the guidance should be done with the aperture in mask 5 greater than about 20 bits.

8. The above conclusions lead to the further conclusion that the previous studies, which used 27-bits precision throughout, were done with adequate precision.

Ames Research Center
National Aeronautics and Space Administration
Moffett Field, Calif., Jan. 10, 1966

APPENDIX A

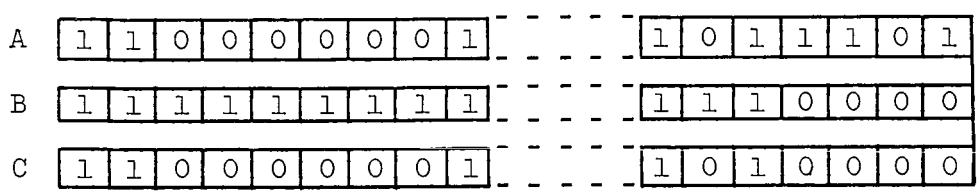
WORD LENGTH REDUCTION

In FORTRAN II, floating-point arithmetic is used throughout (except for indexing and bookkeeping) to maintain the decimal point of numbers automatically. Internally the IBM 7094 uses floating-point binary numbers which are presented as a signed proper fraction, called the mantissa, times some integral power of 2, called the characteristic or exponent. The number or computer word is 36 binary bits,[1] made up, from left to right, of a sign of the mantissa, an 8-bit characteristic, and a 27-bit mantissa. The precision with which a number may be represented may be varied by changing the number of binary bits in the mantissa where 27 is the maximum.

The function of the mask is to specify in the Boolean "and" operation the number of successive least significant bits of the mantissa to be set to zero. This is done by placing in the corresponding bit of the mask a "1" where a bit is to be retained and a "0" where a bit is to be discarded. The following sketch illustrates a masking operation in which a floating-point number, A, is to be truncated from 27 to 23 bits. If in the sketch the mask is the number, B, and the truncated result is C, then the masking operation may be described by the Boolean expression

$$A*B = C$$

where the "*" is the "and" operator.



Thus, the "and" operation places bit by bit a "1" in C if both A and B contain a "1," otherwise C is set to "0." The net effect, then, is to retain the bits of A where the corresponding bit in B is a "1."

[1]Reference Manual IBM 7094 Data Processing System.

APPENDIX B

CORRECTION OF P MATRIX TO COMPENSATE FOR INCORRECT ERROR PROPAGATION MODEL

In order to correct the P matrix to compensate for incorrect error propagation model, suppose the equations of motion for deviations from the reference in the 27-bit case are given by the linear differential equation

$$\dot{x} = Fx \tag{B1}$$

where

$$x = \begin{bmatrix} \delta x \\ \delta y \\ \delta z \\ \delta \dot{x} \\ \delta \dot{y} \\ \delta \dot{z} \end{bmatrix} \tag{B2}$$

and the δ's represent deviations from the reference. Now, in reduced precision, it is assumed that the equations of motion have a term missing which may be accounted for by the addition of the term $F_p x_p$ so that, for the 16-bit case which will be used hereafter as an example, equation (B1) becomes

$$\dot{x}_{27} = F_{16} x_{16} + F_p x_p \tag{B3}$$

Here x_p is some state vector which is not updated and is active only from one observation to the next; F_p is a matrix whose elements are assumed to be functions of time and the precision of the integration of the equations of motion. The transition matrices Φ_{16} and Ψ_{16} are found from:

$$\Phi_{16}(t;t_o) = \int_{t_o}^{t} F_{16}(\tau)\Phi(\tau_o)d\tau \tag{B4}$$

$$\Psi_{16}(t;t_o) = \int_{t_o}^{t} F_p(\tau)\Psi(\tau_o)d\tau \tag{B5}$$

where $\Phi(\tau_o)$ and $\Psi(\tau_o)$ represent the proper initial conditions.

The deviations at some later time, t, can now be found from the expressions

$$\tilde{x}_{16}(t) = \Phi_{16}(t;t_o)\tilde{x}(t_o) \tag{B6}$$

and

$$x_p(t) = \Psi_{16}(t;t_o)x_p(t_o) \tag{B7}$$

Let x be the vector sum of $\hat{x}$ and $\tilde{x}$ where $\hat{x}$ is the estimate of x, and $\tilde{x}$ is the error in the estimate of x. Thus,

$$x_{16}(t) = \hat{x}_{16}(t) + \tilde{x}_{16}(t) \tag{B8a}$$

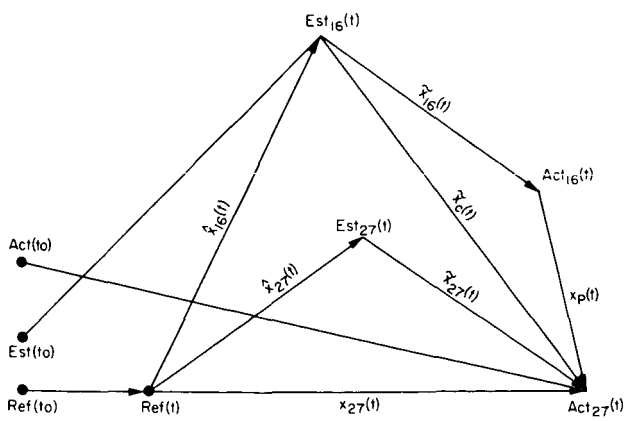
$$x_{27}(t) = \hat{x}_{27}(t) + \tilde{x}_{27}(t) \tag{B8b}$$

To simplify writing the following equations, let

$$\Phi_{16} = \Phi_{16}(t;t_o) \tag{B9a}$$

$$\Psi_{16} = \Psi_{16}(t;t_o) \tag{B9b}$$

Sketch (b) is a vector diagram showing the actual, estimated, and reference states at times t_o and t. The state at t is the result of numerical integration of the equations of motions from t_o to t using the initial conditions at t_o. When 27-bit precision is used, the estimated state at t is shown as the point $Est_{27}(t)$. The estimate of $x_{27}(t)$, $\hat{x}_{27}(t)$, is the vector from the reference state, Ref(t), to $Est_{27}(t)$. The error in the estimate, $\tilde{x}_{27}(t)$, is the vector from $Est_{27}(t)$ to $Act_{27}(t)$ in accordance with equation (B8b). On the other hand, when 16-bit precision is used, the estimated state at t is shown as the point $Est_{16}(t)$. In this case, application of equation (B8a) does not yield $x_{27}(t)$, the true state at t. Instead, it yields the state at $Act_{16}(t)$ which is in error by $x_p(t)$ because of the inability of equation (B6) to produce the true error vector, $\tilde{x}_c(t)$ which is given by



