On the secular variations of the elements of satellite orbits

By G. N. Ward

The College of Aeronautics, Cranfield

(*Communicated by M. J. Lighthill, F.R.S.—Received* 20 *September* 1961)

It is shown that a simple and rigorous theory of the perturbations of the orbits of close earth satellites can be constructed which is valid for all values of orbital eccentricity less than unity. A vector treatment is used which leads naturally to a description of the orbital geometry in terms of the angular momentum, the direction of perigee, and the eccentricity, from which other elements may be obtained without difficulty. The method is applied briefly to the first-order effects of the earth's oblateness, and in more detail to the effects of atmospheric resistance. Certain integrals occurring in the theory of atmospheric resistance are evaluated as asymptotic series whose first two terms are sufficient for determining the perturbations to high accuracy.

1. Introduction

The motion of a close satellite of the earth departs from that of a Keplerian elliptical orbit primarily as the result of perturbing forces caused by the oblateness of the earth and the presence of the earth's atmosphere. There are also secondary perturbing forces caused by the gravitational attractions of the sun and the moon, and by solar radiation pressure, but these are small compared with the primary perturbations, except possibly in the case of low-density balloon satellites for which radiation pressures may be important, provided that the satellite does not move too far away from the earth. These perturbing forces are sufficiently small compared with the central gravitational field to allow the use of series-expansion or iteration methods for the calculation of the perturbed motions, and, further, for most satellites launched to date, the orbital eccentricities have been small enough to allow the theoretical predictions of the motions to be based on series expansions in powers of the eccentricity. The resulting analysis has usually been somewhat complicated, and it is the purpose of the present paper to show that a simplified and rigorous theory of these motions can be developed which does not depend upon the eccentricity being small. The method depends on the possibility of choosing a mean Keplerian orbit whose elements vary slowly with time, the actual orbital motion being oscillatory about this mean orbit, with period approximately equal to the orbital period. The applications in this paper are limited to the determination of the first-order variations of the mean orbit that are caused by the earth's oblateness and atmospheric resistance, but the actual motion is not difficult to obtain. The extension to second-order perturbations is easy in principle, but the algebra is much heavier than that for the first-order solution.

For conciseness, the analysis is in vector form, and this leads naturally to a somewhat unconventional description of the orbital geometry in terms of the angular momentum, the direction of perigee, and the eccentricity as elements, from which the more usual sets of elements can be obtained without difficulty. The expressions found for the rates of change of the angular momentum and of the direction of perigee can be interpreted directly in terms of rotations of these vectors. The advantages of a vector treatment of the perturbation problem, using vector elements which are equivalent to those used here, have been noticed by previous writers. The most comprehensive treatment seems to be that of Musen (1954); more recently, vector equations of the form used below were published by Allan (1961).

2. Equations of motion

If $\mathbf{r}$ denotes the position vector of the centre of mass of the satellite relative to the earth's centre, $r = |\mathbf{r}|$, and $\hat{\mathbf{r}} = \mathbf{r}/r$, then the equation of motion can be written in the form

$$\ddot{\mathbf{r}} = -(\mu/r^2)\,\hat{\mathbf{r}} + \mathbf{F}, \tag{1}$$

where μ is the gravitational constant for the earth, and $\mathbf{F}$ is the sum of the perturbing forces per unit satellite mass. On multiplying this equation vectorially by $\mathbf{r}$, and denoting the angular momentum per unit mass, $\mathbf{r} \wedge \dot{\mathbf{r}}$, by $\mathbf{H}$, we have

$$\dot{\mathbf{H}} = \mathbf{r} \wedge \mathbf{F}, \tag{2}$$

and on multiplying it vectorially by $\mathbf{H}$ and using (2), we have

$$\mathrm{d}(\mathbf{H} \wedge \dot{\mathbf{r}} + \mu\hat{\mathbf{r}})/\mathrm{d}t = (\mathbf{r} \wedge \mathbf{F}) \wedge \dot{\mathbf{r}} + \mathbf{H} \wedge \mathbf{F}. \tag{3}$$

When $\mathbf{F} = 0$, (1) is the equation for Keplerian orbits, and if a suffix zero is used to denote quantities which refer to Keplerian orbits (this convention is used throughout the paper), then (2) and (3) yield the well-known results

$$\mathbf{r}_0 \wedge \dot{\mathbf{r}}_0 = \mathbf{H}_0 = \text{constant}, \tag{4}$$

and

$$\mathbf{H}_0 \wedge \dot{\mathbf{r}}_0 + \mu\hat{\mathbf{r}}_0 = -\mu e_0 \mathbf{a}_0 = \text{constant}, \tag{5}$$

where e_0 is the eccentricity of the orbit and $\mathbf{a}_0$ is the unit vector in the direction of perigee. Scalar multiplication of (5) by $\mathbf{r}_0$ now gives

$$H_0^2 = \mu r_0(1 + e_0 \cos\phi), \tag{6}$$

where ϕ is the true anomaly, defined by the relation $\cos\phi = \mathbf{a}_0 . \hat{\mathbf{r}}_0$, and vector multiplication of (5) by $\mathbf{H}_0$ gives

$$H_0^2 \dot{\mathbf{r}}_0 = \mu \mathbf{H}_0 \wedge (\hat{\mathbf{r}}_0 + e_0 \mathbf{a}_0). \tag{7}$$

The length, a_0, of the semi-major axis is then given by

$$H_0^2 = \mu a_0(1 - e_0^2), \tag{8}$$

and the periodic time, T_0, is given by

$$H_0 T_0 = 2\pi a_0^2 (1 - e_0^2)^{\frac{1}{2}}. \tag{9}$$

The geometry of a Keplerian orbit is defined completely by the values of $\mathbf{H}_0$, $\mathbf{a}_0$, and e_0 (with $\mathbf{a}_0^2 = 1$ and $\mathbf{a}_0 . \mathbf{H}_0 = 0$) giving five independent scalar elements, which may be used as elements of the orbit. The sixth element, defining the position of the satellite in its orbit at any given time, is not required in the present application.

