Error Propagation in Astro-Inertial Guidance Systems for Low-Thrust Missions

HERMANN M DUSEK*
General Motors Corporation, El Segundo, Calif

The propagation of position and velocity errors caused by instrument errors in astro-inertial guidance systems for low-thrust space missions is analyzed by two methods that supplement each other For the case of orbit-sustaining in the upper atmosphere, a strict nonlinear method based on Hamilton-Jacobi's partial differential equation is employed, since the problem becomes separable in parabolic coordinates The regions of possible deviations from the unperturbed orbit are determined by a pure algebraic process and without an explicit knowledge of the perturbed orbits The influence of the second harmonic upon the region of possible motion is analyzed for the axially symmetric case using the existing integrals of motion in this nonseparable case Low-thrust missions beyond orbit-sustaining are analyzed by a first-order perturbation theory An explicit analytic solution is derived using elliptic orbits as unperturbed reference trajectories A general classification of the possible deviation based upon this explicit solution is given The errors due to the linearization are discussed by comparison with numerical integration of the exact differential equation of motion and the results based upon the separation of Hamilton-Jacobi's partial differential equation

Introduction

LOW-THRUST propulsion will very likely play an important role in many future space missions This means, from a guidance point of view, that a precise thrust acceleration measurement in the range of 10^{-3} to 10^{-6} earth-g is required over an extended flight time The applicability of any guidance system in low thrust space missions, therefore, depends strongly upon the satisfactory answer to the following questions:

1) What are the deviations from the nominal position and velocity caused by the errors in the thrust measuring instruments?

2) How can these instrument errors be detected and compensated for during the mission; possibly by the space vehicle instrumentation alone?

This paper concerns itself only with the first part of the problem The second aspect is treated in Ref 1 In the following analysis, it is assumed that the thrust acceleration is measured by accelerometers that are fixed with respect to an inertially stabilized platform Furthermore, it is assumed that appropriate measurements of a star tracker, which is also mounted on the platform, are used to compensate for the platform drift

Since the errors in drift compensation and the nominal acceleration can be considered as small quantities of first order, the errors in the acceleration measurement due to uncompensated platform drift are small quantities of second order The remaining errors of the inertial measurement unit result from the accelerometers and can be written as a power series in the true acceleration in which only the first term, the so-called bias, is small of first order All remaining error terms will be small of second or higher order For low-thrust space missions these terms can be neglected, since the uncertainties in knowledge of the universal gravitational constant and errors in the ephemeris of the planets cause position and velocity uncertainties of the same order of magnitude as the higher-order error terms The accelerometer bias constitutes, therefore, the significant error source From the previously made assumptions about the guidance system configuration, it follows that the accelerometer bias is dynamically equivalent to a homogeneous conservative force field of initially unknown magnitude and direction It is also assumed that the bias is a deterministic quantity for one instrument but a random variable with zero mean when one instrument is compared with another Only the deterministic aspects are treated in this investigation For a statistical analysis, see Ref 1

In order to analyze the problem, first-order perturbation theory could be employed This leaves open the question of the range of validity of the theory, a question of particular importance in missions with long flight times An improved first-order theory, e g, Krylov-Bogoliubov's method of averaging, could be used to discuss the long-term behavior; however, both methods determine the position and velocity deviations by an explicit calculation of the perturbed orbit, in most cases a rather involved and complex procedure This circumstance and the conservative character of the perturbing field, which characterizes the biases, make it worthwhile to search among the low-thrust missions for cases where general conclusions about the deviations can be drawn from rigorous nonlinear methods

The free-fall orbit-sustaining mission in the upper atmosphere constitutes such a class In this case Hamilton-Jacobi's partial differential equation becomes separable, and Staeckel's theory for such systems can be applied Separability implies that the region of possible deviations from the nominal orbit can be determined by a purely algebraic process This simple method is valid for field perturbations of arbitrary magnitude in the whole time interval $0 \leq t \leq \infty$ It gives a general insight into the influence of the bias error upon position deviations without an explicit determination of the perturbed orbits

The influence of the zonal harmonics in the earth's potential upon the error propagation cannot be analyzed by this method, since separability is impossible Nevertheless, for the case of circular, equatorial orbits, there exist two first integrals These are used to derive upper bounds, which are compared with the error propagation for motion in a pure central force field The result indicates that an increase in the region of possible deviations due to the second harmonic in the earth's potential can occur

An explicit first-order perturbation theory in cylindrical coordinates using elliptic orbits of arbitrary eccentricity as unperturbed references and the eccentric anomaly as inde-

Presented as Preprint 2682-62 at the ARS 17th Annual Meeting and Space Flight Exposition, Los Angeles, Calif, November 13–18, 1962; revision received February 17, 1964

* Head, Space Studies, AC Spark Plug Division Member AIAA

pendent variable is developed in the final portion of the paper in order to treat more general missions than orbit-sustaining Comparisons against numerical integration and the results obtained by separation of the variables indicate the range of validity of the theory

