Orbital Perturbations and Control by Solar Radiation Forces

J.C. Van der Ha* and V.J. Modi†
The University of British Columbia, Vancouver, B.C., Canada

Solar radiation pressure induced orbital perturbations are analyzed for a satellite in an elliptic orbit of arbitrary inclination with respect to the ecliptic. The problem is formulated in terms of the two-variable expansion procedure as well as a straightforward perturbation/rectification scheme leading to results valid over a long duration. The vast amount of information generated by a systematic variation of design parameters is condensed in the form of polar plots, from which the long-term in-plane perturbations can easily be visualized. The results show that, in the first order, the semimajor axis remains unchanged in the long run, whereas the eccentricity changes in a periodic fashion. The longitude of the ascending node shows a secular variation which is essentially insensitive to changes in the initial eccentricity. On the other hand, the orbital inclination markedly depends on the initial eccentricity and solar aspect angle. Subsequently, results of controlled on/off switching of the solar radiation force, using several different strategies, are given which represent effective procedures for orbital transfer. A suitable choice of switching strategy can lead to an increase in the semimajor axis by a factor of 10 in less than five years for a spacecraft with an area/mass ratio of 5 m^2/kg. Advantages rendered by this capability in terms of scientific exploration, escape, and launch into heliocentric orbits are apparent.

Nomenclature

a = semimajor axis
a_r = semimajor axis of reference orbit (24-hr period), 42,241 km
c = velocity of light, 2.998×10^8 m/s
c_ϵ = constant, ϵ/δ
e = eccentricity
i = inclination of satellite's orbit with respect to the ecliptic
j = stands for any arbitrary orbital element (including η)
l = semilatus rectum
m = mass of satellite
p = $e\cos\tilde{\omega}$, orbital element
q = $e\sin\tilde{\omega}$, orbital element
r = radial vector of satellite with respect to earth, Fig. 1
t = time
u^s = unit vector along solar radiation
u_r, u_ϕ, u_p = components of u^s along radial, circumferential and orbit-normal directions, Eq. (2)
w_r, w_ϕ, w_p = components of auxiliary rotation vector, Eq. (7)
x,y = auxiliary elements, $e\sin(\tilde{\eta}-\tilde{\omega})$ and $e\cos(\tilde{\eta}-\tilde{\omega})$, respectively
A = total reflective area of satellite and plates; cross-sectional area of spherical satellites
A_{nk}, B_{nk} = integrals, Eqs. (13)
F = solar radiation force, Eq. (1)
$F(e,\lambda)$ = function, Eq. (21)
K_1, K_2 = functions, Eq. (12)
K_{10}, K_{20} = constants, Eq. (16)
K_{30}, K_{40} = constants, Eq. (19)
S = solar constant, amount of solar energy incident upon a unit area normal to the radiation at 1 A.U. from the Sun, 1.35 kW/m^2
S' = solar radiation pressure, S/c, 4.51×10^{-6} N/m^2
α = angle between switching locations and the Sun-Earth line, Eq. (22) and Fig. 2
δ = angular rate of the Sun in the ecliptic, 1/365.2422
ϵ = ratio of solar radiation and gravity forces, Eq. (4)
η = solar aspect angle as measured from autumnal equinox, Fig. 1
$\tilde{\eta}$ = $\eta-\Omega$, solar aspect angle as measured from axis $\nu=0$ (eclipitic orbit only)
θ = $\phi-\omega$, position of satellite as measured from instantaneous perigee axis
κ = material parameter, $1-\tau+\rho$ for plate or $1-\tau$ for sphere: plate with perfect specular reflection, $\kappa=2$; sphere with perfect specular reflection or perfect absorption, $\kappa=1$; sphere with perfect diffuse reflection, $\kappa=1.44$
λ = $\tilde{\eta}-\tilde{\omega}$, angle between position of Sun and perigee axis, Fig. 2 (ecliptic orbit only)
μ = Earth's gravitational parameter, $3.986 \times 10^{14} m^3/s^2$
ν = quasiangle used as independent variable, Eq. (5) and Fig. 1
ν_f = length of interval before rectification
ν_j = switching points ($j=1,\ldots,6$), Fig. 2
ρ = reflectivity of satellite (constant)
τ = transmissivity of satellite (constant)
ϕ = position angle of satellite as measured from line of nodes, Fig. 1
ψ = $\nu-\phi$, Fig. 1
ω = argument of perigee as measured from line of nodes
$\tilde{\omega}$ = $\omega+\psi$, argument of perigee with respect to axis $\nu=0$
Δj = first-order changes in arbitrary element j
Ω = longitude of nodes with respect to the vernal equinox

Presented as Paper 77-32 at the AIAA 15th Aerospace Sciences Meeting, Los Angeles, Calif., Jan. 24-26, 1977; submitted Feb. 8, 1977; revision received Sept. 19, 1977.

Index categories: Earth-Orbital Trajectories; Spacecraft Navigation, Guidance, and Flight-Path Control.

*Graduate Research Fellow, Dept. of Mechanical Engineering. Member AIAA.

†Professor, Dept. of Mechanical Engineering. Member AIAA.

$(\dot{}), (\,)'$ = differentiation with respect to t and ν, respectively

Subscript

00 = initial conditions

Introduction

EVER since the first exhibition of solar radiation perturbation effects upon an Earth satellite (the Vanguard I, launched on March 17, 1958), many aspects of solar radiation induced orbital perturbations have been studied.[1,2] Especially, the anomalous behavior of the Echo I balloon satellite whose perigee varied between about 900 and 1500 km above the Earth, received a great deal of attention.[3,4]

The feasibility of utilizing solar radiation forces for controlled orbital transfer was assessed quite early in the space age. In 1958, Garwin[5] envisioned an exploration of the planetary system by means of large solar sails made of aluminized Mylar. The concept was further substantiated in subsequent studies.[6-9] In a geocentric configuration, Sands[10] proposed to rotate the sail about an axis perpendicular to the orbital plane at half the orbital rate of the satellite around the planet. This strategy enables the satellite to reach escape trajectory eventually. Considering an orbit normal to the ecliptic, Fimple[11] determined the control strategy which maximizes the solar radiation force component along the instantaneous velocity. This strategy proves to be particularly effective as it maximizes the rate of energy increase. On the other hand, Cohen et al.[12] have achieved orbital change in the ecliptic plane by means of an on/off switching program for the solar sail. During the on-phase, when the satellite moves away from the Sun, the plate is aligned with the radius vector and perpendicular to the orbital plane, while the off-phase is characterized by the plate being aligned with the solar radiation. This switching policy leads to a substantial increase in angular momentum and thus in the semilatus rectum.