If $\mathbf{F}$ is a small perturbing force in the sense that $|r^2\mathbf{F}| \ll \mu$, then the satellite's motion can be described in the usual manner in terms of an osculating Keplerian orbit whose elements are varying with time, the equations for the rates of change of these elements being (2), and (3) with $-\mu e_0 \mathbf{a}_0$ substituted for $\mathbf{H} \wedge \dot{\mathbf{r}} + \mu\hat{\mathbf{r}}$. These equations may be solved by iteration, and it is known that the variations of the elements with time may be divided into two classes: first, there are periodic variations which have approximately the same period as the unperturbed Keplerian orbit, which are called the short-period variations; secondly, there are much slower variations whose time-scale is long compared with the orbital period, which are called the long-period or secular variations. This latter class contains both periodic and aperiodic variations in general, but it is not necessary to distinguish between them, and for convenience both are called secular variations here.

If the direct iteration procedure is adopted, then as the eccentricity, e_0, becomes small, the short-period variations of the direction of perigee, which is measured by $\mathbf{a}_0$, dominate the secular variations in general, and the iteration is non-uniformly valid as $e_0 \to 0$. For this reason it is desirable to eliminate the short-period variations from the elements, and this can be achieved by introducing small extra terms to absorb the short-period variations and leave $\mathbf{H}_0$ and $e_0\mathbf{a}_0$ with secular variations only. The elements $\mathbf{H}_0$, $\mathbf{a}_0$, and e_0 then no longer refer to an osculating Keplerian orbit, but rather to a mean Keplerian orbit about which the satellite oscillates with a short-period motion. The resulting process is uniformly valid as $e_0 \to 0$.

This difficulty at $e_0 = 0$ has been overcome in a similar manner by a number of authors, in particular by King-Hele (1958), Brenner & Latta (1960), and Struble (1960). These authors defined non-elliptic motions in moving orbital planes by adding extra terms, but their orbital planes always contain the satellite and consequently are subject to short-period variations in their motions. The method outlined in the previous paragraph defines a mean orbital plane whose motion is always of a purely secular nature, and the satellite is not confined to this plane, but has a short-period motion out of it.

3. The first-order secular equations

It is convenient at this stage to introduce some additional unit vectors. We take the unit vector $\mathbf{n}_0$ in the direction of $\mathbf{H}_0$; we choose $\mathbf{a}_0$ to be perpendicular to $\mathbf{n}_0$, and write $\mathbf{b}_0 = \mathbf{n}_0 \wedge \mathbf{a}_0$, so that $\mathbf{a}_0$, $\mathbf{b}_0$, and $\mathbf{n}_0$ are an orthogonal triad of unit vectors which define a rotating frame of reference in which the short-period variations are described, the plane containing $\mathbf{a}_0$ and $\mathbf{b}_0$ being the mean orbital plane. The unit radial vector, $\hat{\mathbf{r}}_0$, can now be expressed in the form

$$\hat{\mathbf{r}}_0 = \mathbf{a}_0 \cos\phi + \mathbf{b}_0 \sin\phi. \tag{10}$$

To carry out the procedure outlined above, we introduce a small parameter, ϵ, which characterizes the magnitude of the perturbing force, and express all quantities as series in ascending powers of ϵ. We take

$$\mathbf{F} = \epsilon\mathbf{F}_1 + \epsilon^2\mathbf{F}_2 + \dots, \tag{11}$$

and try to find solutions of the equations of motion in the form

$$\mathbf{H} = H_0(\mathbf{n}_0 + \epsilon\mathbf{h}_1 + \epsilon^2\mathbf{h}_2 + \dots), \tag{12}$$

$$\mathbf{H} \wedge \dot{\mathbf{r}} + \mu\hat{\mathbf{r}} = -\mu(e_0\mathbf{a}_0 + \epsilon\mathbf{f}_1 + \epsilon^2\mathbf{f}_2 + \dots), \tag{13}$$

$$\hat{\mathbf{r}} = \hat{\mathbf{r}}_0 + \epsilon r_1\mathbf{n}_0 + \epsilon^2\mathbf{r}_2 + \dots, \tag{14}$$

such that $\mathbf{H}_0$, $\mathbf{a}_0$, and e_0 contain no short-period terms. There is no loss of generality involved in taking the special form of the first-order term, $\epsilon r_1 \mathbf{n}_0$, in (14), since the first-order perturbation of $\hat{\mathbf{r}}$ can always be made perpendicular to the mean orbital plane by choosing $\mathbf{a}_0$ suitably. The coefficients in the three series expansions (12), (13), and (14) are not independent, because the identity $\mathbf{H}.\hat{\mathbf{r}} = 0$ must be satisfied, and another identity is obtained by multiplying (13) scalarly by $\mathbf{H}$ and substituting the series for $\mathbf{H}$ and $\hat{\mathbf{r}}$. The first-order relations obtained in this way are

$$r_1 = -\hat{\mathbf{r}}_0 . \mathbf{h}_1, \tag{15}$$

which determines r_1, and

$$\mathbf{n}_0 . \mathbf{f}_1 + e_0 \mathbf{a}_0 . \mathbf{h}_1 = 0. \tag{16}$$

The second-order terms in these series are not used in the following analysis, and have been displayed only to show the form of the series. Moreover, we shall see that explicit expressions for the first-order terms are not required when only the secular variations are to be calculated, and that it is only necessary to know that they exist and satisfy (16). They can be found easily, however, if they are needed for determining the short-period perturbations or for use in second-order theory.

On substituting the series expansions in (2) and (3), and retaining only first-order terms on the right-hand sides, we obtain the equations

$$\dot{\mathbf{H}}_0 = \epsilon \mathbf{r}_0 \wedge \mathbf{F}_1 - \epsilon H_0 \dot{\mathbf{h}}_1 + O(\epsilon^2) \tag{17}$$

and

$$\mathrm{d}(e_0 \mathbf{a}_0)/\mathrm{d}t = -\epsilon(\mathbf{r}_0 \wedge \mathbf{F}_1) \wedge \dot{\mathbf{r}}_0 - \epsilon \mathbf{H}_0 \wedge \mathbf{F}_1 - \epsilon\mu \dot{\mathbf{f}}_1 + O(\epsilon^2). \tag{18}$$

The right-hand sides of these equations are nearly-periodic functions of time, and can be expressed as Fourier series whose fundamental period is the orbital period, T, defined as the time between successive passages of the ascending node, the period, T, and the coefficients being slowly varying functions of time. The short-period terms in these Fourier series can be cancelled by proper choice of $\dot{\mathbf{h}}_1$ and $\dot{\mathbf{f}}_1$, and then the right-hand sides contain only secular terms, as required. There is an indeterminacy in $\mathbf{h}_1$ and $\mathbf{f}_1$ to the extent of a secular term in each function, which may be removed by imposing the condition that $\mathbf{h}_1$ and $\mathbf{f}_1$ have zero mean value over an orbital period, and with this condition, the identity (16) is satisfied automatically, for the identity

$$\mathrm{d}(e_0 \mathbf{a}_0 . \mathbf{H}_0)/\mathrm{d}t = 0 \tag{19}$$

combined with (17) and (18) shows that

$$\mathbf{n}_0 . \dot{\mathbf{f}}_1 + e_0 \mathbf{a}_0 . \dot{\mathbf{h}}_1 = 0, \tag{20}$$

and on integration with the condition of zero mean value, this yields (16).