Region of Possible Deviations

Case 1: Hamilton-Jacobi's Differential Equation Separable

In order to investigate the error propagation for orbit-sustaining missions in the upper atmosphere, it is assumed that the gravitational potential is Kr^{-1} The more general situation, where higher harmonics are included, is treated as case 2 In order to achieve a free-fall orbit, the low-thrust propulsion has to compensate the drag forces, i e, the accelerometer output has to be zero From the symmetry of Kr^{-1}, it follows that an inertial coordinate system (x,y,z) can be chosen in such a way that the perturbing field **a**, as generated by the accelerometer biases, has the components $(0,0,-a)$ Thus the potential in which the perturbed motion takes place is given by $V = -mKr^{-1} + maz$ From the theory of separable dynamical systems,[2, 3] it is well known that Hamilton-Jacobi's partial differential equation becomes separable for a potential of the form

$$V = \frac{c_1}{r} + c_2 z + \frac{c_2}{2r} \ln\frac{r+z}{r-z} \qquad (1)$$

The potential of the present problem is included in Eq (1) if $a_2 = 0$ The proper separation variables are the parabolic coordinates (ξ,η,Φ), which are related to the Cartesian coordinates (x,y,z) by the transformation

$$\begin{aligned} x &= \xi\eta\cos\Phi & 0 \le \xi \le +\infty \\ y &= \xi\eta\sin\Phi & 0 \le \eta \le +\infty \qquad (2) \\ z &= \tfrac{1}{2}(\xi^2 - \eta^2) & 0 \le \Phi \le 2\pi \end{aligned}$$

or

$$\rho^2 = -2\xi^2[z - (\xi^2/2)] \qquad \rho^2 = 2\eta^2[z + (\eta^2/2)] \qquad (3)$$

where $\rho = (x^2 + y^2)^{1/2}$ The coordinate surfaces $\xi = \text{const}$ and $\eta = \text{const}$ are paraboloids of rotation with their foci at the origin The vertices of the paraboloids $\xi = \text{const}$ are on the positive z axis, and those of the paraboloids $\eta = \text{const}$ are on the negative z axis Separability means that the problem of integrating the equations of motion is reduced to quadratures, since a sufficient number of first integrals can be obtained In particular, the two momenta p_ξ and p_η, conjugated to the parabolic coordinates ξ and η, are found to be (for a deviation see, for instance, Ref 3)

$$p_\xi = [2m\,Wu + 2\alpha_1 - (p_{\Phi_0}{}^2/u) - \epsilon(m^2K/r_0{}^2)u^2]^{1/2} \qquad u = \xi^2$$

$$p_\eta = [2m\,Wv + 2\alpha_2 - (p_{\Phi_0}{}^2/v) + \epsilon(m^2K/r_0{}^2)v^2]^{1/2} \qquad v = \eta^2 \qquad (4)$$

$$\alpha_1 + \alpha_2 = 2m^2K$$

The dimensionless quantity ϵ is defined as

$$\epsilon = ar_o{}^2/K \qquad (5)$$

where ϵ is the ratio of the perturbing force to the central force acting on the vehicle at the beginning of the motion Thus, this parameter is a measure for the magnitude of the perturbation, and it will play a major role in the following analysis Any orbit is defined by values (ξ,η) for which p_ξ and p_η are real or $p_\xi{}^2(u) \ge 0$, $p_\eta{}^2(v) \ge 0$, and the boundaries of the motion are given by

$$p_\xi{}^2(u) = 0 \qquad p_\eta{}^2(v) = 0 \qquad (6)$$

From Eqs (3, 4, and 6), it follows that the envelopes of the

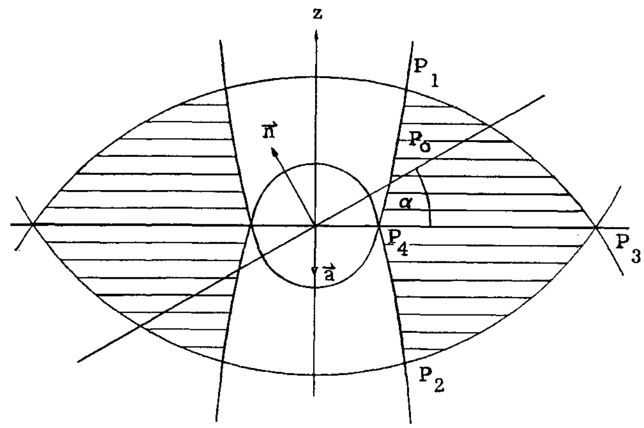

Fig 1 Upper bounds for the deviations

motion are parts of paraboloids which can be found by a purely algebraic process, namely, the discussion of two cubic equations as functions of ϵ and the constants of integration W, $p_{\Phi_0}{}^2$, α_1, and α_2 For a discussion of the different possible cases, it is advantageous to express these constants as functions of the initial position and velocity Since the problems has rotational symmetry with respect to the z axis, the coordinate system can be chosen without any loss in generality such that $\mathbf{r}_0 = (0,y_0,z_0)$ and $\dot{\mathbf{r}}_0 = (\dot{x}_0,y_0,\dot{z}_0)$, and one obtains

$$\alpha_1 = \frac{m^2}{2}\left\{2K\left(1 + \frac{z_0}{r_0}\right) + \frac{\epsilon K y_0{}^2}{r_0{}^2} + 2[y_0 y_0 \dot{z}_0 - \dot{z}_0(\dot{x}_0{}^2 + y_0{}^2)]\right\}$$

$$\alpha_2 = \frac{m^2}{2}\left\{2K\left(1 - \frac{z_0}{r_0}\right) - \frac{\epsilon K y_0{}^2}{r_0{}^2} - 2[y_0\dot{y}_0\dot{z}_0 - z_0(\dot{x}_0{}^2 + \dot{y}_0)^2]\right\} \qquad (7)$$