Sketch (b)

$$\tilde{x}_c(t) = \tilde{x}_{16}(t) + x_p(t)$$

$$= \Phi_{16}\tilde{x}_{16}(t_o) + \Psi_{16}x_p(t_o) \tag{B10}$$

and, as seen from sketch (b), is related to the true state by

$$x_{27}(t) = \hat{x}_{16}(t) + \tilde{x}_c(t) \tag{B11}$$

The true covariance matrix of the error in estimate is defined as

$$P_c(t) = E[\tilde{x}_c \tilde{x}_c{}^T] \tag{B12}$$

When (B10) is substituted into (B12)

$$P_c(t) = E\left\{[\tilde{x}_{16}(t) + x_p(t)][\tilde{x}_{16}^T(t) + x_p{}^T(t)]\right\}$$

$$= E[\tilde{x}_{16}(t)\tilde{x}_{16}^T(t) + \tilde{x}_{16}(t)x_p{}^T(t) + x_p(t)\tilde{x}_{16}^T(t) + x_p(t)x_p{}^T(t)]$$

$$= \Phi_{16}E[\tilde{x}_{16}(t_o)\tilde{x}_{16}^T(t_o)]\Phi_{16}^T + \Phi_{16}E[\tilde{x}_{16}(t_o)x_p(t_o)^T]\psi_{16}^T$$

$$+ \psi_{16}E[x_p(t_o)\tilde{x}_{16}^T(t_o)]\Phi_{16}^T + \psi_{16}E[x_p(t_o)x_p{}^T(t_o)]\psi_{16}^T \tag{B13}$$

Let

$$P_{16}(t_o) = E[\tilde{x}_{16}(t_o)\tilde{x}_{16}^T(t_o)] \tag{B14}$$

$$P_{16}(t) = \Phi_{16}P_{16}(t_o)\Phi_{16}^T \tag{B15}$$

$$C(t_o) = E[\tilde{x}_{16}(t_o)x_p{}^T(t_o)] \tag{B16}$$

$$D(t_o) = E[x_p(t_o)x_p{}^T(t_o)] \tag{B17}$$

Now (B13) becomes

$$P_c(t) = P_{16}(t) + \Phi_{16}C(t_o)\psi_{16}^T + \psi_{16}C(t_o)^T\Phi_{16}^T + \psi_{16}D(t_o)\psi_{16}^T \tag{B18}$$

The matrix C gives the correlation of the vectors $\tilde{x}_{16}$ and x_p. Therefore, C must be updated in order to reflect the change in $\tilde{x}$ due to its translation in time from t_o to t. Thus, at time t, equation (B16) becomes

$$C(t) = E[\tilde{x}_c(t)x_p{}^T(t)] \tag{B19}$$

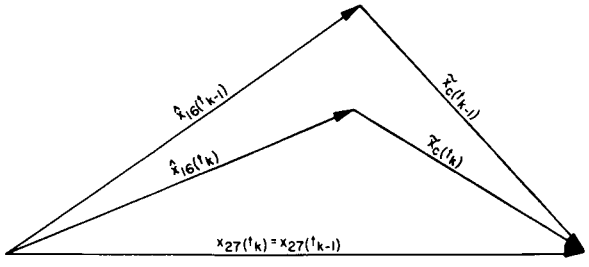
Substituting for $\tilde{x}_c(t)$ from (B10), we have

$$C(t) = E\left\{[\Phi_{16}\tilde{x}_{16}(t_o) + \Psi_{16}x_p(t_o)]x_p(t_o)^T\Psi_{16}^T\right\}$$

$$= E\left[\Phi_{16}\tilde{x}_{16}(t_o)x_p(t_o)^T\Psi_{16}^T + \Psi_{16}x_p(t_o)x_p(t_o)^T\Psi_{16}^T\right]$$

$$= \Phi_{16}C(t_o)\Psi_{16}^T + \Psi_{16}D(t_o)\Psi_{16}^T \qquad (B20)$$

Returning to sketch (b), the situation is shown where the three trajectories (reference, actual, estimated) have been integrated from t_o to t, but only k-1 observations have been made. When the kth observation is made at time t both $P_c(t)$ from (B18) and $C(t)$ from (B20) include information only from the previous k-1 observations. Therefore, to avoid confusion, it is desirable to place an additional subscript on t to indicate whether or not updating has taken place. Thus, $P(t_{k-1})$ and $C(t_{k-1})$ must be updated to time t_k to include the new knowledge gained from the kth observation. This new knowledge allows the estimated state to be corrected as shown in sketch (c).



Sketch (c)

With $P_c(t_{k-1})$ known, $P_c(t_k)$ is computed from the usual equation

$$P_c(t_k) = [I - K(t_k)H(t_k)]P_c(t_{k-1}) \qquad (B21)$$

where

$$K(t_k) = P_c(t_{k-1})H^T(t_k)[H(t_k)P_c(t_{k-1})H^T(t_k) + Q(t_k)]^{-1} \qquad (B22)$$

To update $C(t_{k-1})$ to time t_k note that from sketch (c)

$$\tilde{x}_c(t_{k-1}) = x_{27}(t_{k-1}) - \hat{x}_{16}(t_{k-1}) \qquad (B23)$$

and also

$$\tilde{x}_c(t_k) = x_{27}(t_k) - \hat{x}_{16}(t_k)$$

$$= x_{27}(t_{k-1}) - \hat{x}_{16}(t_k) \qquad (B24)$$

From reference 3, the actual observed space angles, y, are given by

$$y(t_k) = H(t_k)x_{27}(t_{k-1}) + q(t_k) \tag{B25}$$

where $q(t_k)$ is the error in measuring the angles. The estimate of the same space angles, $\hat{y}$, is given by

$$\hat{y}(t_k) = H(t_k)\hat{x}_{16}(t_{k-1}) \tag{B26}$$

The error in the estimate of the space angles is given by subtracting (B26) from (B25). Thus,

$$y(t_k) - \hat{y}(t_k) = H(t_k)[x_{27}(t_{k-1}) - \hat{x}(t_{k-1})] + q(t_k)$$

$$= H(t_k)\tilde{x}_c(t_{k-1}) + q(t_k) \tag{B27}$$

The new estimate of the state at time t_k is from reference 3

$$\hat{x}_{16}(t_k) = \hat{x}_{16}(t_{k-1}) + K(t_k)[H(t_k)\tilde{x}_c(t_{k-1}) + q(t_k)] \tag{B28}$$