The first-order approximation is maintained if the orbital period, T, is replaced by the period, T_0, of the Keplerian orbit, and then we have the general first-order secular equations in the form

$$\frac{\mathrm{d}\mathbf{H}_0}{\mathrm{d}t} = \frac{\epsilon}{T_0} \int_0^{T_0} \mathbf{r}_0 \wedge \mathbf{F}_1 \mathrm{d}t \tag{21}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}(e_0 \mathbf{a}_0) = -\frac{\epsilon}{\mu T_0} \int_0^{T_0} \{(\mathbf{r}_0 \wedge \mathbf{F}_1) \wedge \dot{\mathbf{r}}_0 + \mathbf{H}_0 \wedge \mathbf{F}_1\} \mathrm{d}t. \tag{22}$$

In applications to particular problems, it is usually convenient to replace the time as the variable of integration by the true anomaly, ϕ, or the eccentric anomaly,

E, of the Keplerian orbit, and for this purpose the values of $\mathrm{d}t/\mathrm{d}\phi$ and $\mathrm{d}t/\mathrm{d}E$ are required. For the present first-order theory, the Keplerian values of these derivatives suffice, and these are

$$\frac{\mathrm{d}t}{\mathrm{d}\phi} = \frac{r_0^2}{H_0} \quad \text{and} \quad \frac{\mathrm{d}t}{\mathrm{d}E} = \frac{a_0(1-e_0^2)^{\frac{1}{2}}\, r_0}{H_0}. \tag{23}$$

But for the second-order theory, better approximations are required, and these seem to be obtainable most easily from the values of $(\mathrm{d}\hat{\mathbf{r}}/\mathrm{d}t)^2$, $(\mathrm{d}\hat{\mathbf{r}}/\mathrm{d}\phi)^2$, and $(\mathrm{d}\hat{\mathbf{r}}/\mathrm{d}E)^2$. For example, in the case of $\mathrm{d}t/\mathrm{d}\phi$, we have

$$\left(\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}t}\right)^2 = \left(\frac{\mathbf{H}\wedge\hat{\mathbf{r}}}{r^2}\right)^2 = \frac{\mathbf{H}^2}{r^4}, \tag{24}$$

and from (14), on taking account of the orthogonality of $\mathbf{a}_0$, $\mathbf{b}_0$, and $\mathbf{n}_0$, it is not difficult to show that

$$\left(\frac{\mathrm{d}\hat{\mathbf{r}}}{\mathrm{d}t}\right)^2 = 1 + 2\mathbf{b}_0 \cdot \frac{\mathrm{d}\mathbf{a}_0}{\mathrm{d}t} + O(\epsilon^2), \tag{25}$$

from which it follows that

$$\frac{\mathrm{d}t}{\mathrm{d}\phi} = \frac{r^2}{H}\left(1 + \mathbf{b}_0 \cdot \frac{\mathrm{d}\mathbf{a}_0}{\mathrm{d}t}\right) + O(\epsilon^2). \tag{26}$$

It seems appropriate to remark about the uniform validity of the expansions as $e_0 \to 0$ at this stage, because the difficulties which can arise if the process is not uniformly valid will not be apparent in what follows. It will be seen that (18) gives the rate of change of $e_0\mathbf{a}_0$, so the calculation of the rate of change of $\mathbf{a}_0$ necessitates dividing the right-hand side of this equation by e_0. Now in general there are terms in $(\mathbf{r}_0 \wedge \mathbf{F}_1)\wedge\dot{\mathbf{r}}_0 + \mathbf{H}_0 \wedge \mathbf{F}_1$ which are independent of e_0, which would lead to terms with factors $1/e_0$ in the equation for $\mathrm{d}\mathbf{a}_0/\mathrm{d}t$ if they were not removed by $\mathbf{f}_1$, and would give a non-uniformity in the theory as $e_0 \to 0$. Similarly, in second-order theory, such terms are removed by $\mathbf{f}_2$, and it may be supposed that they can also be removed in higher-order approximations, thus giving uniformly valid expansions as $e_0 \to 0$. Whether the resulting series converge or are of an asymptotic nature has not been determined.

Before proceeding to apply (21) and (22) to special problems, we must notice that although these equations are linear in the perturbing force, they are non-linear in the orbital elements, so perturbations in these elements which arise from separate causes cannot be superposed by addition in general, and it is only the rates of change of the elements that are additive.

4. The effect of the earth's oblateness

This effect has been investigated previously by many authors, so the treatment here is brief.

If $\mathbf{k}$ is the unit vector parallel to the earth's polar axis, directed from south to north, then the gravitational potential of the earth can be represented as a series of spherical harmonics in the form

$$\frac{\mu}{r}\left\{1 - \sum_{n=2}^{\infty} J_n\left(\frac{R}{r}\right)^n P_n(\mathbf{k}.\hat{\mathbf{r}})\right\}, \tag{27}$$

where R is the equatorial radius of the earth, and the J_n are constants. J_2 is about 10^{-3} and the others are roughly 1000 times smaller (or less), so the parameter ϵ of § 3 may be replaced by J_2. The force per unit mass acting on the satellite, from the potential (27), is

$$-\frac{\mu}{r^2}\hat{\mathbf{r}} - \frac{3J_2\mu R^2}{2r^4}\{[1-5(\mathbf{k}.\hat{\mathbf{r}})^2]\,\hat{\mathbf{r}} + 2\mathbf{k}.\hat{\mathbf{r}}\mathbf{k}\} + O(J_2^2), \tag{28}$$

so we have

$$\epsilon\mathbf{F}_1 = -\frac{3J_2\mu R^2}{2r_0^4}\{[1-5(\mathbf{k}.\hat{\mathbf{r}}_0)^2]\,\hat{\mathbf{r}}_0 + 2\mathbf{k}.\hat{\mathbf{r}}_0\mathbf{k}\}. \tag{29}$$