$$W = -\frac{mK}{r_0} + m\frac{\epsilon K_z}{r_0{}^2} + \frac{m}{2}(\dot{x}_0{}^2 + y_0{}^2 + \dot{z}_0{}^2)$$

$$p_{\Phi_0}{}^2 = m^2 y_0{}^2 \dot{x}_0{}^2$$

where W is the total energy and p_{Φ_0} the angular momentum which is a constant of motion, since the variable Φ is cyclic Orbit-sustaining in the upper atmosphere means that the unperturbed orbits are closed Thus

$$W(\epsilon = 0) < 0 \qquad (8)$$

Since there must exist values of ξ and η for the unperturbed motion where $p_\xi{}^2 > 0$ and $p_\eta{}^2 > 0$, it follows from Eqs (4) and (8) that

$$\alpha_1(\epsilon = 0) > 0 \qquad \alpha_2(\epsilon = 0) > 0 \qquad (9)$$

For the perturbed motion it follows from Eq (7) that

$$\lim_{\epsilon\to+\infty} \alpha_1(y_0 \ne 0) = +\infty \qquad \lim_{\epsilon\to+\infty} \alpha_2(y_0 \ne 0) = -\infty$$
$$\lim_{\epsilon\to+\infty} W(z_0 \ne 0) = +\infty \qquad (10)$$

A general classification of the region of possible deviations can be made depending on whether $p_{\Phi_0}{}^2$ is equal to zero or not

If $p_{\Phi_0}{}^2 \ne 0$, the orbit will be nonplanar, since there exists a velocity component perpendicular to the $\mathbf{r}_0\,\mathbf{a}$ plane at the beginning of the motion From Eq (4), it follows that

$$\lim_{u\to0} p_\xi{}^2(\epsilon \ge 0) = -\infty \qquad \lim_{v\to0} p_\eta{}^2(\epsilon \ge 0) = -\infty \qquad (11)$$

and

$$\lim_{u\to+\infty} p_\xi{}^2(\epsilon \ge 0) = -\infty \qquad \lim_{v\to+\infty} p_\eta{}^2(\epsilon = 0) = -\infty$$
$$\lim_{v\to+\infty} p\eta^2(\epsilon > 0) = +\infty \qquad (12)$$

From Eqs (4 and 10–12), one concludes the following:

1) For $p_{\Phi_0}{}^2 \neq 0$, the center of attraction is excluded from the region of possible motion for all $\epsilon \geq 0$

2) The region of possible motion is finite for $\epsilon \leq \epsilon_1$ and infinite for $\epsilon > \epsilon_1$, where ϵ_1 is a critical value which is a function of the initial conditions

For $p_{\Phi_0}{}^2 = 0$ the motion becomes planar, and the regions of possible motion exhibit a quite different behavior From Eqs (4) and (7), it follows that

$$\left.\begin{aligned}
&\lim_{u\to 0} p_\xi{}^2(\epsilon \geq 0) > 0 \\
&\lim_{u\to +\infty} p_\xi{}^2(\epsilon \geq 0) = -\infty \\
&\lim_{v\to 0} p_\eta{}^2(0 \leq \epsilon \leq \bar{\epsilon},\, y_0 \neq 0) > 0 \\
&\lim_{v\to 0} p_\eta{}^2(\epsilon > \bar{\epsilon}, y_0 \neq 0) < 0 \\
&\lim_{v\to 0} p_\eta{}^2(\epsilon \geq 0,\, y_0 = 0) > 0 \\
&\lim_{v\to +\infty} p_\eta{}^2(\epsilon > 0) = +\infty
\end{aligned}\right\} \qquad (13)$$

From Eqs (4) and (13), one concludes the following:

3) For $p_{\Phi_0}{}^2 = 0$, the center of attraction is included in the region of possible motion for all $\epsilon > 0$ if $y_0 = 0$ and for all $0 \leq \epsilon < \epsilon_2$ if $y_0 \neq 0$

4) The region of possible motion is finite for $\epsilon < \epsilon_1'$ and infinite for $\epsilon > \epsilon_1'$, where ϵ_1', ϵ_2, and $\bar{\epsilon}$ are critical values of the perturbing field which depend upon the initial conditions For a special case they are given in Eq (28)

From the theory of multiple periodic motions in separable systems it follows that, in general, in the case of a finite region of motion, the particle comes arbitrarily close at each point as time approaches infinity [4] The transition from finite to infinite regions is marked by a critical value of the perturbing field For this critical value, one of the envelopes is determined by a double root of the cubic equation $p_\eta{}^2(v) = 0$, and it can be shown[5] that the particle cannot reach this boundary in a finite time In order to obtain more detailed results, the general theory will be applied to circular nominal orbits