In this paper, a satellite in an arbitrary geocentric orbit is considered while the solar radiation force has a constant magnitude and is directed along the instantaneous Sun-Earth line. This model is sufficiently realistic for communications satellites with large solar panels kept normal to the radiation for maximum on-board power production. Also, spherical satellites with homogeneous material characteristics would be covered by this formulation. A few approaches for studying the long-term orbital behavior of the spacecraft are formulated, namely the two-variable expansion procedure[13,14] and the straightforward perturbation method in conjunction with repeated rectification. The former approach yielded interesting results in small-thrust problems[15-17] and led to long-term analytical approximations for the orbital elements under solar radiation forces for an orbit in the ecliptic plane.[18]

After establishing the long-term perturbations under continuous exposure to sunlight pressure, the responses of a few on/off switching strategies are established. During the on-phase, the plate is assumed to be normal to the incident radiation while the off-phase is characterized by the plate being aligned with the Sun-Earth line. Alternatively, the analysis is valid for a Mylar-coated plastic sphere with a pumping device for inflating and deflating the balloon.

The results established here are of interest for orbital correction and transfer of a large class of spacecraft. For instance, stationkeeping of communications satellites with controllable solar arrays could be accomplished in this manner. Their normal operation would remain unaffected during the on-phase, i.e. about half the time. The switching strategies are also relevant for raising a solar sail or Satellite Solar Power Station (SSPS) from a low or intermediate geocentric orbit into a heliocentric trajectory by means of solar radiation forces.

In the analysis no restrictions have been placed on the eccentricity of the initial nor of the ensuing osculating ellipses and the singularity in the equation for the argument of the perigee is avoided by a suitable transformation. The solar parameter ϵ, denoting the ratio of solar radiation and Earth's gravity force at geosynchronous altitude, is assumed to be of the same order of magnitude as δ, the angular rate of the Sun in the ecliptic. The solar aspect ratio angle is effectively treated as if it were an orbital element.

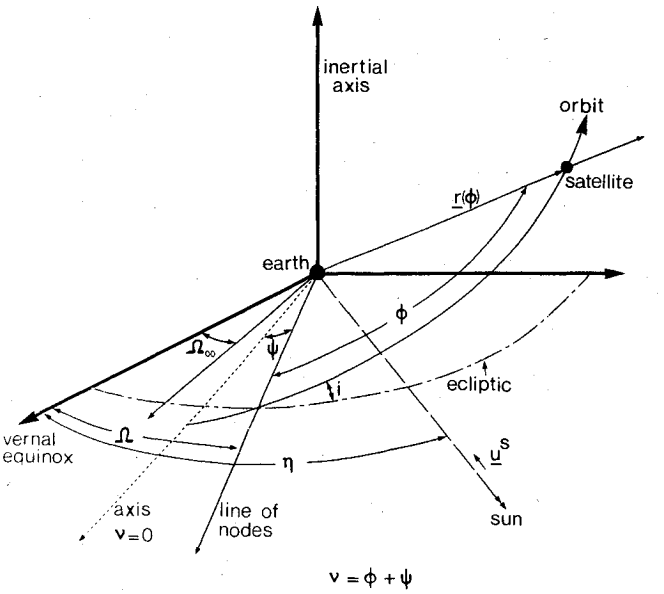

Fig. 1 General three-dimensional configuration of Earth, satellite, and Sun.

Formulation of the Problem

Consider a satellite in an arbitrary orbit as shown in Fig. 1. The inertial reference is represented by the vernal equinox, the direction perpendicular to it in the ecliptic and that normal to the ecliptic. The position of the satellite is specified by the vector $r(\phi)$ where ϕ is measured from the line of nodes. Assuming homogeneous reflecting properties, the solar radiation force upon a satellite of the balloon type or one modelled as a plate normal to the Sun in an orbit around the Earth, is of the form

$$F = S' A \kappa u^s \tag{1}$$

with u_r, u_ϕ, u_p representing components of the unit vector u^s in radial, circumferential, and orbit-normal directions:

$$\begin{aligned} u_r &= -[\cos\phi\cos(\eta-\Omega) + \sin\phi\cos i\sin(\eta-\Omega)] \\ u_\phi &= [\sin\phi\cos(\eta-\Omega) - \cos\phi\cos i\sin(\eta-\Omega)] \\ u_p &= \sin i\sin(\eta-\Omega) \end{aligned} \tag{2}$$

Usually the instantaneous orbital elements are described in terms of the classical Lagrange's perturbation equations involving the variables l, e, ω, i, and Ω. In nondimensional form, these equations can be written as

$$\begin{aligned} \dot{l} &= 2\epsilon r l^{1/2} u_\phi \\ (\dot{i}) &= \epsilon r\cos\phi\, u_p / l^{1/2} \\ \dot{\Omega} &= \epsilon r\sin\phi\, u_p / (l^{1/2}\sin i) \\ \dot{e} &= \epsilon l^{1/2}\{u_r\sin\theta + [(\cos\theta+e)r/l + \cos\theta]u_\phi\} \\ \dot{\omega} &= \epsilon l^{1/2}\{-u_r\cos\theta + [(2+e\cos\theta)\sin\theta]u_\phi r/l \\ &\quad - e\sin\phi\cot i\, u_p r/l\}/e \end{aligned} \tag{3}$$

where $\dot{\phi} = l^{1/2}/r^2 - \dot{\Omega}\cos i$ and $\theta = \phi - \omega$.

The reference length and time units are a_r and $(a_r^3/\mu)^{1/2}$, respectively; the small nondimensional parameter ϵ designates

the ratio of the solar radiation and gravitational forces,

$$\epsilon = S'(A/m)\kappa(a_r^2/\mu) \tag{4}$$

The form of the system of Eqs. (3) is not particularly convenient in exploring analytical representations for the perturbed orbital elements. One problem concerns the choice of a suitable independent variable. In this regard, the quasiangle ν defined by the differential relation,

$$\dot{\nu} = \dot{\phi} + \dot{\Omega}\cos i = l^{1/2}/r^2 \qquad \nu(0) = 0 \tag{5}$$

turns out to be very attractive both in terms of mathematical convenience and physical interpretation of results. Another difficulty is the singularity in $\dot{\omega}$ when the instantaneous orbit becomes circular. Therefore, a different formulation is explored. The spatial equations for orbital motion under the influence of the solar radiation force may be obtained from Newton's second law, while accounting for the rotation of the local reference axes in the inertial frame,

$$\ddot{r} + 1/r^2 + rw_p^2 = \epsilon u_r$$

$$2\dot{r}w_p + r\dot{w}_p = \epsilon u_\phi$$

$$rw_p w_r = \epsilon u_p \tag{6}$$

where

$$\begin{pmatrix} w_r \\ \hline w_\phi \\ \hline w_p \end{pmatrix} = \begin{pmatrix} \dot{\Omega}\sin\phi\sin(i) + (i)\cos\phi \\ \hline \dot{\Omega}\cos\phi\sin(i) - (i)\sin\phi \\ \hline \dot{\phi} + \dot{\Omega}\cos i = \dot{\nu} \end{pmatrix} \tag{7}$$