Substituting (B28) into (B24) and using (B23) yields

$$\tilde{x}_c(t_k) = \tilde{x}_c(t_{k-1}) - K(t_k)[H(t_k)\tilde{x}_c(t_{k-1}) + q(t_k)] \tag{B29}$$

which is the equation for updating $\tilde{x}_c$ from time t_{k-1} to t_k. Before proceeding with the derivation of $C(t_k)$, it is noted that $q(t_k)$ is a vector such that

$$Q(t_k) = E[q(t_k)q^T(t_k)] \tag{B30}$$

where $Q(t_k)$ is the covariance matrix of the observational errors and is assumed to be a known function of time. Also, it will be assumed that $q(t)$ and $x_p(t)$ are uncorrelated which results in

$$E[q(t_k)x_p^T(t_{k-1})] = 0 \tag{B31}$$

Now, to find $C(t_k)$ equation (B29) is substituted into equation (B19) and using equation (B31) it is found that

$$C(t_k) = E\left\{\tilde{x}_c(t_{k-1})x_p^T(t_{k-1}) - K(t_k)[H(t_k)\tilde{x}_c(t_{k-1}) + q(t_k)]x_p^T(t_{k-1})\right\}$$

$$= E[\tilde{x}_c(t_{k-1})x_p^T(t_{k-1})] - K(t_k)H(t_k)E[\tilde{x}_c(t_{k-1})x_p^T(t_{k-1})]$$

$$- K(t_k)E[q(t_k)x_p^T(t_{k-1})]$$

$$= C(t_{k-1}) - K(t_k)H(t_k)C(t_{k-1}) \qquad \text{(B32)}$$

To find the corrected P matrix, P_c, where there is increased uncertainty due to increased error in the estimated state, the computational steps are summarized below:

(1) To update in time from t_o to t, compute:

$$C(t) = \Phi_{16}C(t_o)\psi_{16}^T + \psi_{16}D(t_o)\psi_{16}^T$$

$$P_c(t) = P_{16}(t) + \Phi_{16}C(t_o)\psi_{16}^T + \psi_{16}C(t_o)^T\Phi_{16}^T + \psi_{16}D(t_o)\psi_{16}^T$$

(2) When t is the time of an observation t_k, then step (1) gives $C(t_{k-1})$ and $P(t_{k-1})$. To update as a result of the observation compute:

$$P_c(t_k) = [I - K(t_k)H(t_k)]P_c(t_{k-1})$$

$$C(t_k) = [I - K(t_k)H(t_k)]C(t_{k-1})$$

This method allows the P matrix to be corrected for errors due to reduced precision providing F_p and x_p in equation (B3) can be found. Unfortunately, this may not be easy even for relatively simple situations.

REFERENCES

1. McLean, John D.; Schmidt, Stanley F.; and McGee, Leonard A.: Optimal Filtering and Linear Prediction Applied to a Midcourse Navigation System for the Circumlunar Mission. NASA TN D-1208, 1962.

2. Smith, Gerald L.; Schmidt, Stanley F.; and McGee, Leonard A.: Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle. NASA TR R-135, 1962.

3. Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems. J. Basic Engr., vol. 82, no. 1, March 1960, pp. 35-45. (Also available as: Monograph 60-11, RIAS, Inc.; Office Sci. Res. TN 59-268; and ASME Paper 59-IRD-11.)

4. Kalman, R. E.; and Bucy, R. S.: New Results in Linear Filtering and Prediction Theory. J. Basic Engr., vol. 83, no. 1, March 1961, pp. 95-108. (Also available as: Monograph 61-8, RIAS, Inc.; and ASME Paper 60-JAC-12.)

5. Battin, Richard H.: Astronautical Guidance. McGraw-Hill Book Co., N. Y., 1964, p. 306.

TABLE I.- H MATRIX PARTIALS

	$\frac{\partial}{\partial X}$	$\frac{\partial}{\partial Y}$	$\frac{\partial}{\partial Z}$
α_E	$\frac{X_{EV}Z_{EV}}{R_{EV}^2(X_{EV}^2 + Y_{EV}^2)^{1/2}}$	$\frac{Y_{EV}Z_{EV}}{R_{EV}^2(X_{EV}^2 + Y_{EV}^2)^{1/2}}$	$\frac{Z_{EV}^2 - R_{EV}^2}{R_{EV}^2(X_{EV}^2 + Y_{EV}^2)^{1/2}}$
β_E	$\frac{-Y_{EV}}{X_{EV}^2 + Y_{EV}^2}$	$\frac{X_{EV}}{X_{EV}^2 + Y_{EV}^2}$	0
γ_E	$\frac{-R_E X_{EV}}{R_{EV}^2(R_{EV}^2 - R_E^2)^{1/2}}$	$\frac{-R_E Y_{EV}}{R_{EV}^2(R_{EV}^2 - R_E^2)^{1/2}}$	$\frac{-R_E Z_{EV}}{R_{EV}^2(R_{EV}^2 - R_E^2)^{1/2}}$
α_M	$\frac{X_{MV}Z_{MV}}{R_{MV}^2(X_{MV}^2 + Y_{MV}^2)^{1/2}}$	$\frac{Y_{MV}Z_{MV}}{R_{MV}^2(X_{MV}^2 + Y_{MV}^2)^{1/2}}$	$\frac{Z_{MV}^2 - R_{MV}^2}{R_{MV}^2(X_{MV}^2 + Y_{MV}^2)^{1/2}}$
β_M	$\frac{-Y_{MV}}{X_{MV}^2 + Y_{MV}^2}$	$\frac{X_{MV}}{X_{MV}^2 + Y_{MV}^2}$	0
γ_M	$\frac{-R_M X_{MV}}{R_{MV}^2(R_{MV}^2 - R_M^2)^{1/2}}$	$\frac{-R_M Y_{MV}}{R_{MV}^2(R_{MV}^2 - R_M^2)^{1/2}}$	$\frac{-R_M Z_{MV}}{R_{MV}^2(R_{MV}^2 - R_M^2)^{1/2}}$

R_E radius of earth

R_M radius of moon

$\overline{R_{EV}}$ = earth-vehicle vector = $\begin{Bmatrix} X_{EV} \\ Y_{EV} \\ Z_{EV} \end{Bmatrix}$

$\overline{R_{EM}}$ earth-moon distance = $\begin{Bmatrix} X_{EM} \\ Y_{EM} \\ Z_{EM} \end{Bmatrix}$

Moon-vehicle vector = $\overline{R_{EV}} - \overline{R_{EM}} = \begin{Bmatrix} X_{MV} \\ Y_{MV} \\ Z_{MV} \end{Bmatrix}$

TABLE II.- OBSERVATION ANGLES

Observed body	α	β	γ
Earth	$-\sin^{-1}\left(\frac{Z_{EV}}{R_{EV}}\right)$	$\sin^{-1}\frac{-Y_{EV}}{(X_{EV}^2 + Y_{EV}^2)^{1/2}}$	$\sin^{-1}\left(\frac{R_E}{R_{EV}}\right)$
Moon	$-\sin^{-1}\left(\frac{Z_{MV}}{R_{MV}}\right)$	$\sin^{-1}\frac{-Y_{MV}}{(X_{MV}^2 + Y_{MV}^2)^{1/2}}$	$\sin^{-1}\left(\frac{R_M}{R_{MV}}\right)$

α = declination of observed body

β = right ascension of observed body

γ = $\frac{1}{2}$ subtended angle of observed body

R_E = radius of the earth, R_M = radius of the moon

$\overline{R_{EV}}$ = earth-vehicle vector = $\begin{bmatrix} X_{EV} \\ Y_{EV} \\ Z_{EV} \end{bmatrix}$

$\overline{R_{MV}}$ = moon-vehicle vector = $\begin{bmatrix} X_{MV} \\ Y_{MV} \\ Z_{MV} \end{bmatrix} = \overline{R_{EV}} - \overline{R_{EM}}$