Circular Nominal Orbits

The initial conditions

$$\mathbf{r}_0 = (0,y_0,0) \qquad \mathbf{v}_0 = (-v_0 \cos\alpha, 0, v_0 \sin\alpha) \qquad (14)$$

are chosen The constants of integration become

$$2\alpha_1 = m^2K(2+\epsilon) \qquad 2\alpha_2 = m^2K(2-\epsilon)$$
$$W = -(mK/2r_0) \qquad (15)$$

and the cubic equations that determine the region of possible deviation can be written in the nondimensional form

$$U^3 + \gamma U^2 - (1+2\gamma)U + \gamma\cos^2\alpha = 0 \qquad U = \xi^2/r_0$$
$$(16)$$
$$V^3 - \gamma V^2 - (1-2\gamma)V - \gamma\cos^2\alpha = 0 \qquad V = \eta^2/r_0$$

where

$$\gamma = \epsilon^{-1} = K/r_0{}^2a = (g/a)(R/r_0)^2 \qquad (17)$$

R and g in Eq (17) are average values for the earth's radius and the acceleration at the surface of the earth, respectively Equation (16) shows that the cubic in V can be obtained from the cubic in U by a change of the sign in γ The problem has originally the four parameters r_0,v_0,α, and a But, as one observes from Eq (16), only the parameters γ and ϕ are essential In particular, for a given inclination α, the roots of the cubic equations will remain unchanged for all values of a and r_0 which satisfy the condition

$$(g/a)(R/r_0)^2 = \text{const} \qquad (18)$$

or, expressed in terms of initial conditions,

$$r_{0k} = r_{0i}(a_i/a_k)^{1/2} \qquad v_{0k} = v_{0i}(a_k/a_i)^{1/4} \qquad \alpha_i = \alpha_k \quad (19)$$

The indices i and k indicate two sets of initial conditions corresponding to two different perturbing fields a_i and a_k For families of parameters satisfying Eq (19), the envelopes of the region of possible motion will remain fixed in the U-V-Φ space, whereas in the physical ξ-η-Φ space a homothetic transformation

$$\xi_k = \xi_i(a_i/a_k)^{1/4} \qquad \eta_k = \eta_i(a_i/a_k)^{1/4} \qquad \Phi_k = \Phi_i \quad (20)$$

relates the envelopes belonging to a_i and a_k, respectively Because of the invariance properties, the following discussion is carried out in the U-V-Φ space Transformation in the physical space can always be achieved by using Eq. (20) The following nondimensional measures for the upper bounds in the altitude and out-of-plane variations are introduced (see Fig 1, where P_0 indicates the initial point of the orbit):

$$\Delta r^+/r_0 = (1/r_0)|r(P_3) - r(P_0)| = \tfrac{1}{2}|U_{\max} + V_{\max} - 2|$$
$$(21)$$
$$\Delta r^-/r_0 = (1/r_0)|r(P_4) - r(P_0)| = \tfrac{1}{2}|U_{\min} + V_{\min} - 2|$$

and

$$\Delta n^+/r_0 = (1/r_0)|\mathbf{n}\ \mathbf{r}(P_1)| = |[U_{\max}V_{\min}]^{1/2}\sin\alpha\cos\Phi + \tfrac{1}{2}(U_{\max} - V_{\min})\cos\alpha|$$

$$\Delta n^-/r_0 = (1/r_0)|\mathbf{n}\ \mathbf{r}(P_2)| = |[U_{\min}V_{\max}]^{1/2}\sin\alpha\cos\Phi + \tfrac{1}{2}(U_{\min} - V_{m\ x})\cos\alpha| \quad (22)$$

$$\Delta n_1{}^-/r_0 = (1/r_0)|\mathbf{n}\ \mathbf{r}(P_3)| = |[U_{m\ x}V_{\max}]^{1/2}\sin\alpha\cos\Phi + \tfrac{1}{2}(U_{\max} - V_{\max})\cos\alpha|$$

The hatched area in Fig 1 indicates the region of possible deviations It is typical for the case $p_{\Phi_0}{}^2 \neq 0$, $\epsilon < \epsilon_1$ The subscripts max and min indicate the bigger and smaller roots of the cubics which are significant for the envelopes Equations (21) and (22) show that the defined quantities are the same for all values of a and r_0 which satisfy condition (18) General explicit algebraic expressions for $\Delta r^+|r_0$ $\Delta n_1{}^-/r_0$ as a function of ϵ and α can be derived but are omitted for the sake of brevity Furthermore, the discussion of the limiting cases $\alpha = 0$, $\alpha = \pm\pi/2$ for arbitrary magnitude of ϵ indicates clearly the functional relationship between the regions of possible motion and the inclination of the plane of unperturbed motion

For $\alpha = 0$, one obtains

$$\Delta r^+/r_0 = \Delta r^-/r_0 = \epsilon + \qquad \Delta n_0{}^-/r_0 = 2\epsilon +$$
$$\Delta n_1{}^-/r_0 = \epsilon + \qquad (23)$$

and $\Delta n^+/r_0$ becomes zero In Fig 2 the intersection of the region of possible motion with a plane $\Phi = \text{const}$ is indicated by the cross-hatched area The first order series expansion will suffice in many cases, since $\epsilon \leq 10^{-5}$ However, $\gamma = \epsilon^{-1}$ will be very large, and one has to expect a strong increase in the region of possible motion even for small angles $\alpha \neq 0$ For a specific ϵ, $\Delta r^-/r_0$ is plotted in Fig 3 For $\alpha = \pm\pi/2$, one obtains in a first approximation

$$\Delta r^+/r_0 = 1 + (\epsilon^2) \qquad \Delta r^-/r_0 = 1 \qquad (24)$$

whereas $\Delta n_i{}^\pm$ $(i = 1,2)$ are zero, since $\Phi = \pi/2$ One con-

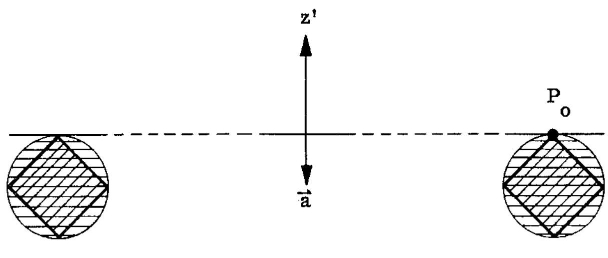

Fig 2 Regions of possible motion

cludes from Eqs (23) and (24) and from Fig 3 that, for practical purposes, the deviations are unbounded if **a** lies in the plane of the unperturbed motion or if it has an appreciable component in this plane The deviations are negligible if **a** is perpendicular to the plane of the unperturbed motion as long as ϵ is in the order of 10^{-5} earth g