Using the transformation $u = 1/r$ as in the derivation of the unperturbed Kepler ellipse and writing $u(\nu) = (1 + p\cos\nu + q\sin\nu)/l$ with osculating elements p, q, and l, a closed system of first-order differential equations in terms of the independent variable ν is obtained from Eqs. (5-7):

$$\begin{aligned}
l'(\nu) &= 2\epsilon r^3 u_\phi \\
p'(\nu) &= \epsilon r^2 \{u_r\sin\nu + u_\phi[\cos\nu + (p+\cos\nu)r/l]\} \\
q'(\nu) &= \epsilon r^2 \{-u_r\cos\nu + u_\phi[\sin\nu + (q+\sin\nu)r/l]\} \\
\Omega'(\nu) &= \epsilon r^3 u_p \sin(\nu-\psi)/(l\sin i) \\
i'(\nu) &= \epsilon r^3 u_p \cos(\nu-\psi)/l \\
\psi'(\nu) &= \epsilon r^3 u_p \sin(\nu-\psi)\cot i/l = \Omega'(\nu)\cos i
\end{aligned} \tag{8}$$

The elements $p = e\cos\tilde{\omega}$ and $q = e\sin\tilde{\omega}$ are preferred to e and ω in order to have a uniformly valid representation for all eccentricities, including $e = 0$ where ω becomes indeterminate. Physically, the angle $\tilde{\omega} = \omega + \psi$ defines the position of the instantaneous perigee as measured from the axis $\nu = 0$. It may be pointed out that the apparent singularity in the equation for Ω for small i is of no consequence as $\sin i$ cancels due to its appearance in u_p. It is convenient to treat the solar aspect angle η as a quasiorbital element defined by the differential equation

$$\eta'(\nu) = \delta r^2/l^{1/2} \qquad \eta(0) = \eta_{00} \tag{9}$$

with $\delta = 1/365.2422$ denoting the apparent angular rate of the Sun in the ecliptic plane.

Two-Variable Expansion Procedure

Assuming that ϵ is of the same order of magnitude as δ, the orbital elements and the solar aspect angle η are expanded as

$$j(\nu) = j_0(\nu,\bar{\nu}) + \epsilon j_1(\nu,\bar{\nu}) + \epsilon^2 j_2(\nu,\bar{\nu}) + \ldots \tag{10}$$

where j denotes any one of the elements p, q, l, Ω, i, ψ, η, and $\bar{\nu}$ stands for $\epsilon\nu$. Substituting the series above into Eqs. (8) and (9), and collecting terms of like order in ϵ gives for the zeroth order: $j_0(\nu,\bar{\nu}) = j_0(\bar{\nu})$ and $j_0(0) = j_{00}$. The flexibility gained through addition of the independent slow variable $\bar{\nu}$ shall be utilized by requiring that the first-order functions $j_1(\nu,\bar{\nu})$ remain bounded in ν. Thereto, the right-hand sides of the first-order equations are expanded in terms of Fourier series with slowly varying coefficients. To eliminate unbounded contributions, the nonharmonic terms must vanish leading to the following set of equations for the zeroth-order approximations:

$$\begin{aligned}
dp_0/d\bar{\nu} &= (l_0^2/2\pi)\{K_1[p_0 B_{31}(2\pi) + \tfrac{1}{2}B_{32}(2\pi)] \\
&\quad - K_2[A_{20}(2\pi) + p_0 A_{31}(2\pi) + \tfrac{1}{2}A_{30}(2\pi) \\
&\quad + \tfrac{1}{2}A_{32}(2\pi)]\} \\
dq_0/d\bar{\nu} &= -(l_0^2/2\pi)\{K_2[q_0 A_{31}(2\pi) + \tfrac{1}{2}B_{32}(2\pi)] \\
&\quad - K_1[A_{20}(2\pi) + q_0 B_{31}(2\pi) + \tfrac{1}{2}A_{30}(2\pi) \\
&\quad - \tfrac{1}{2}A_{32}(2\pi)]\} \\
dl_0/d\bar{\nu} &= 2(l_0^3/2\pi)[K_1 B_{31}(2\pi) - K_2 A_{31}(2\pi)] \\
d\Omega_0/d\bar{\nu} &= (l_0^2/2\pi)\sin(\eta_0 - \Omega_0)[\cos\psi_0 B_{31}(2\pi) \\
&\quad - \sin\psi_0 A_{31}(2\pi)] \\
di_0/d\bar{\nu} &= (l_0^2/2\pi)\sin i_0 \sin(\eta_0 - \Omega_0)[\cos\psi_0 A_{31}(2\pi) \\
&\quad + \sin\psi_0 B_{31}(2\pi)] \\
d\psi_0/d\bar{\nu} &= \cos i_0\, d\Omega_0/d\bar{\nu} \\
d\eta_0/d\bar{\nu} &= \frac{(\delta/\epsilon)l_0^{3/2}}{(2\pi)A_{20}(2\pi)}
\end{aligned} \tag{11}$$

Here K_1 and K_2 represent functions of the slow variable $\bar{\nu}$, defined by the relations

$$\begin{aligned}
K_1(\bar{\nu}) &= \sin^2(i_0/2)\cos(\eta_0 - \Omega_0 - \psi_0) \\
&\quad + \cos^2(i_0/2)\cos(\eta_0 - \Omega_0 + \psi_0) \\
K_2(\bar{\nu}) &= \cos^2(i_0/2)\sin(\eta_0 - \Omega_0 + \psi_0) \\
&\quad - \sin^2(i_0/2)\sin(\eta_0 - \Omega_0 - \psi_0)
\end{aligned} \tag{12}$$

The integrals,

$$\begin{aligned}
A_{nk}(p,q;\nu) &= \int_0^\nu \frac{\cos(k\tau)\,d\tau}{(1 + p\cos\tau + q\sin\tau)^n} \\
B_{nk}(p,q;\nu) &= \int_0^\nu \frac{\sin(k\tau)\,d\tau}{(1 + p\cos\tau + q\sin\tau)^n}
\end{aligned} \tag{13}$$

for $n = 1,2,3,\ldots$, $k = (0), 1, 2, \ldots$, and $p^2 + q^2 < 1$ can be evaluated in a straightforward manner by means of recurrence relations.[19] Note that $A_{nk}(2\pi)$ stands for $A_{nk}[p_0(\bar{\nu}), q_0(\bar{\nu}); 2\pi]$ in Eq. (11). Upon substitution of the resulting expressions for the integrals of Eqs. (13) into Eqs. (11), a system of nonlinear first-order differential equations is obtained. The semimajor axis $a_0(\bar{\nu})$ is related to l_0, p_0 and q_0 giving

$$\frac{da_0}{d\bar{\nu}} = \frac{l_0'(\bar{\nu})}{(1 - e_0^2)} + \frac{2l_0[p_0 p_0'(\bar{\nu}) + q_0 q_0'(\bar{\nu})]}{(1 - e_0^2)^2} = 0 \tag{14}$$

so that $a_0(\bar{\nu}) = a_{00}$ and over a long time the total energy of the satellite remains conserved (in zeroth-order approximation).