$\overline{R_{EM}}$ = earth-moon vector = $\begin{bmatrix} X_{EM} \\ Y_{EM} \\ Z_{EM} \end{bmatrix}$

TABLE III.- COMPARISON OF $\tilde{r}_{rms}$ AND $\tilde{v}_{rms}$ FOR 16 AND 27 BITS

		16-bit word length		27-bit word length	
Time, hr	Comments	$\tilde{r}_{rms}$, km	$\tilde{v}_{rms}$, m/s	$\tilde{r}_{rms}$, km	$\tilde{v}_{rms}$, m/s
3.5	At time of first observation	48.57	4.820	48.62	4.830
12	Immediately after executing first velocity correction	16.82	.533	16.83	.533
24	At time of sixteenth observation	43.21	.700	43.26	.701
36	Immediately after executing second velocity correction	9.21	.089	9.23	.089
64	At time of twenty-eighth observation	18.55	.113	18.61	.114
72	Immediately after executing third velocity correction	6.94	.054	6.98	.054
74	At time of thirty-sixth observation	7.36	.077	7.40	.078
78.847	At end point	1.71	.163	1.79	.170

TABLE IV.- MAGNITUDE OF $\hat{r}$, KM

At time of	16 bits 16-bit A matrix	16 bits 27-bit A matrix
Resumption of observations after first velocity correction (24 hours)	216	201
Resumption of observations after second velocity correction (64 hours)	363	72
Resumption of observations after third velocity correction (74 hours)	259	299
End point (78.847 hours)	1177	571

TABLE V.- COMPARISON OF ACTUAL AND ESTIMATED TRAJECTORY POSITION DEVIATIONS FROM THE REFERENCE TRAJECTORY

	27 bits 27-bit estimated trajectory 27-bit A matrix		27 bits 16-bit estimated trajectory 27-bit A matrix		16 bits 27-bit estimated trajectory 27-bit A matrix		16 bits 16-bit estimated trajectory 27-bit A matrix		16 bits 16-bit estimated trajectory 16-bit A matrix	
	$\lvert\hat{r}\rvert$	$\lvert r\rvert$	$\lvert\hat{r}\rvert$	$\lvert r\rvert$	$\lvert\hat{r}\rvert$	$\lvert r\rvert$	$\lvert\hat{r}\rvert$	$\lvert r\rvert$	$\lvert\hat{r}\rvert$	$\lvert r\rvert$
First velocity correction (12 hours)	36.8	36.3	103	36.3	36.8	36.3	93.0	36.3	93.0	36
Second velocity correction (36 hours)	22.3	18.6	241	422	24.0	21.5	189	387	218	401
Third velocity correction (72 hours)	16.1	12.7	468	728	40.9	41.6	443	713	253	454
End point (78.847 hours)	8.0	7.5	395	406	52.4	49.8	469	584	1177	1196

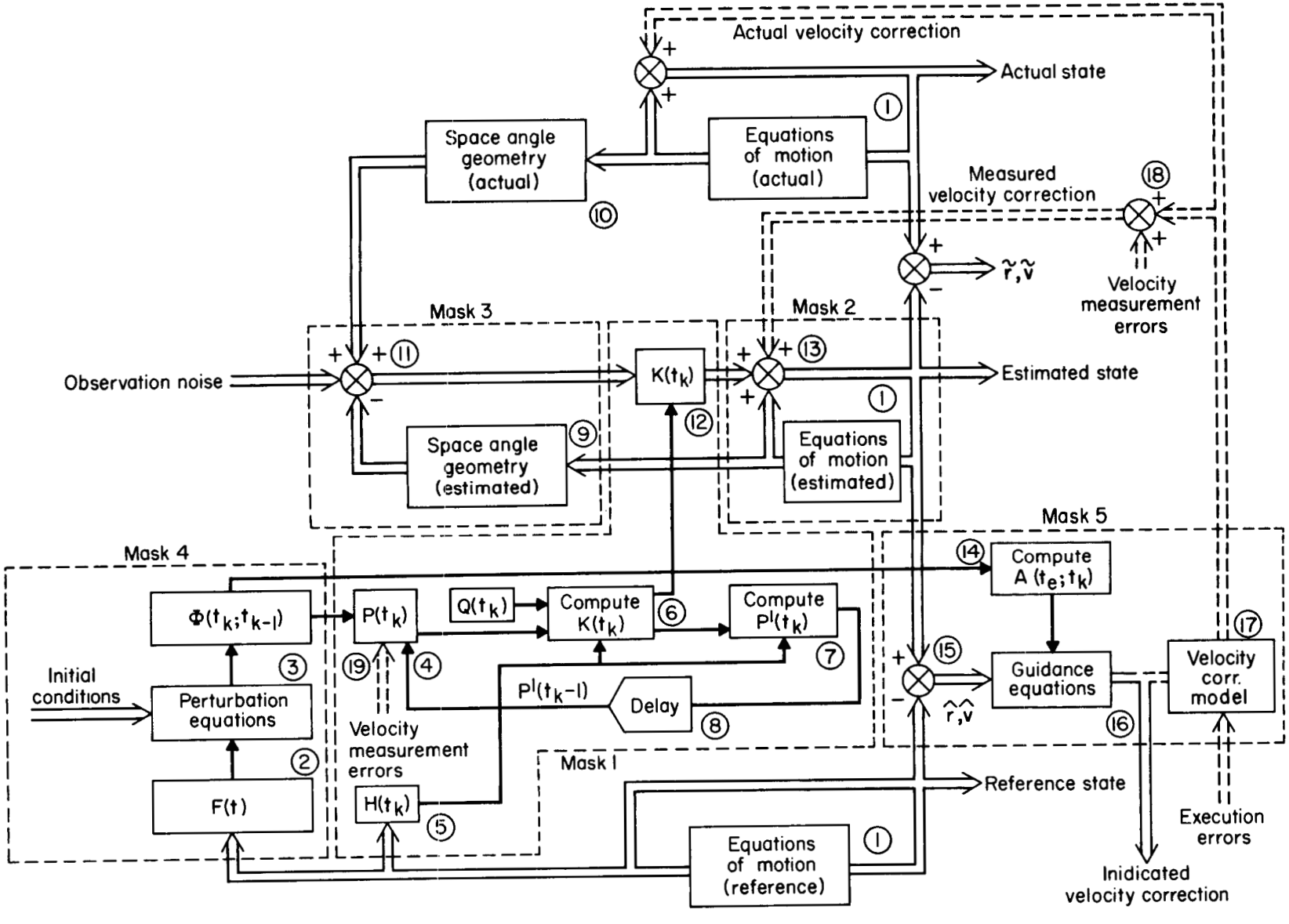


Figure 1.- Block diagram of complete system.