So far the analysis was carried out under the assumption that the regions of possible deviations are finite Inspection of Eq (16) shows that there must exist a critical value a_c, for which only one of the roots of $p_\eta{}^2(v) = 0$ is positive That means that the region of possible motion becomes infinite for $a > a_c$ This critical value is determined by the biggest root γ_{max} of the polynomial

$$\gamma^4(4 \sin^2\alpha) - \gamma^3(36 \sin^2\alpha) - \gamma^2(18 \cos^2\alpha + 27 \cos^4\alpha - 49) - 24\gamma + 4 = 0 \quad (25)$$

Equation (25) is identical with the condition that the cubic for V possesses a double root The critical value of a is

$$a_{cr} = (1/\gamma_{max})(R/r_0)^2 \quad (26)$$

and a detailed analysis shows that a_c is, for fixed r_0, a monotonically decreasing function of α bounded between

$$\frac{g}{4 + 2(3)^{1/2}}\left[\frac{R}{r_0}\right]^2 \leq a_{cr} \leq \frac{g}{3 + 2(2)^{1/2}}\left[\frac{R}{r_0}\right]^2 \quad (27)$$

In Fig 4, a_c is plotted as a function of R/r_0 for $\alpha = 0$ and $\alpha = \pi/2$ These curves separate the (a,r_0) values that determine bounded regions from those that cause unbounded regions All separation curves corresponding to $0 < |\alpha| < \pi/2$ lie between the two plotted curves The curves indicate a rather insensitive behavior against a variation of α Figure 4 also shows that in a strict sense the deviations from a circular trajectory caused by bias errors $|\mathbf{a}| \leq 10^{-5}\, g$ will be bounded for a close-earth satellite, i e , $r_0/R \leq 5$

If the developed formulas are used to compute critical parameters ϵ_1', ϵ_2, and $\bar{\epsilon}$ as used in the general discussion, one obtains

$$\epsilon_1' = [(2 - (3)^{1/2}]/2 \qquad \epsilon_2 = [2 + (3)^{1/2}]/2 \qquad \bar{\epsilon} = 2 \quad (28)$$

ϵ_1' is equal to $\gamma_{max}{}^{-1}$ in this case Exactly the same analytical methods as applied to the circular orbit can be used for elliptical orbits

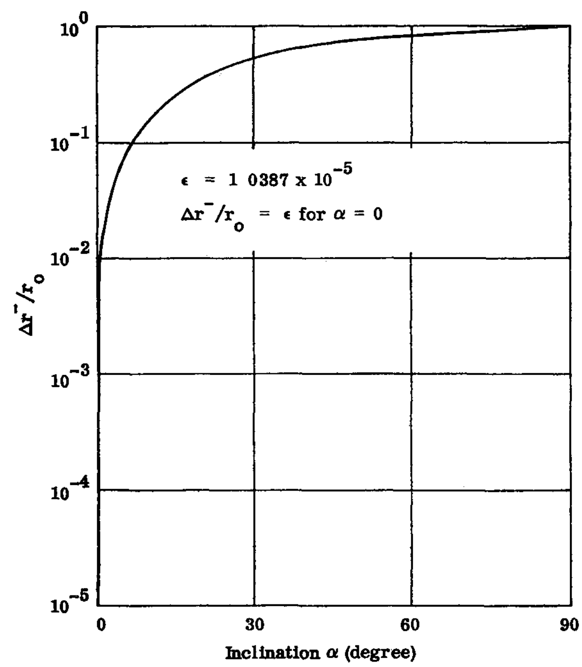

Fig 3 $\Delta r^-/r_0$ as function of inclination

Case 2: Deviations from Circular Orbits in Axially Symmetric Fields

To estimate the influence of higher harmonics in the earth's potential upon the regions of possible motion, it is assumed that the earth's potential is axially symmetric, the vector **a** is oriented along the axis of rotation, and the unperturbed orbit is a circle in the equatorial plane, i e , only even zonal harmonics in the earth's potential are considered In cylindrical coordinates (r,z,Φ), the potential is

$$V = \frac{-mK}{(r^2 + z^2)^{1/2}}\left\{1 - J_2 \frac{R^2}{r^2 + z^2} P_2(\cos\theta) + \qquad \right\} + maz \quad (29)$$

$$\cos\theta = z/(r^2 + z^2)^{1/2}$$

For the sake of brevity, higher harmonics are not included, although the method can easily be extended to that case For this dynamical problem, the Hamilton-Jacobi partial differential equation is not separable, but there still exist two first integrals of motion, namely,

$$(m/2)(\dot{r}^2 + r^2\dot{\Phi}^2 + \dot{z}^2) + V = W \qquad r^2\dot{\Phi} = p_{\Phi_0} \quad (30)$$