Physically, the energy added while moving away from the Sun balances that removed during the motion towards the Sun in the present approximation.

Perturbation/Rectification Procedure

Although analytical solutions have not been obtained for the orbital elements described in Eqs. (11), numerical integration could readily be performed. However, an alternate convenient and highly effective procedure is available. Recognizing that the rate of change in the orbital elements is small, a straightforward perturbation expansion of the form $j(\nu)=j_{00}+\Delta j(\nu)+\ldots$ provides a valid approximation over some limited interval of ν. Using the conventional approach, the first-order perturbations can be calculated:

$$\begin{aligned}
\Delta p(\nu) &= \epsilon l_{00}^2\{K_{10}[p_{00}B_{31}(\nu)+\tfrac{1}{2}B_{32}(\nu)] \\
&\quad -K_{20}[A_{20}(\nu)+p_{00}A_{31}(\nu)+\tfrac{1}{2}A_{30}(\nu)+\tfrac{1}{2}A_{32}(\nu)]\} \\
\Delta q(\nu) &= -\epsilon l_{00}^2\{K_{20}[q_{00}A_{31}(\nu)+\tfrac{1}{2}B_{32}(\nu)] \\
&\quad -K_{10}[A_{20}(\nu)+q_{00}B_{31}(\nu)+\tfrac{1}{2}A_{30}(\nu)-\tfrac{1}{2}A_{32}(\nu)]\} \\
\Delta l(\nu) &= 2\epsilon l_{00}^3[K_{10}B_{31}(\nu)-K_{20}A_{31}(\nu)] \\
\Delta\Omega(\nu) &= \epsilon l_{00}^2\sin(\eta_{00}-\Omega_{00})[\cos\psi_{00}B_{31}(\nu)-\sin\psi_{00}A_{31}(\nu)] \\
\Delta i(\nu) &= \epsilon l_{00}^2\sin i_{00}\sin(\eta_{00}-\Omega_{00})[\cos\psi_{00}A_{31}(\nu) \\
&\quad +\sin\psi_{00}B_{31}(\nu)] \\
\Delta\psi(\nu) &= \cos i_{00}\Delta\Omega(\nu) \\
\Delta\eta(\nu) &= \delta l_{00}^{3/2}A_{20}(\nu)
\end{aligned} \tag{15}$$

where K_{10} and K_{20} are constants defined as follows:

$$\begin{aligned}
K_{10} &= \sin^2(i_{00}/2)\cos(\eta_{00}-\Omega_{00}-\psi_{00}) \\
&\quad +\cos^2(i_{00}/2)\cos(\eta_{00}-\Omega_{00}+\psi_{00}) \\
K_{20} &= \cos^2(i_{00}/2)\sin(\eta_{00}-\Omega_{00}+\psi_{00}) \\
&\quad -\sin^2(i_{00}/2)\sin(\eta_{00}-\Omega_{00}-\psi_{00})
\end{aligned} \tag{16}$$

The expressions in Eqs. (15) are perfectly suitable for obtaining changes in the orbital elements for time spans of the order of one or even a few revolutions.

To obtain changes in the elements valid over a long duration, a certain interval $(0, \nu_f)$ is chosen and the elements are updated (or rectified) at $\nu=\nu_f$ by putting $(j_{00})_{\text{rect}}=j_{00}+\Delta j(\nu_f)$. This iteration procedure can be performed by a digital computer at a considerable saving of time and effort compared to a numerical integration of the original equations or the zeroth order Eqs. (11) obtained through the two-variable expansion method. It should be noted that in practice it is convenient to take the interval ν_f equal to π, 2π, etc. because this leads to relatively simpler expressions for the integrals. Comparison with numerical results obtained by double-precision Runge-Kutta integration of the exact system of Eqs. (8) suggests that rectification after one revolution ($\nu_f=2\pi$) results in two significant digits for the eccentricity over a period of 1200 days when $\epsilon=0.0002$ is taken. This compares closely with the accuracy of three decimal places achieved through the two-variable expansion procedure in the case of an orbit in the ecliptic.[18]

It is interesting to compare the structure of Eqs. (15) for $\nu=2\pi$ with Eqs. (11). Whereas the straightforward perturbation method in Eqs. (15) simply calculates changes $\Delta j(2\pi)$ by keeping all elements constant ($j=j_{00}$) during the integration, Eqs. (11) are still in differential form with respect to the slow variable $\bar{\nu}$. One may interpret the zeroth-order two-variable Eqs. (11) as obtained from rectification of the perturbation Eqs. (8) with the dependence upon ν eliminated by averaging and an infinitesimal interval before rectification, $d\bar{\nu}$. Consequently, the zeroth-order two-variable results should be expected to yield better approximations to the exact solution than those from rectification (after one revolution) and iteration of the first-order straightforward perturbation solution. On the other hand, in order to improve upon a certain accuracy, one needs to solve for the *higher order* terms in case of two-variable expansion procedure, while the accuracy of the perturbation/iteration method can be enhanced by simply choosing a smaller interval before rectification of the *first-order* results. Finally, it may mentioned that Eqs. (11) can also be obtained through the classical method of averaging, since the nonharmonic term in the Fourier expansion represents the average value of the function under consideration.

The results from the perturbation/rectification approach with $\nu_f=2\pi$ can be reduced, after substitution of the integrals, to the following compact form:

$$\begin{aligned}
\Delta p &= -3\pi\epsilon a_{00}^2K_{20}(1-e_{00}^2)^{1/2} \\
\Delta q &= 3\pi\epsilon a_{00}^2K_{10}(1-e_{00}^2)^{1/2} \\
\Delta l &= 6\pi\epsilon a_{00}^2[K_{20}p_{00}-K_{10}q_{00}](1-e_{00}^2)^{1/2} \\
\Delta\Omega &= 3\pi\epsilon a_{00}^2\sin(\eta_{00}-\Omega_{00})[p_{00}\sin\psi_{00} \\
&\quad -q_{00}\cos\psi_{00}]/(1-e_{00}^2)^{1/2} \\
\Delta i &= -3\pi\epsilon a_{00}^2\sin i_{00}\sin(\eta_{00}-\Omega_{00})[p_{00}\cos\psi_{00} \\
&\quad +q_{00}\sin\psi_{00}]/(1-e_{00}^2)^{1/2} \\
\Delta\psi &= \cos i_{00}\Delta\Omega \\
\Delta\eta &= 2\pi\delta a_{00}^{3/2}
\end{aligned}$$