Because $\mathbf{F}_1$ contains an inverse power of r_0, the true anomaly, ϕ, is the most convenient variable of integration in (21) and (22), which take the form

$$\frac{\mathrm{d}\mathbf{H}_0}{\mathrm{d}t} = -\frac{3J_2\mu R^2}{H_0 T_0}\int_0^{2\pi}\frac{1}{r_0}\mathbf{k}.\hat{\mathbf{r}}_0\hat{\mathbf{r}}_0\wedge\mathbf{k}\,\mathrm{d}\phi \tag{30}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}t}(e_0\mathbf{a}_0) = -\frac{3J_2R^2}{2H_0T_0}\int_0^{2\pi}\frac{1}{r_0^2}\{2r_0\mathbf{k}.\hat{\mathbf{r}}_0(\hat{\mathbf{r}}_0\wedge\mathbf{k})\wedge\dot{\mathbf{r}}_0$$
$$+[1-5(\mathbf{k}.\hat{\mathbf{r}}_0)^2]\,\mathbf{H}_0\wedge\hat{\mathbf{r}}_0 + 2\mathbf{k}.\hat{\mathbf{r}}_0\mathbf{H}_0\wedge\mathbf{k}\}\,\mathrm{d}\phi. \tag{31}$$

After substituting the values of r_0 and $\hat{\mathbf{r}}_0$ from (6) and (7), the evaluation of these integrals is a simple matter, and the final results are

$$\mathrm{d}\mathbf{H}_0/\mathrm{d}t = -2K\cos i_0\,\mathbf{k}\wedge\mathbf{H}_0 \tag{32}$$

and

$$\mathrm{d}(e_0\mathbf{a}_0)/\mathrm{d}t = -K\{(1-5\cos^2 i_0)\,\mathbf{n}_0 + 2\cos i_0\,\mathbf{k}\}\wedge e_0\mathbf{a}_0, \tag{33}$$

where $\cos i_0 = \mathbf{k}.\mathbf{n}_0$, so that i_0 is the inclination of the mean orbital plane to the earth's equatorial plane, and

$$K = \frac{3\pi J_2 R^2}{2a_0^2T_0(1-e_0^2)^2}. \tag{34}$$

It follows from (32) that $\mathbf{H}_0$ is a vector of constant magnitude which rotates about the earth's polar axis with constant angular velocity $2K\cos i_0$ in the sense opposite to that of the satellite's motion, the inclination i_0 remaining constant. The rate of change of $e_0\mathbf{a}_0$ due to this rotation is $-2K\cos i_0\,\mathbf{k}\wedge(e_0\mathbf{a}_0)$, so the rate of change of $e_0\mathbf{a}_0$ in the mean orbital plane is

$$-K(1-5\cos^2 i_0)\,\mathbf{n}_0\wedge(e_0\mathbf{a}_0), \tag{35}$$

which shows that $e_0\mathbf{a}_0$ is a vector of constant magnitude and rotates in the mean orbital plane with angular velocity $K(5\cos^2 i_0 - 1)$ in the same sense as that of the satellite's motion.

5. Atmospheric drag

A full discussion of the form of the drag force acting on a satellite due to atmospheric resistance has been given by Cook, King-Hele & Walker (1960), which need not be repeated here. For the purpose of calculating the principal effects of atmospheric resistance, we assume that the drag force is proportional to the local atmospheric density and the square of the satellite's speed relative to the atmosphere, that this force acts in the direction opposite to that of the satellite's motion relative to the atmosphere, and that the atmospheric density varies exponentially with distance from the earth's centre and does not vary with time. Thus any effects caused by the rotation of the satellite and the oblateness of the earth's atmosphere are neglected. Also the atmospheric density is known to vary exponentially only over rather small ranges of altitude above the earth's surface (King-Hele 1959), and the variations over the ranges under consideration here are far from exponential; however, the approximation is adopted because it is analytically convenient, and, provided that the values of the parameters are chosen properly, gives fairly accurate answers over a large proportion of the satellite's life-time.

With these assumptions, if $\omega\mathbf{k}$ is the angular velocity of the atmosphere, the drag force per unit mass acting on the satellite, being the appropriate form of $\epsilon\mathbf{F}_1$ in this case, is

$$\epsilon\mathbf{F}_1 = -\delta\rho v(\dot{\mathbf{r}}_0 - \omega\mathbf{k}\wedge\mathbf{r}_0), \tag{36}$$

where

$$v = |\dot{\mathbf{r}}_0 - \omega\mathbf{k}\wedge\mathbf{r}_0|, \tag{37}$$

ρ is the atmospheric density, given by

$$\rho = \rho_p \exp\{-\beta(r_0 - r_p)\}, \tag{38}$$

the suffix p referring to conditions at perigee, and δ is a constant for a given satellite. The constant δ can be expressed in the form $\delta = \frac{1}{2}SC_D/m$, where S is the effective cross-sectional area of the satellite, C_D is its drag coefficient based on S, and m is its mass.

From (37) we have

$$\begin{aligned} v^2 &= \dot{\mathbf{r}}_0^2 - 2\omega\mathbf{k}\wedge\mathbf{r}_0.\dot{\mathbf{r}}_0 + \omega^2(\mathbf{k}\wedge\mathbf{r}_0)^2 \\ &= v_0^2\left\{\left(1-\frac{\omega\mathbf{k}.\mathbf{H}_0}{v_0^2}\right)^2 + \frac{\omega^2}{v_0^4}[\dot{\mathbf{r}}_0\wedge(\mathbf{k}\wedge\mathbf{r}_0)]^2\right\}, \end{aligned} \tag{39}$$

where $v_0 = |\dot{\mathbf{r}}_0|$. At perigee, we have

$$\left[\frac{\omega^2}{v_0^4}\{\dot{\mathbf{r}}_0\wedge(\mathbf{k}\wedge\mathbf{r}_0)\}^2\right]_p = \left(\frac{1-e_0}{1+e_0}\right)\left(\frac{\omega}{2\pi}\right)^2 T_p^2(\mathbf{k}.\mathbf{b}_0)^2, \tag{40}$$

where T_p is the period of a circular orbit whose radius is equal to the perigee distance, $a_0(1-e_0)$, of the actual orbit, and for perigee altitudes less than 1000 km, say, and taking ω to be the angular velocity of the earth, we have

$$(\omega/2\pi)^2 T_p^2 < 0{\cdot}003, \tag{41}$$

showing that the last term in (39) is very small at perigee. Away from perigee this term increases, but there are two compensating factors; first, the angular velocity of the atmosphere almost certainly decreases with increasing altitude, which will tend to keep the term small, and secondly, the density decreases rapidly with altitude, so the integrated effect of this term is of less and less importance as the altitude increases. Thus we may conclude that this term always has a very small effect and that we can make the approximation

$$v = v_0(1-\omega\mathbf{k}.\mathbf{H}_0/v_0^2) \tag{42}$$

with an error of about 1 part in 1000 in most cases, and then (36) becomes

$$\epsilon \mathbf{F}_1 = -\delta\rho v_0(1-\omega \mathbf{k}.\mathbf{H}_0/v_0^2)(\dot{\mathbf{r}}_0 - \omega \mathbf{k}\wedge \mathbf{r}_0). \tag{43}$$