Consider a particle that starts at the same initial point $r_0 = (0,y_0,0)$ as in case 1 In order to obtain a circular orbit in presence of J_2, $\dot{\Phi}_0 = Kr_0{}^{-3}(1 + 1\ 5J_2R^2r_0{}^{-2})$ Using these initial conditions, one derives from Eq (30)

$$f(r',z') = \frac{r_0}{K}(\dot{r}^2 + \dot{z}^2) = -1 - \frac{1}{r'^2} + \frac{2}{(r'^2 + z'^2)} - 2\epsilon z' + J_2\frac{1}{2}\left(\frac{R}{r_0}\right)^2\left[1 - \frac{3}{r'^2} - \frac{4}{(r'^2 + z'^2)^{3/2}} P_2(\cos\theta)\right] \quad (31)$$

where $r' = r/r_0$, $z' = z/r_0$ The boundaries of the region of possible deviation are given by $f(r',z') = 0$, since $f(r',z') \geq 0$ for any actual motion For $J_2 = 0$ and $\dot{\Phi}_0 = Kr_0{}^{-3}$, the boundaries are indicated by Eq (23) (see also the double-hatched area in Fig 2) In order to obtain an explicit expression for $J_2 \neq 0$ and $\epsilon \ll 1$, new coordinates $(\bar{u},w)$ are introduced by the transformation

$$r' = \pm 1 + \bar{u} \cos w \qquad z' = \bar{u} \sin w \quad (32)$$

Considering J_2 as a small quantity of first order and expanding $f(r',z')$ to second-order terms, one obtains

$$\bar{u}^2 + 2\epsilon\bar{u} \sin w = 0 \quad (33)$$

This equation represents a circle of radius ϵ with its center located in the r'-z' plane at $(r' = 1, z' = -\epsilon)$ The situation is depicted in Fig 2 Thus, the region of possible motion is increased It is remarkable that J_2 does not enter explicitly in Eq. (33) It modifies the envelopes only indirectly by destroying the separability and thus reducing the number of first integrals The nonseparability implies that for $t \rightarrow \infty$, the particle does not necessarily approach infinitely close to each point of the region as is the case in separable systems with finite regions of possible motion

Explicit Solution by a First-Order Perturbation Theory

The previous analysis yielded the upper bounds for the deviations from the unperturbed orbit, but not the explicit functional relation between these deviations and the time This could be accomplished, for instance, in a rigorous manner starting from Eq (4) The results can be expressed by elliptic functions and are rigorously valid for an arbitrary time period However, the employed model was not completely realistic, since gravitational perturbations were neglected, and only orbit-sustaining missions were considered Furthermore, the gravitational perturbations as well as the low-thrust forces can be considered small compared to the central force field These facts dictate the development of a first-order perturbation theory

It is assumed that the dominant central force field defines the unperturbed problem Gravitational deviations from the central force field, the nominal low-thrust program, and the bias error are treated as perturbations Since first-order theory is employed, the corresponding terms appear in the solution in an additive way, and the bias error can be investigated independently from the other perturbations Spatial cylindrical coordinates (R,ϕ,Z) are used with the Z axis perpendicular to the plane of the unperturbed motion This simplifies the uncoupling of the resulting system of perturbation differential equations The dimensionless variables

$$r = R/a_0 \qquad z = Z/a_0 \qquad t = n_0 T \qquad n_0 = K/a_0^3 \tag{34}$$

and the dimensionless forces

$$F = F_R(K/a_0^2)^{-1} \qquad F_\phi = F_\phi(K/a_0^2)^{-1}$$
$$F = F_Z(K/a_0^2)^{-1} \tag{35}$$

are used Thus the semimajor axis a_0 is the unit of length, the period of the unperturbed motion is the unit of time, and the magnitude of the central force on the circle with a radius equal to the semimajor axis a_o of the unperturbed motion is the unit of force The exact differential equations of motion in the nondimensional cylindrical coordinates (r,ϕ,z) can be written as

$$\frac{d^2r}{dt^2} + \frac{r}{(r^2+z^2)^{3/2}} - r\left(\frac{d\phi}{dt}\right)^2 = \frac{\partial R_b}{\partial r} + \sum_{i=1}^{N} F_i$$
$$\frac{d}{dt}\left(r^2\frac{d\phi}{dt}\right) = \frac{\partial R_b}{\partial \phi} + \sum_{i=1}^{N} F_{i\phi} \tag{36}$$
$$\frac{d^2z}{dt^2} + \frac{z}{(r^2+z^2)^{3/2}} = \frac{\partial R_b}{\partial z} + \sum_{i=1}^{N} F_i$$

The terms containing F_i, $F_{i\phi}$, and F_i indicate the perturbations due to forces beside the bias They appear only in an additive way in the solution and will not be treated in this paper $R_b = R_b(r,\phi,z)$ is the potential function describing the bias error and can be written as

$$R_b = \epsilon_x r\cos\phi + \epsilon_y r\sin\phi + \epsilon\, z \tag{37}$$

$\epsilon_x,\epsilon_y,\epsilon_z$ are the nondimensionalized bias error components in the Cartesian coordinate system (x,y,z), which is correlated to the employed cylindrical in-plane coordinate system by $x = r\cos\phi$, $y = r\sin\phi$, $z = z$ The system (36) possesses the energy integral

$$\dot r^2 + r^2\dot\phi^2 + \dot z^2 - [2/(r^2+z^2)^{1/2}] + 1 = 2R_b + C_3 \tag{38}$$