It should be emphasized that $\Delta a=0$, meaning that the total energy (and thus the major axis) returns to its original value after one revolution. For completeness the results for Δe and $\Delta\tilde{\omega}$ (if $e_{00}\neq 0$) are given:

$$\begin{aligned}
\Delta e &= (p_{00}\Delta p+q_{00}\Delta q)/e_{00}=3\pi\epsilon a_{00}^2K_{30}(1-e_{00}^2)^{1/2} \\
\Delta\tilde{\omega} &= (p_{00}\Delta q-q_{00}\Delta p)/e_{00}^2=3\pi\epsilon a_{00}^2K_{40}(1-e_{00}^2)^{1/2}/e_{00}
\end{aligned} \tag{18}$$

with the constants K_{30} and K_{40} given by

$$\begin{aligned}
K_{30} &= K_{10}\sin\tilde{\omega}_{00}-K_{20}\cos\tilde{\omega}_{00} \\
K_{40} &= K_{10}\cos\tilde{\omega}_{00}+K_{20}\sin\tilde{\omega}_{00}
\end{aligned} \tag{19}$$

Orbital Control by Switching

With some appreciation as to the orbital perturbations due to solar radiation pressure, the next logical step would be to explore the possibility of achieving desired orbital changes by means of the solar radiation force. However, the opportunities are limited in the case of a continuously acting force, directed along the Sun-Earth line because the average change in the semimajor axis over a long period vanishes. This severely restricts the character of attainable target orbits. On the other hand, substantial changes in all orbital elements can be obtained if the force were to be switched off during a certain part of the orbit. The switchoff may be accomplished by aligning the plate parallel to the radiation (or by deflation in case of a balloon satellite), while the on-phase represents plate normal to the radiation. The switching instants can be chosen so as to meet a given objective, e.g., one may want to

enlarge the orbit by keeping changes in the semilatus rectum positive during the on-phase. The switching may be assumed to occur instantaneously as the time taken to complete the maneuver represents a negligible fraction of the orbital period. For a given strategy, changes in the orbital elements can be calculated by integration of Eqs. (8) over the relevant interval. In the cases considered here, the interval is always less than $3\pi/2$, and sufficiently accurate results are obtained by taking the elements on the right-hand side of Eqs. (8) to be constant during integration. This corresponds to expanding the elements in a first-order straightforward perturbation series as discussed in the previous section. In general, changes in the elements are determined from Eqs. (15) by taking the difference $\Delta j(\nu_i) - \Delta j(\nu_{i-1})$. In all, three procedures were attempted with the on/off control activated at (Fig. 2): 1) the Sun-Earth line (points 1 and 2), 2) positions where the velocity vector is normal to the Sun-Earth line (points 3 and 4) 3) apogee and perigee (points 5 and 6). Although the problem can be formulated and solved numerically for general three-dimensional orbital control, here several simpler cases for an ecliptic orbit are considered to illustrate the effectiveness. Furthermore, they yield results which can be expressed in compact analytical formulas and hence give better appreciation of the control process as to the various input parameters.

A. Increase in the Semilatus Rectum

From Eqs. (8) it is seen that $l'(\nu) > 0$ as long as u_ϕ is positive. Hence, by keeping the plate switched on during the interval where $l'(\nu) > 0$ and off when $l'(\nu) < 0$, a transfer program can be established whereby the semilatus rectum l is made to grow more rapidly than under any other bang-bang switching program. The switching points ν_1 and ν_2, representing the zeros of u_ϕ can readily be determined. In the case $i_{00} = 0$, u_ϕ equals sin $(\nu - \tilde{\eta})$ and the switching points $\nu_1 = \tilde{\eta}$ and $\nu_2 = \tilde{\eta} + \pi$ lie on the Sun-Earth line; ν_1 between the Sun and Earth and ν_2 on the opposite side of the Earth (Fig. 2). It is convenient to express the results in terms of the angle $\lambda = \tilde{\eta} - \tilde{\omega}$, at which the Sun-Earth line is inclined with respect to the instantaneous major axis. A few of the typical results are

$$\begin{aligned}
\Delta l &= 2\epsilon a^2 l\{3 - (1-e^2)/(1-e^2\cos^2\lambda) + 3eF(e,\lambda)\sin\lambda\} \\
\Delta a &= 4\epsilon a^2 l/(1-e^2\cos^2\lambda) \\
\Delta e &= -3\epsilon a\, l\sin\lambda\{e\sin\lambda/(1-e^2\cos^2\lambda) + F(e,\lambda)\} \\
e\Delta\tilde{\omega} &= 3\epsilon a l\cos\lambda\{e\sin\lambda/(1-e^2\cos^2\lambda) + F(e,\lambda)\}
\end{aligned} \tag{20}$$

where the variables j denote the instantaneous orbital elements at switchon. The function $F(e,\lambda)$ stands for

$$F(e,\lambda) = \frac{\pi/2 + \arctan[e\sin\lambda/(1-e^2)^{1/2}]}{(1-e^2)^{1/2}} \tag{21}$$

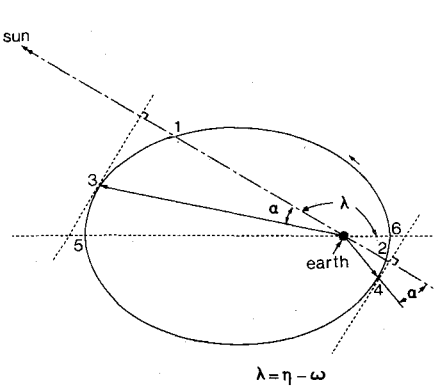

Fig. 2 Configuration of switching points for controlled orbital change.

While these expressions designate the changes in the elements over one revolution, the long-term behavior can be found by repeated rectification of the elements and iteration of the results in Eqs. (20).