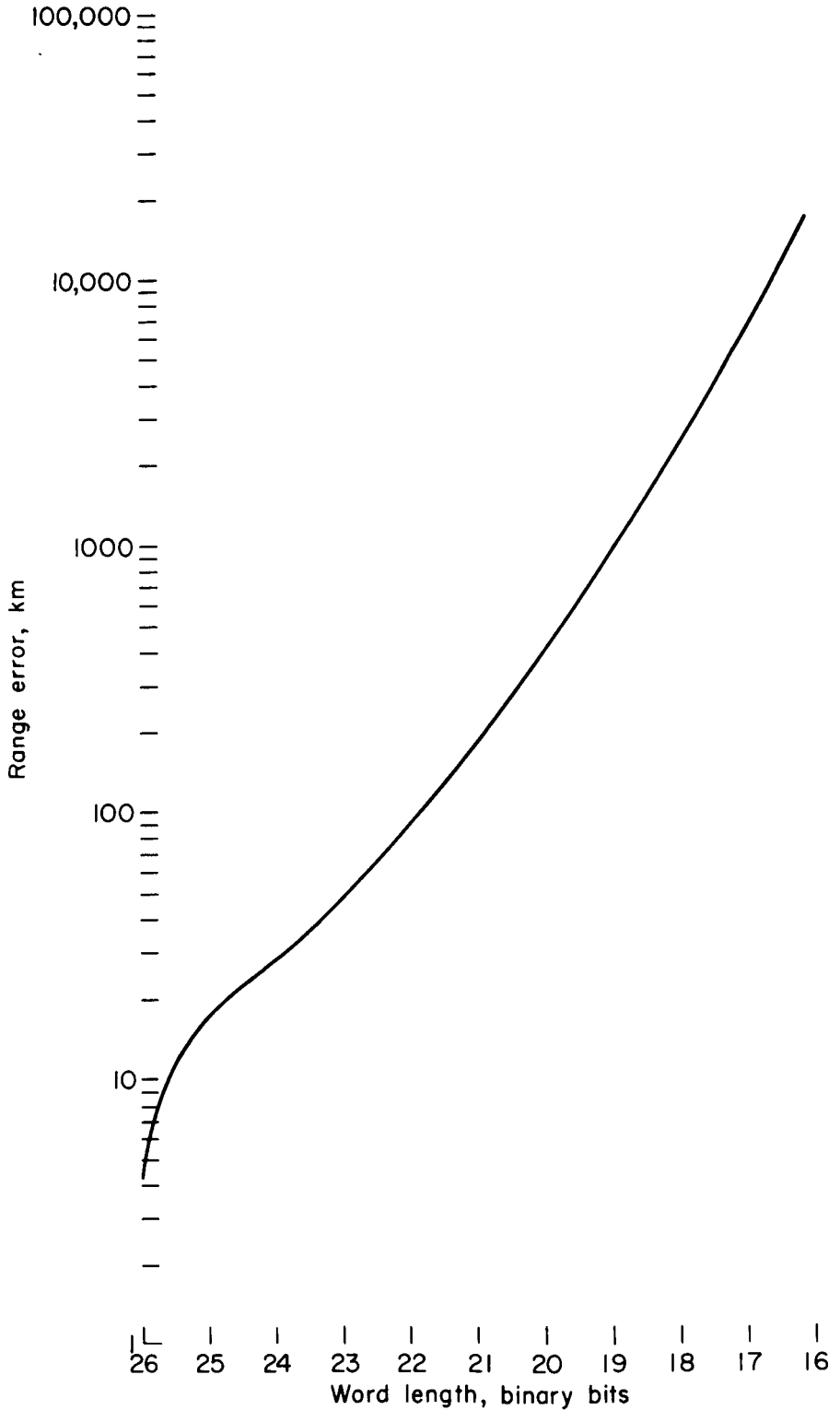


Figure 2.- Deviation of $|\hat{r}|$ from the value obtained with 27-bit precision (earth-to-moon trajectory).

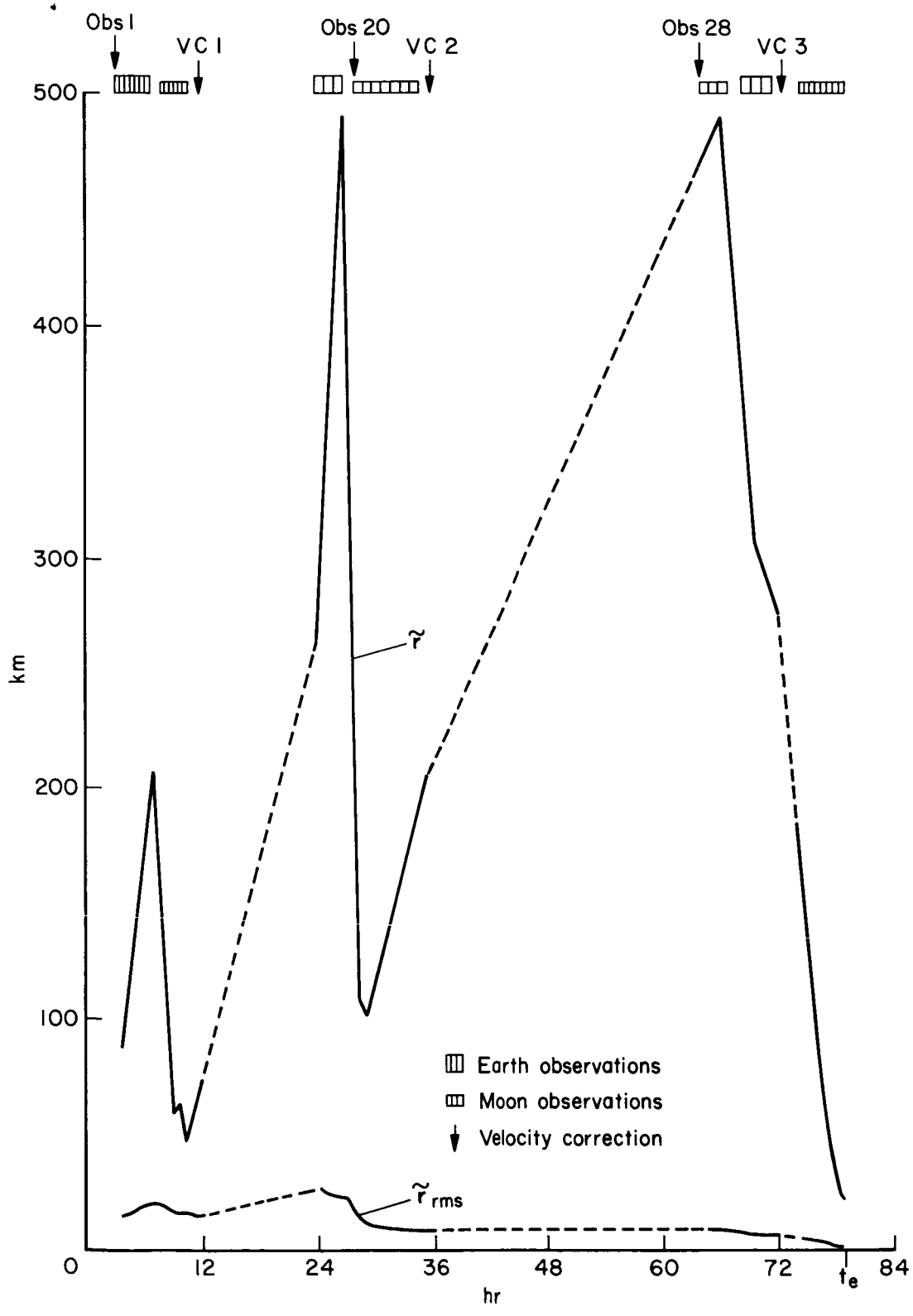


Figure 3.- Comparison of $\tilde{r}$ and $\tilde{r}_{rms}$ for 16 bits.