C_3 is a constant of integration Let $r = r_0(t)$, $\phi = \phi_0(t)$, $z = z_0(t) = 0$ be the unperturbed solution For $R_b \neq 0$, let $r = r_0 + \delta r$, $\dot r = \dot r_0 + \delta\dot r$, $\phi = \phi_0 + \delta\phi$, $\dot\phi = \dot\phi + \delta\dot\phi$, $\dot z = \dot z_0 + \delta\dot z$ To derive a system of linearized differential equations, one observes first that the term containing $\dot\phi$ in the first equation of (36) can be eliminated if the differential equation is multiplied by r and added to Eq (38) This yields, after linearization,

$$\frac{d^2(r_0\delta r)}{dt^2} + \frac{1}{r_0^3}(r_0\delta r) = 2R_{b0} + r_0\left(\frac{\partial R}{\partial r}\right)_0 + C_3 \tag{39}$$

By a twofold subtraction of Eq (39) from the linearized energy integral (38), one obtains

$$\frac{d(\delta\phi)}{dt} = (1-e^2)^{1/2}\left[-\frac{d}{dt}\left(2\frac{d(r_0\delta r)}{dt} - \delta r\frac{dr_0}{dt}\right) - 3R_0 - 2r_0\left(\frac{\partial R_b}{\partial r}\right)_0 - 3\frac{C_3}{2}\right] \tag{40}$$

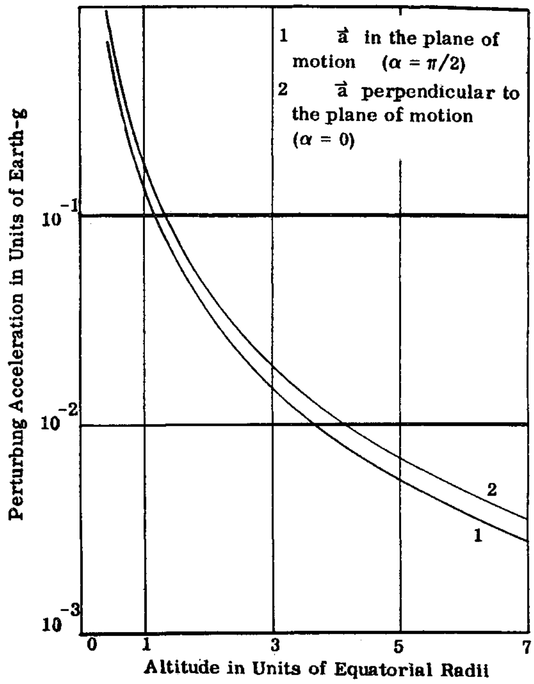

Fig 4 Critical values of $|a|$

e is the eccentricity of the unperturbed orbit The differential equation for δz follows from the third Eq (36):

$$\frac{d^2(\delta z)}{dt^2} + \frac{1}{r_0^3}\delta z = \left(\frac{\partial R_b}{\partial z}\right)_0 \tag{41}$$

The subscript zero indicates that the unperturbed trajectory must be used in evaluating the functions This follows from the assumption that ϵ_x, ϵ_y, and ϵ_z in Eq (37) are small quantities of the first order, and only first-order perturbation theory is employed Thus the right-hand sides in Eqs (39–41) are known functions of time This system of three linear differential equations with time-varying coefficient determines the error propagation It is of fifth order, but the integration problem is reduced to quadratures if the differential equation $\ddot q + r_0^{-3}q = 0$ can be solved for elliptic trajectories of arbitrary eccentricity This is indeed possible if the eccentric anomaly E is introduced as new independent variable (see, for instance, Ref 6) The solution for Eq (39) and Eq (41) can be written as

$$q^{(i)} = C_1^{(i)}q_2 + C_2^{(i)}q_2 + q_2\int_0^E \left(q_1Q_i\frac{dt}{dE}\right)_0 dE - q_1\int_0^E\left(q_2Q_i\frac{dt}{dE}\right)_0 dE \tag{42}$$

where $q^{(1)} = r_0\delta r$, $q^{(2)} = \delta z$, $q_1 = \cos E - e$, $q_2 = \sin E$, and

$$Q_1 = 2R_{b0} + r_0(\partial R/\partial r)_0 + C_3 \qquad Q_2 = (\partial R_b/\partial z)_0 \tag{43}$$

The time and the new variable are related to each other by $dt = (1 - e\cos E)dE$ The solution for $\delta\phi$ can then be found by an additional quadrature, and, since $R_b = \epsilon_x(\cos E - e) + \epsilon_y(1-e^2)^{1/2}\sin E + \epsilon\, z$, all quadratures can be expressed in rational and trigonometric functions of E The general explicit expressions are omitted, since the dependence of the error propagation upon ϵ_x, ϵ_y, and ϵ_z can be deduced from the analysis of some simple cases

Consider first the situation $\epsilon_x = \epsilon_y = 0$, $\epsilon \neq 0$, i e , the perturbing forces are perpendicular to the plane of the unperturbed motion It follows immediately from Eqs (39) and (40) that the in-plane errors vanish The boundedness or unboundedness of the out-of-plane errors are determined by the integrals

$$I_1 = \int_0^E (\cos E - e)(1 - e\cos E)dE$$
$$I_2 = \int_0^E \sin E(1 - e\cos E)dE \quad (44)$$