B. Increase in Semimajor axis

Another interesting strategy would be the one in which switching takes place at the instants when the velocity vector is normal to the Sun-Earth line. Switchon occurs when the satellite is (roughly) between Sun and Earth (point 3) and switchoff at point 4 on the other side. This policy is effective in the sense that the rate of change of total energy (and major axis) is always positive during the on-phase, since the component of the perturbing force along the instantaneous velocity vector is positive. This approach makes the change in the semimajor axis larger than any other bang-bang switching program. It is clear that if the orbit is circular, this switching program is identical to the previous one. The switching instants are determined as $\nu_3 = \tilde{\eta} + \alpha$ and $\nu_4 = \tilde{\eta} + \pi - \alpha$ with α defined as

$$\alpha = \arcsin(q\cos\tilde{\eta} - p\sin\tilde{\eta}) = \arcsin(e\sin\lambda) \tag{22}$$

If $0 < \lambda < \pi$, the on-phase is less than π radians, whereas for $\pi < \lambda < 2\pi$ it is more. The change in orbital elements at the end of one on/off cycle may be written as

$$\begin{aligned}
\Delta a &= 4\epsilon a^3(1-e^2\sin^2\lambda)^{1/2} \\
\Delta l &= 3\pi\epsilon a^3 e\sin\lambda(1-e^2)^{1/2} \\
&\quad + 4\epsilon a^2 l(1+e^2\sin^2\lambda)/(1-e^2\sin^2\lambda)^{1/2} \\
\Delta e &= -\epsilon a^2\cos\lambda\{3\pi(1-e^2)^{1/2}/2 \\
&\quad + 4e(1-e^2)\sin\lambda/(1-e^2\sin^2\lambda)^{1/2}\} \\
e\Delta\tilde{\omega} &= \epsilon a^2(1-e^2)^{1/2}\cos\lambda\{3\pi/2 \\
&\quad + 4e(1-e^2)^{1/2}\sin\lambda/(1-e^2\sin^2\lambda)^{1/2}\} \\
&\quad + \epsilon a^2 e^3\cos(2\lambda)/(1-e^2\sin^2\lambda)^{1/2}
\end{aligned} \tag{23}$$

It can be shown that Δa is indeed larger and Δl smaller than the corresponding results in Eqs. (20) for $e > 0$ and that the results coincide for $e = 0$.

While the long-term implications of the present switching strategy can be assessed by repeated rectification and iteration of the results of Eqs. (23), additional insight is obtained by application of the two-variable expansion procedure. Although this method is usually employed for equations with continuous right-hand side, in the present case one may interpret the zeroth-order results as those obtained by analytical averaging (where the right-hand side of the first-order differential equations is replaced by their average rate of change, i.e., they are integrated over the on-interval and divided by 2π). In the evaluation of higher-order contributions, many terms in the Fourier expansion would need to be carried, however, for obtaining acceptable accuracy. In the present case, ϵ is of the order 10^{-4} to 10^{-3} so that the zeroth-order approximation is expected to be a uniformly valid (ϵ-band) representation for the actual solution over a (nondimensional) time interval of the order 10^3 to 10^4, corresponding to upper and lower limits of 1600 and 160 days. For an assessment of the effectiveness of the control strategies, accuracies of this order would be sufficient, especially since the errors due to simplifying assumptions (e.g. instantaneous switching) may lead to comparable discrepancies in a practical application.

The zeroth-order equations obtained by the two-variable expansion procedure can be written in the form

$$a_0'(\bar{\nu}) = 2a_0^3(1-x_0^2)^{1/2}/\pi$$

$$x_0'(\bar{\nu}) = y_0 a_0^{3/2}/c_\epsilon - x_0 a_0^2 [1+(1-x_0^2-y_0^2)/(1-x_0^2)^{1/2}]/\pi$$

$$y_0'(\bar{\nu}) = -x_0 a_0^{3/2}/c_\epsilon + y_0 a_0^2 x_0^2/[\pi(1-x_0^2)^{1/2}] \quad (24)$$

These equations were integrated numerically using a double precision Runge-Kutta routine with error control. The solution was found to be in good agreement with the one from the rectification/iteration method. Over approximately four years, the results are consistent up to the first decimal place of the semimajor axis. An interesting estimate for the major axis can be found for near-circular orbits

$$a_0(\bar{\nu}) \leq a_{00}/(1-4\bar{\nu}a_{00}^2/\pi)^{1/2}$$

predicting that an escape trajectory is not reached before $\nu = \pi/(4\epsilon a_{00}^2)$, i.e. approximately 625 revolutions (7 years) for $\epsilon = 0.0002$ and $a_{00} = 1$.

C. Switching at Perigee and Apogee

Finally, a control program is studied where the swtich-on takes place at the instantaneous perigee $\nu_6 = \tilde{\omega}$ and switchoff at the apogee $\nu_5 = \nu_6 + \pi$. The orbit is assumed to remain elliptic throughout. The first-order perturbations were obtained as

$$\Delta I = 3\pi\epsilon e a^3 (1-e^2)^{1/2}\sin\lambda + 4\epsilon a^3 (1-e^2)\cos\lambda$$

$$\Delta a = 4\epsilon a^3 \cos\lambda$$

$$\Delta e = -3\pi\epsilon a^2 (1-e^2)\sin\lambda/2$$

$$e\Delta\tilde{\omega} = 3\pi\epsilon a^2 (1-e^2)^{1/2}\cos\lambda/2 + 2\epsilon e a^2 \sin\lambda \quad (25)$$

The change in Δe is exactly one-half of that when the force acts continuously. Although Eqs. (25) give the results only over one revolution, the long-term perturbations may be obtained by iteration and rectification.

Discussion of Results

Figure 3 shows the secular variations of Ω, the longitude of the ascending node, for four different values of initial solar aspect angle (η_{00}). Of particular interest is the insensitivity of the plot to the initial orbital eccentricity in the range 0-0.5. Long-term variations of the orbital inclination for two different values of initial eccentricity, zero and 0.5, are presented in Fig. 4 and 5, respectively. For the case of a circular orbit, the oscillating perturbations are confined to within 0.5 deg. Note that influence of the initial position of the Sun when located on the line of nodes is the same ($\eta_{00} = 0, \pi$). This is also true for the starting position of the Sun at right angles to the line of nodes ($\eta_{00} = \pi/2, 3\pi/2$), but here the perturbations are of higher frequency and lower amplitude. However, for $e_{00} = 0.5$ (Fig. 5) there is a marked dependence on initial solar aspect angle with substantially larger perturbations of about one year period.

Long-term variations of e (eccentricity) and $\tilde{\omega}$ (argument of the perigee) are presented in the form of polar plots ($p = e\cos\tilde{\omega}$, $q = e\sin\tilde{\omega}$) in Fig. 6. It is emphasized that considerable condensation of information is involved as four different values of $\eta_{00}(0, \pi, \pi/2, 3\pi/2)$ and two values of initial eccentricity ($e_{00} = 0, 0.5$) are represented. The eccentricity is periodic with a period of 363 days, and the evolution of the major axis depends critically upon the initial eccentricity. For e_{00} sufficiently large, the orbit remains elliptic with its axis exhibiting periodic oscillations (amplitude of about 12 deg for $e_{00} = 0.5$) as well as a slow clockwise rotation (about 2 deg per year), as in Fig. 6b. For an initially circular orbit (Fig. 6a) the behavior of $\tilde{\omega}$ is completely different, showing an increase of 180 deg over one year followed

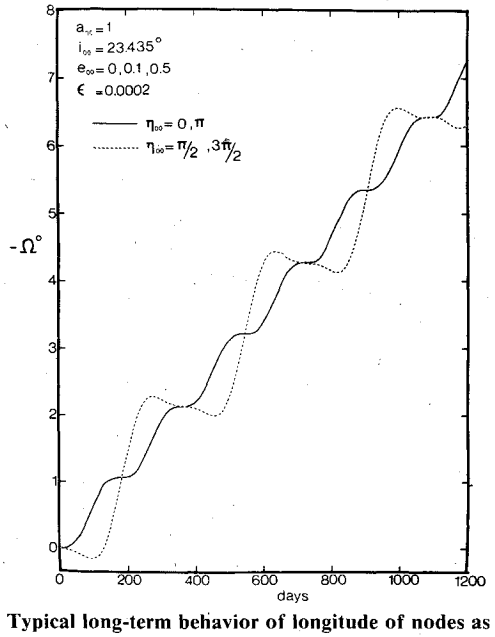

Fig. 3 Typical long-term behavior of longitude of nodes as affected by initial solar aspect angle.