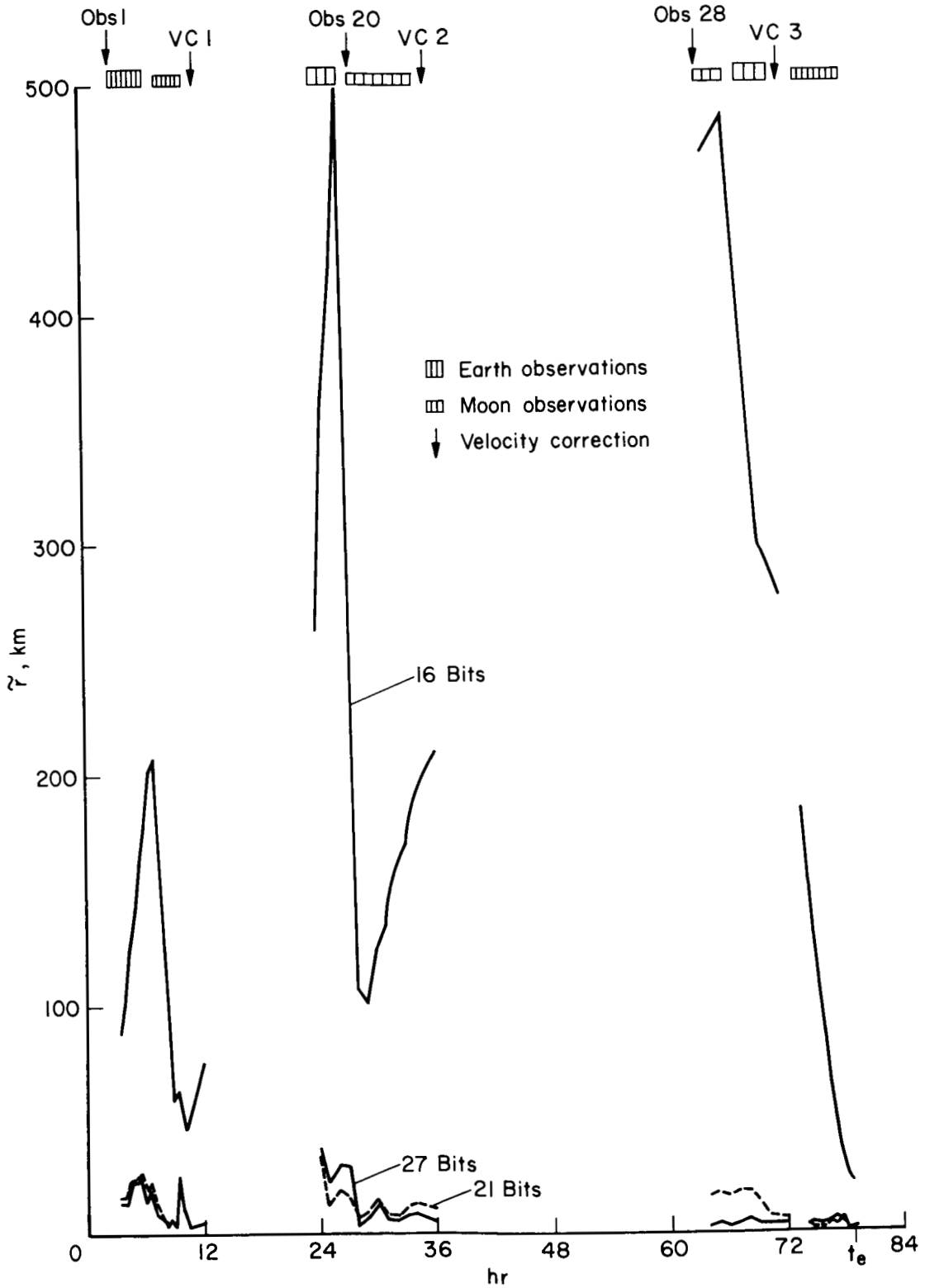


Figure 4.- Comparison of $\tilde{r}$ for 16, 21, and 27 bits.