The eccentric anomaly, E, proves to be the most convenient variable of integration in (21) and (22) in this case. From the theory of Keplerian orbits we have

$$v_0(\mathrm{d}t/\mathrm{d}E) = a_0(1-e_0^2\cos^2 E)^{\frac{1}{2}}, \tag{44}$$

and then, after a little reduction, the secular equations for a spherical earth take the form

$$\frac{\mathrm{d}\mathbf{H}_0}{\mathrm{d}t} = -\frac{\delta a_0}{T_0}\int_0^{2\pi}(1-e_0^2\cos^2 E)^{\frac{1}{2}}(1-\omega\mathbf{k}.\mathbf{H}_0/v_0^2)\{\mathbf{H}_0-\omega\mathbf{r}_0\wedge(\mathbf{k}\wedge\mathbf{r}_0)\}\rho\,\mathrm{d}E \tag{45}$$

and
$$\frac{\mathrm{d}}{\mathrm{d}t}(e_0\mathbf{a}_0) = \frac{\delta a_0}{\mu T_0}\int_0^{2\pi}(1-e\cos_0^2 E)^{\frac{1}{2}}(1-\omega\mathbf{k}.\mathbf{H}_0/v_0^2)$$
$$\times(2\mathbf{H}_0\wedge\dot{\mathbf{r}}_0-\omega r_0^2\mathbf{k}\wedge\dot{\mathbf{r}}_0+\omega\mathbf{k}.\mathbf{r}_0\mathbf{H}_0+\omega\mathbf{k}.\mathbf{H}_0\mathbf{r}_0)\rho\,\mathrm{d}E. \tag{46}$$

In order to evaluate these integrals, we need the following results for Keplerian orbits:

$$r_0 = a_0(1-e_0\cos E), \tag{47}$$

$$\mathbf{r}_0 = a_0\{(\cos E - e_0)\mathbf{a}_0 + (1-e_0^2)^{\frac{1}{2}}\sin E\mathbf{b}_0\}, \tag{48}$$

$$\dot{\mathbf{r}}_0 = (H_0/r_0)\{\cos E\mathbf{b}_0 - (1-e_0^2)^{-\frac{1}{2}}\sin E\mathbf{a}_0\}, \tag{49}$$

$$v_0^2 = \frac{H_0^2}{a_0^2(1-e_0^2)}\left(\frac{1+e_0\cos E}{1-e_0\cos E}\right). \tag{50}$$

Then, if we omit those terms in the integrands which vanish identically on integration because they are odd functions of E, and use (38) for the density, (45) and (46) become

$$\frac{\mathrm{d}\mathbf{H}_0}{\mathrm{d}t} = -\frac{\delta\rho_p a_0 H_0}{T_0}\exp\{\beta(r_p-a_0)\}\int_0^{2\pi}(1-e_0^2\cos^2 E)^{\frac{1}{2}}\exp(\beta a_0 e_0\cos E)$$
$$\times\left\{1-\frac{\omega a_0^2(1-e_0^2)\cos i_0}{H_0}\left(\frac{1-e_0\cos E}{1+e_0\cos E}\right)\right\}$$
$$\times\{\mathbf{n}_0-(\omega a_0^2/H_0)[(\cos E-e_0)^2\mathbf{a}_0\wedge(\mathbf{k}\wedge\mathbf{a}_0)+(1-e_0^2)\sin^2 E\mathbf{b}_0\wedge(\mathbf{k}\wedge\mathbf{b}_0)]\}\,\mathrm{d}E \tag{51}$$

and
$$\frac{\mathrm{d}}{\mathrm{d}t}(e_0\mathbf{a}_0) = -\frac{2\delta\rho_p a_0(1-e_0^2)}{T_0}\exp\{\beta(r_p-a_0)\}\int_0^{2\pi}(1-e_0^2\cos^2 E)^{\frac{1}{2}}\exp(\beta a_0 e_0\cos E)$$
$$\times\left\{1-\frac{\omega a_0^2(1-e_0^2)\cos i_0}{H_0}\left(\frac{1-e_0\cos E}{1+e_0\cos E}\right)\right\}$$
$$\times\left\{\frac{\cos E}{1-e_0\cos E}\mathbf{a}_0-\frac{\omega a_0^2}{H_0}[\cos i_0(\cos E-e_0)\mathbf{a}_0-\tfrac{1}{2}e_0\sin^2 E\mathbf{k}\wedge\mathbf{b}_0]\right\}\mathrm{d}E. \tag{52}$$

The rates of change of H_0 and e_0 can be obtained from these equations by multiplying scalarly by $\mathbf{n}_0$ and $\mathbf{a}_0$, respectively: the terms involving ω^2 can be shown to be negligible as before, and, on omitting them, we obtain the equations

$$\frac{\mathrm{d}H_0}{\mathrm{d}t} = -\frac{\delta\rho_p a_0 H_0}{T_0}\exp\{\beta(r_p-a_0)\}\int_0^{2\pi}(1-e_0^2\cos^2 E)^{\frac{1}{2}}\exp(\beta a_0 e_0\cos E)$$
$$\times\left\{1-\frac{2\omega a_0^2(1-e_0^2)\cos i_0}{H_0}\left(\frac{1-e_0\cos E}{1+e_0\cos E}\right)\left(1+\frac{1}{2}\frac{e_0^2}{1-e_0^2}\sin^2 E\right)\right\}\mathrm{d}E \quad (53)$$

and

$$\frac{\mathrm{d}e_0}{\mathrm{d}t} = -\frac{2\delta\rho_p a_0(1-e_0^2)}{T_0}\exp\{\beta(r_p-a_0)\}\int_0^{2\pi}(1-e_0^2\cos^2 E)^{\frac{1}{2}}\exp(\beta a_0 e_0\cos E)$$
$$\times\left\{\frac{\cos E}{1-e_0\cos E}-\frac{2\omega a_0^2\cos i_0}{H_0}\left[\frac{(1-e_0^2)\cos E}{1+e_0\cos E}\left(1+\frac{1}{2}\frac{e_0^2}{1-e_0^2}\sin^2 E\right)-\frac{e_0}{4}\sin^2 E\right]\right\}\mathrm{d}E. \quad (54)$$