[see Eq (42)] This means that, according to the first-order perturbation theory, the out-of-plane errors are bounded if and only if the nominal orbit is circular For elliptic orbits, the errors are unbounded and increase proportional to $e\epsilon$ For $e \ll 1$, this will only be a second-order effect

In order to discuss the in-plane error propagation, the case of a circular nominal orbit is considered first In this case one obtains from Eqs (42) and (43)

$$\delta r = \epsilon_x[2(\cos t - 1) + \tfrac{3}{2}t\sin t] + \epsilon_y[\tfrac{3}{2}\sin t - \tfrac{3}{2}t\cos t] \quad (45)$$

Equation (45) indicates that δr, as well as $\delta\dot{r}$, is unbounded even for circular orbits Consequently $\delta\phi$ and $\delta\dot{\phi}$ are unbounded This follows immediately from Eqs (45) and (40) The increase is proportional to ϵ_x and ϵ_y For elliptical nominal orbits, the unboundedness and the proportionality to ϵ_x and ϵ_y are preserved, since the influence of the bias along such an orbit will be stronger than the influence along a circular orbit with a radius smaller than the perigee distance $a_0(1 - e)$ of the elliptic orbit The explicit analytical expressions δr, $\delta\dot{r}$, $\delta\phi$, and $\delta\dot{\phi}$ are omitted for the sake of brevity

The following general conclusions about the error propagation can be drawn:

1) The deviations will be bounded if and only if the unperturbed trajectory is circular and $\epsilon_x = \epsilon_y = 0$

2) For $\epsilon_x \neq 0$ and/or $\epsilon_y \neq 0$, the in-plane errors become unbounded for $0 \leq e < 1$ and increase proportionally to ϵ_x, ϵ_y The out-of-plane errors become unbounded for $0 < e < 1$ and increase proportionally to $e\epsilon$

The first statement does not contradict the results obtained by the method of the separation of variables, whereas the second one does But for the time intervals that are of interest for guidance purposes, the approximation will still be excellent because the increase in the perturbation solution is proportional to ϵ_x and ϵ_y, whereas the boundaries for the region of possible motion are $\Delta r^{\pm}/r_0 = 1$ for the pure in plane case [see Eq (24)] For the case $\epsilon_z \neq 0$, $e = 0$, the out-of-plane position error becomes $\delta z = -\epsilon[1 - \cos t]$ Thus $|\delta z_{m\ x}| = 2\epsilon$, and comparison with Eq (23) shows that the difference between exact and first-order theory for the maximal deviations is a second-order effect for all times It is also interesting to note that the variations in the in-plane errors are second-order effects according to the perturbation theory Actually, they are of first order as Eq (23) indicates if $t \to \infty$

The perturbation solution has been compared with results from numerical integration of the exact differential equations using

$$|\mathbf{a}| = 10^{-5}g \qquad r_0 = 3510\,085 \text{ naut miles}$$
$$R = 3444\,085 \text{ naut miles} \quad (46)$$
$$K = 2\,2488848 \times 10^8 \text{ (naut miles)}^3/\text{min}^2$$

The selected parameters correspond to an altitude of 66 naut miles above the earth, and $\epsilon = 1\,0386938 \times 10^{-5}$ The perturbation theory yields $\delta z_{m\ x} = 443\,059$ ft The numerical integration yielded 443 01197 ft The percentage error is approximately 0 01 Comparison of in-plane deviations over five revolutions resulted in an error that was always less than 0 05%

Conclusion

Even for accelerometer biases $|\mathbf{a}| = 10^{-5}g$ and moderate mission times, the position and velocity error in astro-inertial guidance systems for low-thrust space missions can become intolerably large The critical parameter is the orientation of the force field vector generated by the bias error, with respect to the plane of the unperturbed motion Small components of the perturbing field in this plane cause very large position and velocity deviations The deviations will be negligible for all times only if the perturbing field is perpendicular to the plane of the unperturbed motion This was shown rigorously for orbit-sustaining missions in the upper atmosphere using the method of separation of variables in Hamilton-Jacobi's partial differential equation The developed linear perturbation theory was compared against the rigorous results It shows sufficient validity over the time interval that is of interest for guidance purposes The solutions also suggest several methods that promise an effective error compensation, e g , rotation of the accelerometer with respect to the inertially stabilized platform, and/or employment of redundant position, and/or velocity information These techniques and the statistical aspects of the problem are discussed in Ref 1

References

[1] Dusek, H M , "Theory of error compensation in astro-inertial guidance systems for low-thrust space mission," *AIAA Progress in Astronautics and Aeronautics; Guidance and Control—II*, edited by R C Langford and C Mundo (Academic Press Inc , New York, in press), Vol 13

[2] Eisenhart, L P , *Riemannian Geometry* (Princeton University Press, Princeton, N J , 1960) Appendix 13, p 271

[3] Epstein, P , Zur Theorie des Starkeffektes " Ann Physik **50**, 498 ff (1916)

[4] Born M , *The Mechanics of the Atom* (Frederick Ungar Publishing Co , New York, 1960), p 212 ff

[5] Charlier, C L , *Die Mechanik des Himmels* (Viet, Berlin, 1927), Vol 1, p 85 ff

[6] Brouwer, D and Clemence, G M *Methods of Celestial Mechanics* (Academic Press Inc , New York, 1961), p 384 ff