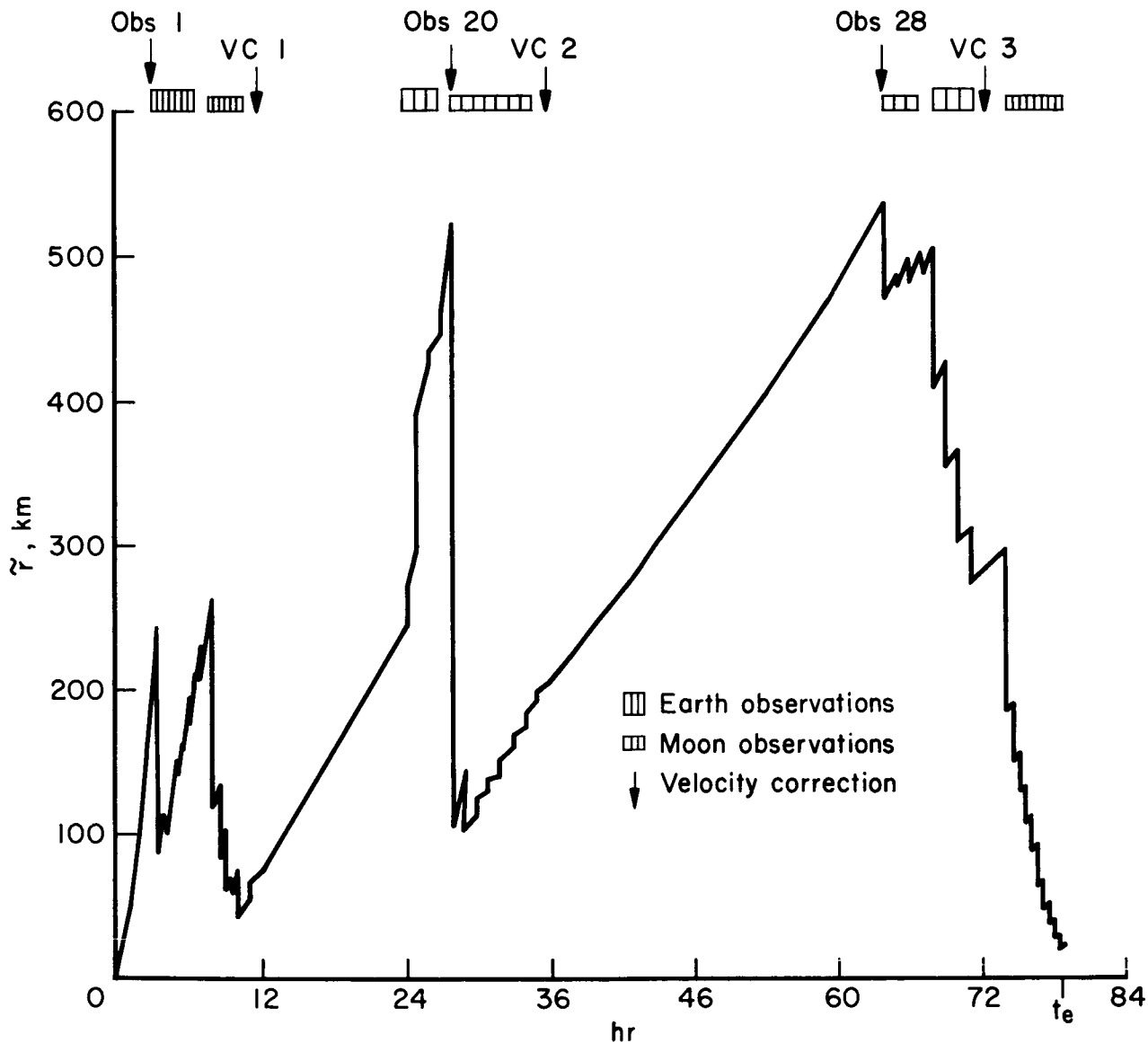


Figure 5.- $\tilde{r}$ before and after each observation for 16 bits.

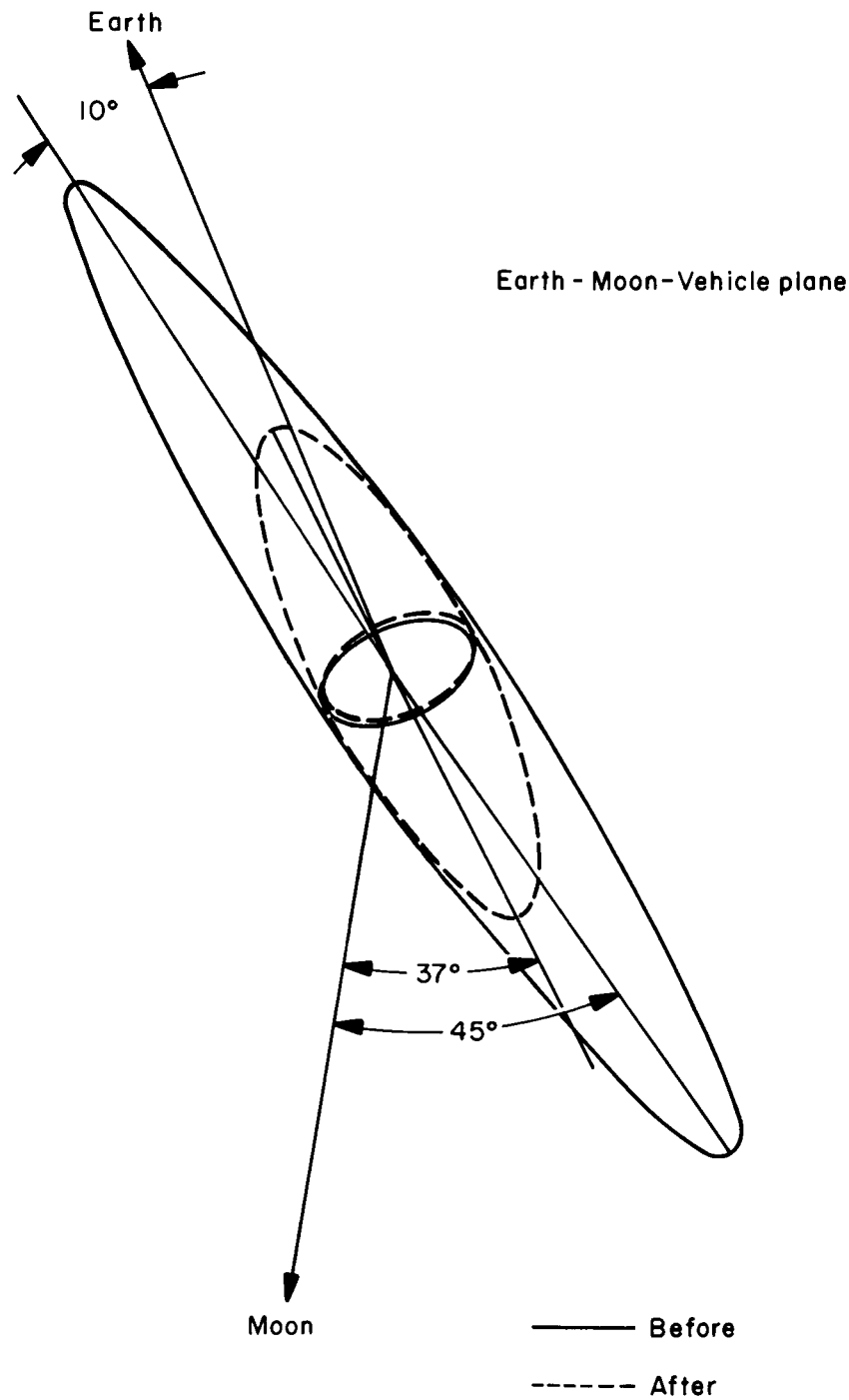


Figure 6.- Position error ellipsoid before and after moon observation 20.