Results equivalent to these with $\omega = 0$ have been given previously by Elyasberg (1959) and Cook *et al.* (1960),† and with $\omega \neq 0$ by Cook & Plimmer (1960), who assumed small eccentricities and evaluated the integrals by expanding all but the exponential factor in the integrands in ascending powers of e_0. These integrals can be evaluated to a high degree of accuracy when βa_0 is large, for all values of e_0, by expanding in ascending powers of $\sin^2 E$, a process which is suggested by the method of steepest descents. By using this method, we have, for example

$$\int_0^{2\pi}(1-e_0^2\cos^2 E)^{\frac{1}{2}}\exp(\beta a_0 e_0\cos E)\,\mathrm{d}E$$
$$= (1-e_0^2)^{\frac{1}{2}}\int_0^{2\pi}\left(1+\frac{1}{2}\frac{e_0^2}{1-e_0^2}\sin^2 E-\dots\right)\exp(\beta a_0 e_0\cos E)\,\mathrm{d}E$$
$$= 2\pi(1-e_0^2)^{\frac{1}{2}}\left\{I_0(\beta a_0 e_0)+\frac{e_0}{2\beta a_0(1-e_0^2)}I_1(\beta a_0 e_0)\right\}\left\{1+O\left(\frac{e_0^2}{\beta^2 a_0^2(1-e_0^2)^2}\right)\right\}, \quad (55)$$

on making use of the integral representation of the modified Bessel functions of the first kind in the form (Watson 1952)

$$\int_0^{\pi}\sin^{2n}E\exp(\xi\cos E)\,\mathrm{d}E = \frac{(n-\frac{1}{2})!\,(-\frac{1}{2})!}{(\frac{1}{2}\xi)^n}I_n(\xi). \quad (56)$$

In the following analysis, we shall also make use of the formula

$$\int_0^{\pi}\cos E\sin^{2n}E\exp(\xi\cos E)\,\mathrm{d}E = \frac{(n-\frac{1}{2})!\,(-\frac{1}{2})!}{(\frac{1}{2}\xi)^n}I_{n+1}(\xi). \quad (57)$$

The O-term in (55) is very small; for an orbit with $e_0 = 0{\cdot}5$, perigee height 400 km, and with $\beta = 1/80\,\mathrm{km}^{-1}$, for example, the O-term is less than 2×10^{-5}, and it becomes smaller as e_0 becomes smaller. Thus further terms in the expansion (55), can be safely neglected, the resulting errors being smaller than those already encountered in using (42). For some purposes it may be sufficient to use only the dominant terms in such expansions; the error then is usually less than 1 %.

† Cook *et al.* incorporated the effects of atmospheric rotation in their equations by modifying the definition of δ, so their δ differs from that used here.

As a further aid to the evaluation of these integrals, we notice that the factors

$$1+\frac{1}{2}\frac{e_0^2}{1-e_0^2}\sin^2 E$$

in (53) and (54) may be replaced by

$$\left(\frac{1-e_0^2\cos^2 E}{1-e_0^2}\right)^{\frac{1}{2}}$$

without incurring greater errors than those indicated in (55). Also, in (53) and (54), we can neglect those terms containing $\omega/\beta a_0$ as a factor which arise when evaluating the integrals, because they too give very small contributions.

In the following analysis, the Bessel functions are written simply as I_0, etc., since the argument is always $\beta a_0 e_0$.

By using the method described above to evaluate the integrals, and making use of (9), equations (53) and (54) become

$$\frac{\mathrm{d}H_0}{\mathrm{d}t} = -\frac{\delta\rho_0 H_0^2}{a_0}\exp\{\beta(r_p-a_0)\} \times\left\{I_0+\frac{e_0}{2\beta a_0(1-e_0^2)}I_1-\frac{2\omega a_0^2\cos i_0}{H_0}[(1+e_0^2)I_0-2e_0 I_1]\right\} \tag{58}$$

and

$$\frac{\mathrm{d}e_0}{\mathrm{d}t} = -\frac{2\delta\rho_p H_0}{a_0}\exp\{\beta(r_p-a_0)\} \times\left\{\left(1-\frac{2-e_0^2}{2\beta a_0(1-e_0^2)}\right)I_1+\left(1-\frac{1}{2\beta a_0(1-e_0^2)}\right)e_0 I_0-\frac{2\omega a_0^2(1-e_0^2)\cos i_0}{H_0}(I_1-e_0 I_0)\right. . \tag{59}$$

The rates of change of $\mathbf{n}_0$ and $\mathbf{a}_0$, giving the angular velocities of the mean orbital plane and the direction of perigee, can be obtained from (51) and (52) by multiplying vectorially twice by $\mathbf{n}_0$ and $\mathbf{a}_0$ respectively. Then we have

$$\frac{\mathrm{d}\mathbf{n}_0}{\mathrm{d}t} = \frac{\delta\rho_p\omega a_0^3}{H_0 T_0}\exp\{\beta(r_p-a_0)\}\int_0^{2\pi}(1-e_0^2\cos^2 E)^{\frac{1}{2}}\exp(\beta a_0 e_0\cos E) \times\left\{1-\frac{\omega a_0^2(1-e_0^2)\cos i_0}{H_0}\left(\frac{1-e_0\cos E}{1+e_0\cos E}\right)\right\} \times\{(\cos E-e_0)^2(\mathbf{b}_0.\mathbf{k})\,\mathbf{a}_0-(1-e_0^2)\sin^2 E(\mathbf{a}_0.\mathbf{k})\,\mathbf{b}_0\}\wedge\mathbf{n}_0\,\mathrm{d}E \tag{60}$$

and

$$\frac{\mathrm{d}\mathbf{a}_0}{\mathrm{d}t} = -\frac{\delta\rho_p\omega a_0^3(1-e_0^2)}{H_0 T_0}\exp\{\beta(r_p-a_0)\}\int_0^{2\pi}(1-e_0^2\cos^2 E)^{\frac{1}{2}}\exp(\beta a_0 e_0\cos E) \times\left\{1-\frac{\omega a_0^2(1-e_0^2)\cos i_0}{H_0}\left(\frac{1-e_0\cos E}{1+e_0\cos E}\right)\right\}\sin^2 E(\mathbf{a}_0.\mathbf{k})\,\mathbf{b}_0\wedge\mathbf{a}_0\,\mathrm{d}E. \tag{61}$$