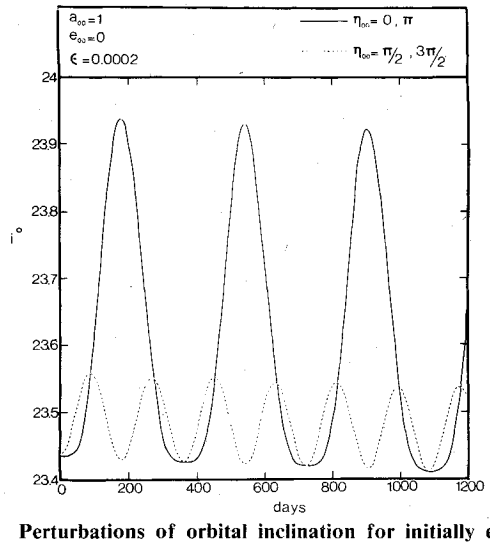

Fig. 4 Perturbations of orbital inclination for initially equatorial circular trajectory.

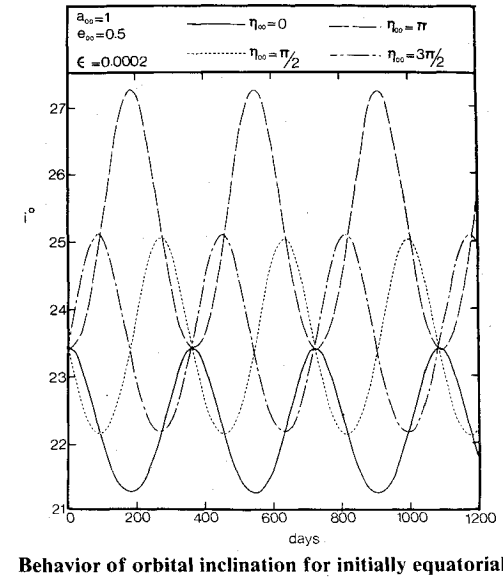

Fig. 5 Behavior of orbital inclination for initially equatorial orbit of eccentricity 0.5.

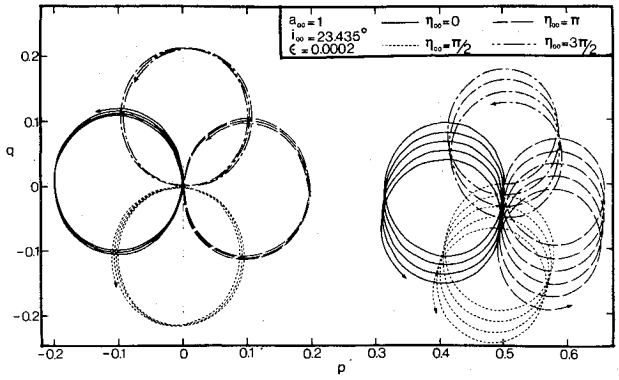

Fig. 6 Polar plots for initial eccentricity of zero and 0.5. Here $p=e\cos\bar{\omega}$ and $q=e\sin\bar{\omega}$.

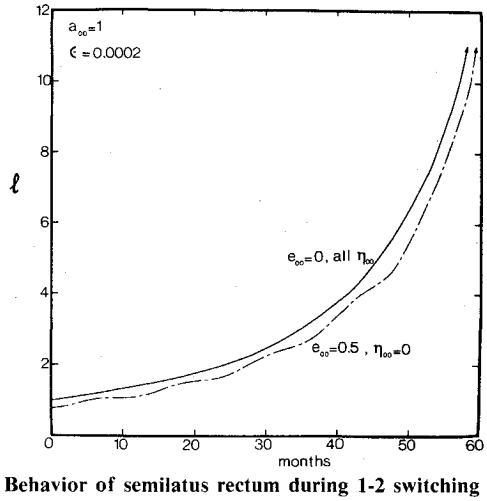

Fig. 7 Behavior of semilatus rectum during 1-2 switching program (case A).

by an instantaneous jump of 180 deg when the eccentricity passes through zero again. Also the slow clockwise rotation is apparent.

The most striking outcome of the switching procedures A and B was a spectacular change in the semilatus rectum (l) and and the semimajor axis (a) (Figs. 7 and 8). Note that starting from the geosynchronous altitude, l and a change by a factor of 10 in about five years! The behavior remained essentially similar for starting eccentricity in the range of 0-0.5 and for all η_{00}.

Figures 8 and 9 present the results of switching at points 3 and 4 as mentioned in case B. Variation of the semimajor axis is quite large for geosynchronous orbits. However, for near-Earth satellites, where gravitational influence predominates, the change in orbit is relatively small as expected. Nevertheless, starting out from $a_{00}=0.34$ (i.e. 8000 km above the Earth), the geosynchronous altitude can be reached in less than five years. This could be of interest for future space stations, such as Solar Satellite Power Station (SSPS), which may be constructed at a lower altitude orbit and may propel itself to a geosynchronous position using this switching program. Note also the rather large variations in eccentricity (Fig. 9), thus permitting a sweep over a wide expanse of space. Advantages rendered by this capability in terms of trajectory transfer, scientific exploration, escape, and launch into heliocentric orbits are apparent.

Concluding Remarks

Solar radiation induced perturbations for a satellite in an arbitrary trajectory are studied using the two-variable expansion procedure and a perturbation/rectification scheme. Controlled orbital transfer through several switching strategies is also explored. Based on the analysis the following general remarks can be made.

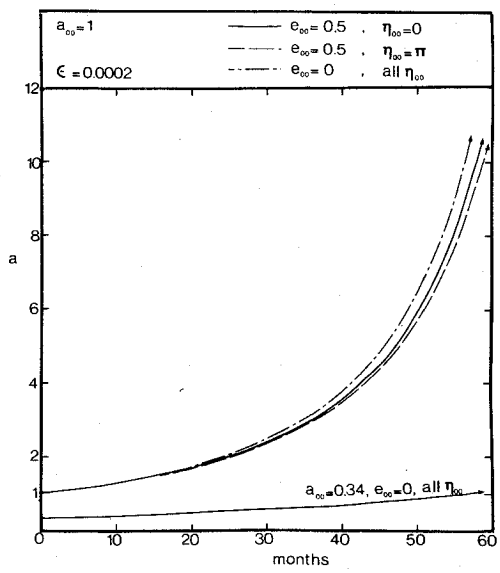

Fig. 8 Controlled variation of the semimajor axis for 3-4 switching program (case B).