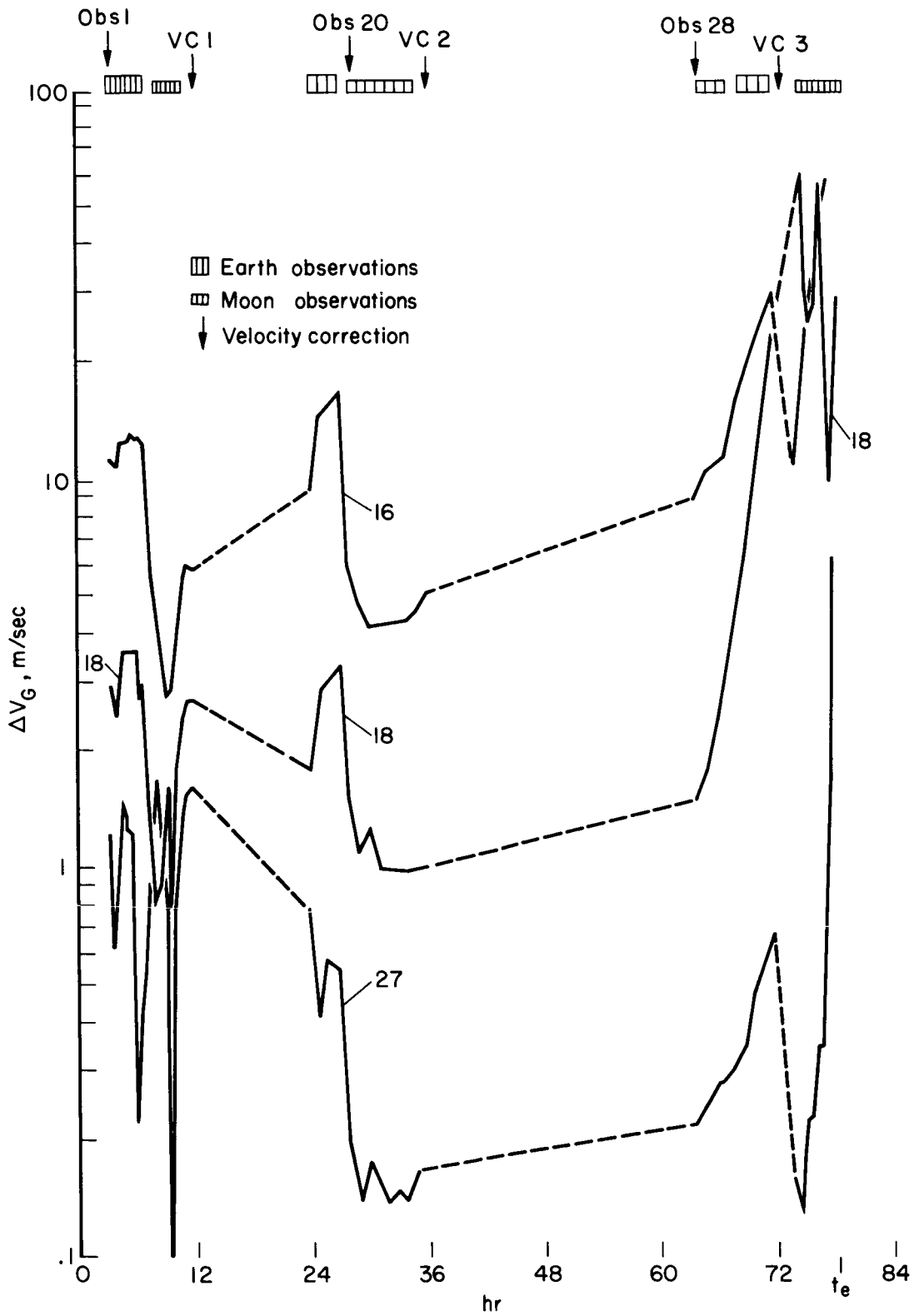


Figure 7.- Variation of $|\Delta V_G|$ with word lengths of 16, 18, and 27 bits.

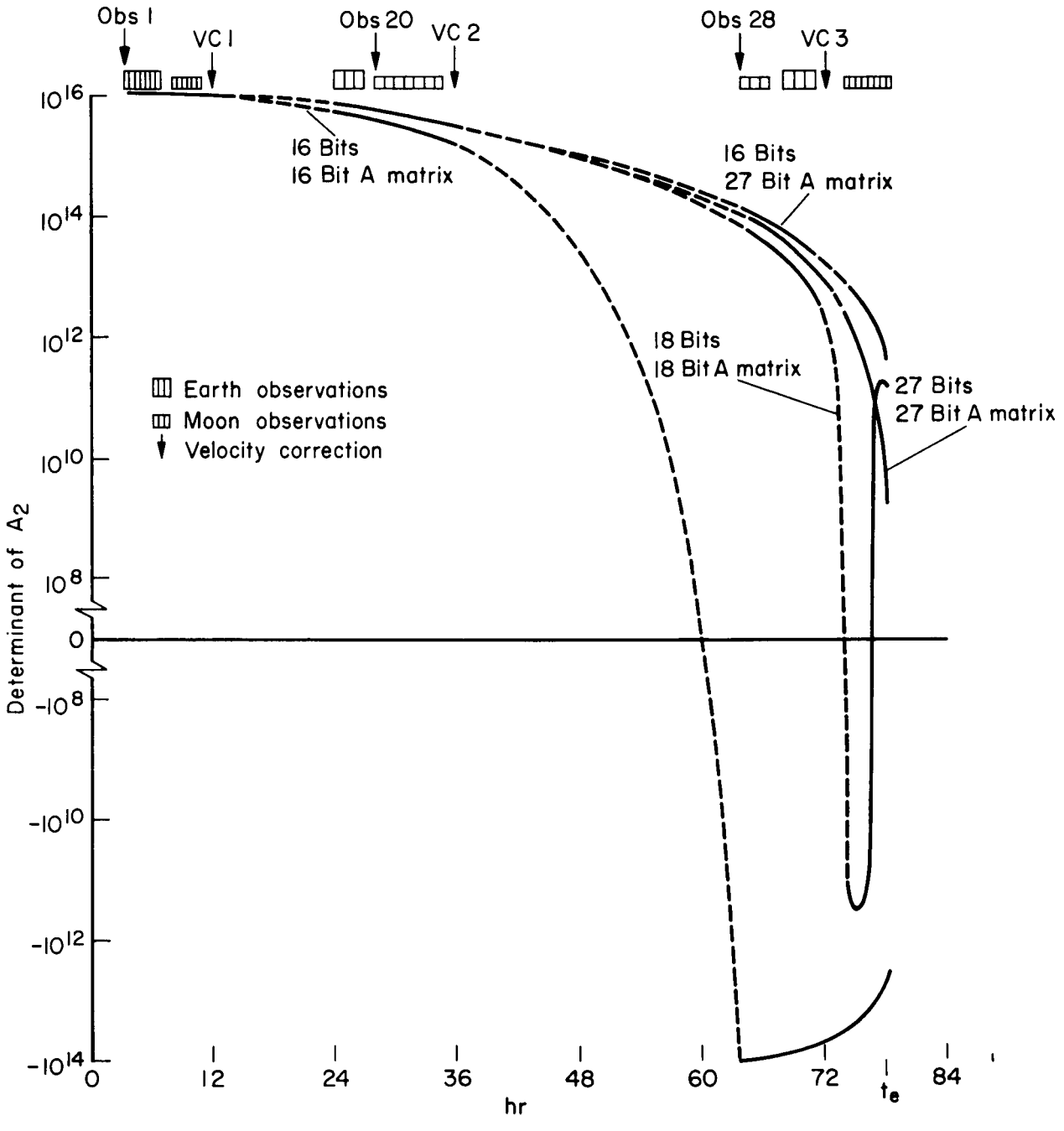


Figure 8.- Variation of the determinant of the submatrix A_2 with word length.

NASA-Langley, 1966 A-2119

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Technical information generated in connection with a NASA contract or grant and released under NASA auspices.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

TECHNICAL REPRINTS: Information derived from NASA activities and initially published in the form of journal articles.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities but not necessarily reporting the results of individual NASA-programmed scientific efforts. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION DIVISION

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546