These integrals can be evaluated by the same method as before, and on neglecting terms in ω^2 and $\omega/\beta a_0$, and using (9) again, (60) becomes

$$\frac{\mathrm{d}\mathbf{n}_0}{\mathrm{d}t} = \delta\rho_p\omega a_0\exp\{\beta(r_p-a_0)\} \times\{[(1+e_0^2)I_0-2e_0 I_1](\mathbf{b}_0.\mathbf{k})\,\mathbf{a}_0-\tfrac{1}{2}(1-e_0^2)(I_0-I_2)(\mathbf{a}_0.\mathbf{k})\,\mathbf{b}_0\}\wedge\mathbf{n}_0, \tag{62}$$

which shows that the mean orbital plane rotates about a line in its own plane, parallel to the vector in the braces on the right-hand side of this equation. For all except very small eccentricities (say $\beta a_0 e_0 > 10$) and when $\mathbf{b}_0 . \mathbf{k}$ is not small compared with $\mathbf{a}_0 . \mathbf{k}$, the term containing the factor $(I_0 - I_2)$ is small compared with the other, because the Bessel functions are nearly equal, and then the mean orbital plane rotates about the major axis approximately.

The rate of change of $\mathbf{a}_0$ in the mean orbital plane is obtained from (61) by subtracting the contribution from the rotation of the mean orbital plane; the only term in (60) which contributes to this rotation is that involving the vector $\mathbf{b}_0 \wedge \mathbf{n}_0$, and it is easy to see that this cancels the right-hand side of (61) identically. Thus there is no movement of the major axis in the mean orbital plane as a result of atmospheric resistance.† Of course, for the actual oblate earth, there is a rotation of $\mathbf{a}_0$ in the mean orbital plane as well as an extra rotation of the mean orbital plane, as shown (to the first order) in § 4. These rotations must be added to that due to atmospheric resistance in order to obtain more complete expressions for the rates of change of $\mathbf{n}_0$ and $\mathbf{a}_0$.

The equations obtained above are so complicated that an analytical solution in closed form does not seem to be feasible, but they can be solved numerically without undue difficulty. It turns out that (58) and (59) are not the most convenient for numerical integration and that a better pair is obtained from the rates of change of perigee distance, $a_0(1-e_0)$, which we denote by x_0, and the square of the eccentricity, e_0^2. The equation for $\mathrm{d}x_0/\mathrm{d}t$ is

$$\frac{\mathrm{d}x_0}{\mathrm{d}t} = -2\delta\rho_p H_0 \left(\frac{1-e_0}{1+e_0}\right) \exp\{\beta(r_p - a_0)\}$$
$$\times \left\{\left(1 + \frac{e_0}{2\beta a_0(1-e_0^2)} - \frac{2\omega a_0^2(1+e_0)^2 \cos i_0}{H_0}\right)(I_0 - I_1) + \frac{I_1}{\beta a_0(1-e_0)}\right\}, \quad (63)$$

and the equation for $\mathrm{d}(e_0^2)/\mathrm{d}t$ is easily obtained from (49).

When $\beta a_0 e_0$ is sufficiently large (say > 10), the factor $(I_0 - I_1)$ in (63) is small compared with I_0 and the perigee distance changes very slowly; also a remarkable property of e_0^2 is that it varies almost linearly with time throughout the life of the satellite. These properties of x_0 and e_0^2 make the equations for them very suitable for numerical integration because a large integration interval can be used, except towards the end of the satellite's life when perigee distance decreases rapidly. The property of e_0^2 referred to above was noticed by Cook *et al.* (1960) for small eccentricities.

The dominant terms in (58), (59) and (63) for a non-rotating atmosphere were found by Nonweiler (1958) in a paper which does not seem to have attracted the attention that it deserves. Nonweiler used the true anomaly as the variable of integration, which leads to the same results as the use of the eccentric anomaly, but he did not discover the term $I_1/\beta a_0(1-e_0)$ in (63), which is important when the eccentricity is not very small.

† When the oblateness of the atmosphere is taken into account, there is a rotation of the major axis, as has been shown by Cook (1961).

The results of some typical numerical integrations of (63) and the equation for e_0^2 are shown in figure 1; the initial values of perigee height and eccentricity are 400 km and 0·6, respectively, $\beta = 1/80\,\mathrm{km}^{-1}$, and the resistance is adjusted to give a lifetime of 5000 days when $i_0 = 90°$, that is, for a polar orbit. The curves show the variations of perigee height and e_0^2 with time for $i_0 = 0°$, 90°, and 180°. The theoretical lifetimes for $i_0 = 0°$ and 180° are 5773 days and 4409 days, respectively. The orbital inclinations have been taken to be constants.