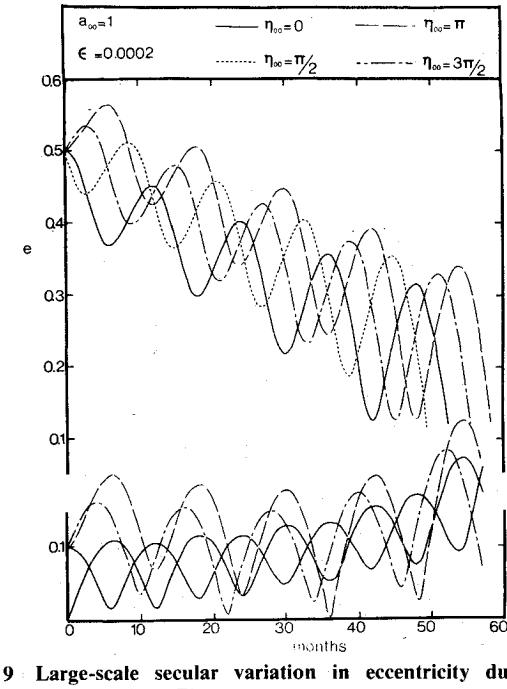

Fig. 9 Large-scale secular variation in eccentricity during 3-4 switching program (case B).

1) Both the two-variable expansion and perturbation/rectification procedures yield long-term valid results of good accuracy. If the resulting equations can be integrated analytically, the former would provide elegant closed-form solutions. On the other hand, however, the latter is more attractive in terms of computational time and effort compared to a numerical integration of the two-variable or original equations.

2) The longitude of the ascending node shows a secular variation which is essentially insensitive to the initial orbital eccentricity in the range 0-0.5.

3) The perturbations in orbital inclination show a marked dependence on both initial solar aspect angles and eccentricity.

4) Polar plots provide a concise and attractive way for visualizing the long-term behavior of the in-plane orbital elements; the variation in eccentricity is periodic while the argument of the perigee shows a slow secular trend. The semimajor axis remains essentially constant in the long run.

5) A suitable choice of switching strategies can lead to a spectacular change in the major axis. This can be used to advantage in the planning of scientific space exploration, trajectory transfer, launching of space stations, and many other applications.

References

[1]Musen, P., "The Influence of Solar Radiation Pressure on the Motion of an Artificial Satellite," *Journal of Geophysical Research,* Vol. 65, May 1960, pp. 1391-1396.

[2]Shapiro, I.I., "The Prediction of Satellite Orbits," *Dynamics of Satellites,* edited by M. Roy, Springer Verlag, Berlin, 1963, pp. 257-312.

[3]Musen, P., Bryant, R., and Baillie, A., "Perturbations in Perigee Height of Vanguard I," *Science,* Vol. 131, March 1960, pp. 935-936.

[4]Bryant, R., "A Comparison of Theory and Observation of the Echo I Satellite," *Journal of Geophysical Research,* Vol. 66, Sept. 1961, pp. 3066-3069.

[5]Garwin, R.L., "Solar Sailing—A Practical Method of Propulsion Within the Solar System," *Jet Propulsion,* Vol. 28, March 1958, pp. 188-190.

[6]Tsu, T.C., "Interplanetary Travel by Solar Sail," *Journal of the American Rocket Society,* Vol. 29, June 1959, pp. 422-427.

[7]London, H.S., "Some Exact Solutions of the Equations of Motion of a Solar Sail with Constant Sail Setting," *Journal of the American Rocket Society,* Vol. 30, Feb. 1960, pp. 198-200.

[8]Kiefer, J. W., "Feasibility Considerations for a Sail Powered Multi-Mission Solar Probe," *Proceedings of the XVth International Astronautical Congress,* Vol. 1, edited by Michael Lunc, Gauthier-Villars, Paris, 1965, pp. 383-416.

[9]Modi, V.J., Pande, K.C., and Nicks, G.W., "On the Semi-Passive Propulsion of Space Vehicles Using Solar Radiation Pressure," *Proceedings of the 10th International Symposium on Space Technology and Science,* edited by S. Kobayashi, AGNE Publishing Co., Tokyo, pp. 375-382.

[10]Sands, N., "Escape from Planetary Gravitational Fields by Use of Solar Sails," *Journal of the American Rocket Society,* Vol. 31, April 1961, pp. 527-531.

[11]Fimple, W.R., "Generalized Three-Dimensional Trajectory Analysis of Planetary Escape by Solar Sail," *Journal of the American Rocket Society,* Vol. 32, June 1962, pp. 883-887.

[12]Cohen, M.J. and Freestone, M.M., "Spiral Trajectories in the Plane of the Ecliptic Using Solar Sails," *Proceedings of the XVth International Astronautical Congress,* Vol. 1, edited by Michael Lunc, Gauthier-Villars, Paris, 1965, pp. 363-382.

[13]Kevorkian, J., "The Two-Variable Expansion Procedure for the Approximate Solution of Certain Nonlinear Differential Equations," *Lectures in Applied Mathematics,* edited by J.B. Rosser, Vol. 7, Part 3, 1966, pp. 206-275.

[14]Nayfeh, A., *Perturbation Methods,* John Wiley and Sons, New York, 1973, Chap. 6.

[15]Shi, Y.Y. and Eckstein, M.C., "Ascent or Descent from Satellite Orbit by Low Thrust," *AIAA Journal,* Vol. 4, Dec. 1966, pp. 2203-2208.

[16]Moss, J.B., "Perturbation Techniques and Orbit Expansion by Continuous Low Thrust," *Journal of the British Interplanetary Society,* Vol. 27, March 1974, pp. 213-225.

[17]Flandro, G.A., "Analytic Solution for Low Thrust Out of Ecliptic Mission Design," *AIAA Journal,* Vol. 14, Jan. 1976, pp. 33-38.

[18]Van der Ha, J.C. and Modi, V.J., "Long Term Orbital Behavior of a Satellite in the Ecliptic Plane under the Influence of Solar Radiation," *XXVIIth International Astronautical Congress,* International Astronautical Federation, Anaheim, Calif., Paper No. 76-006, Oct. 1976.

[19]Van der Ha, J.C., "Solar Radiation Induced Perturbations and Control of Satellite Trajectories," Ph.D. Dissertation, Department of Mechanical Engineering, University of British Columbia, June 1977.