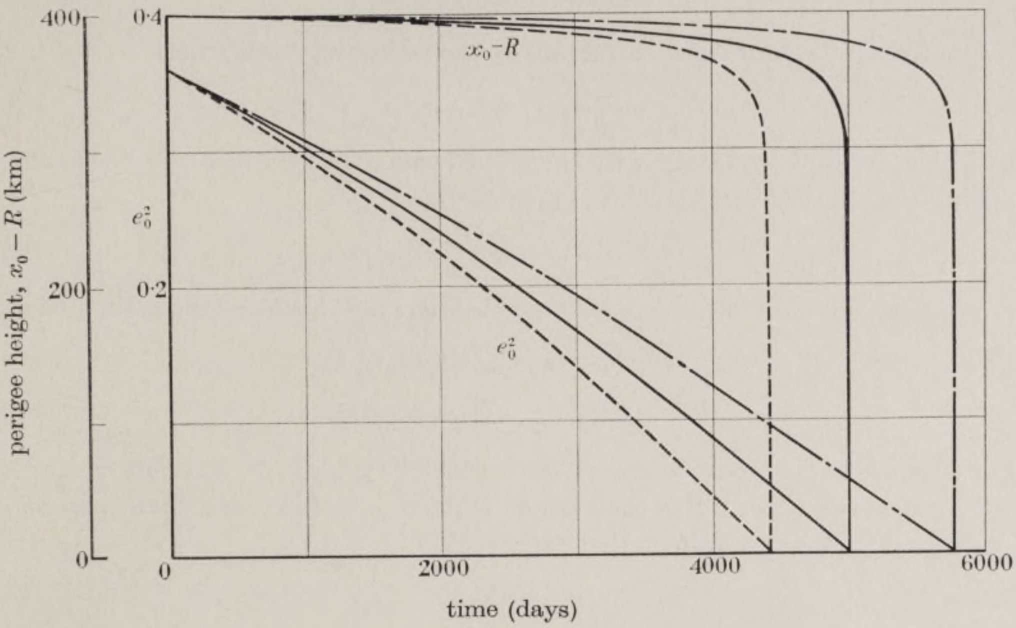

FIGURE 1. Variation of perigee height and e_0^2 with time for some typical orbits. —·—·, $i_0 = 0°$; ——, $i_0 = 90°$; ----, $i_0 = 180°$. $\beta = 1/80\ \mathrm{km}^{-1}$.

It will be seen that e_0^2 varies linearly with time, on the scale of the curves, when $e_0^2 < 0{\cdot}1$, that is, when e^2 is less than about 0·3. An estimate of the remaining lifetime, t_L,of the satellite at any stage can be obtained by linear extrapolation, using the empirical formula

$$t_L = -\tfrac{1}{2}e_0/(\mathrm{d}e_0/\mathrm{d}t); \tag{64}$$

the values of t_L obtained from this formula are found to agree with the *theoretical* lifetimes within about $\pm 2\,\%$ provided that $e_0 < 0{\cdot}3$.

These numerical results will be in error for several reasons, the most important of which is that the atmospheric density does not really vary exponentially. The greatest effects of this departure from reality will occur when the perigee height starts to decrease rapidly near the end of the satellite's lifetime. Because the density increases more rapidly than exponentially as the perigee height decreases, this effect will shorten the life of the satellite in comparison with the results shown in figure 1, but if the value of ρ_p and β appropriate to the actual perigee height are used in (64), the effect is compensated, and the formula may still yield useful predictions.

6. Equations for alternative elements

It is usual to describe the geometry of an orbit in terms of the semi-major axis, a_0, the eccentricity, e_0, the longitude of the ascending node, Ω_0, the inclination, i_0, and the longitude of perigee, ϖ_0. Of these, the rate of change of e_0 is already determined, and the rates of change of a_0 and i_0 are easily derived from (8) in the case of a_0 and from the equation $\cos i_0 = \mathbf{k}.\mathbf{n}_0$ in the case of i_0, the latter giving

$$\mathrm{d}i_0/\mathrm{d}t = -\mathbf{k}.\dot{\mathbf{n}}_0/\sin i_0. \tag{65}$$

If $\mathbf{l}_0$ is the unit vector in the direction of the ascending node, then

$$\sin i_0 \mathbf{l}_0 = \mathbf{k} \wedge \mathbf{n}_0 \tag{66}$$

from which, by differentiating with respect to time, we have the rate of change of Ω_0, being the angular velocity of $\mathbf{l}_0$, in the form

$$\dot{\Omega}_0 = \mathbf{k}.\mathbf{l}_0 \wedge \dot{\mathbf{l}}_0 = \mathbf{k}.\mathbf{n}_0 \wedge \dot{\mathbf{n}}_0/\sin^2 i_0. \tag{67}$$

Also, the angular velocity of $\mathbf{a}_0$ relative to $\mathbf{l}_0$ is $\mathbf{a}_0 \wedge \mathbf{a}_0 - \mathbf{l}_0 \wedge \dot{\mathbf{l}}_0$, so the rate of change of ϖ_0 is

$$\begin{aligned}\dot{\varpi}_0 &= \mathbf{n}_0.(\mathbf{a}_0 \wedge \dot{\mathbf{a}}_0 - \mathbf{l}_0 \wedge \dot{\mathbf{l}}_0) + \dot{\Omega}_0 \\ &= \mathbf{b}_0.\dot{\mathbf{a}}_0 + (1-\cos i_0)\,\dot{\Omega}_0. \end{aligned} \tag{68}$$

For reasons of economy of space, the equations for $\mathrm{d}i_0/\mathrm{d}t$, $\dot{\Omega}_0$, and $\dot{\varpi}_0$ are not written out in full here. When they are so written out, it is found that expressions for $\mathbf{a}_0.\mathbf{k}$ and $\mathbf{b}_0.\mathbf{k}$ are required: these are

$$\mathbf{a}_0.\mathbf{k} = \sin i_0 \sin(\varpi_0 - \Omega_0) \tag{69}$$

and

$$\mathbf{b}_0.\mathbf{k} = \sin i_0 \cos(\varpi_0 - \Omega_0). \tag{70}$$

The author is indebted to Mr D. G. King-Hele for some valuable comments on the first draft of this paper, and for drawing his attention to the work of Musen.

References

Allan, R. R. 1961 *Nature, Lond.* **190**, 615.
Brenner, J. L. & Latta, G. E. 1960 *Proc. Roy. Soc.* A, **258**, 470.
Cook, G. E., King-Hele, D. G. & Walker, D. M. C. 1960 *Proc. Roy. Soc.* A, **257**, 224.
Cook, G. E. & Plimmer, R. N. A. 1960 *Proc. Roy. Soc.* A, **258**, 516.
Cook, G. E. 1961 *Proc. Roy. Soc.* A, **261**, 246.
Elyasberg, P. E. 1959 *Iskusstvennyye Sputniki Zemli,* no. 3, Academy of Sciences USSR (Trans: NASA TT F-47, 1960).
King-Hele, D. G. 1958 *Proc. Roy. Soc.* A, **247**, 49.
King-Hele, D. G. 1959 *Nature, Lond.* **184**, 1267.
Musen, P. 1954 *Astron. J.* **59**, 262.
Nonweiler, T. R. F. 1958 *J. Brit. Interplan. Soc.* **16**, 368.
Struble, R. A. 1960 A rigorous theory of satellite motion. *Tenth International Congress of Applied Mechanics, Stresa, Italy.* Also: 1961 *Arch. Rat. Mech. Anal.* **7**, 87.
Watson, G. N. 1952 *Bessel functions,* 2nd ed. Cambridge University Press.