Elementary derivation of the perturbation equations of celestial mechanics

Joseph A. Burns*
Ames Research Center, National Aeronautics and Space Administration, Moffett Field, California 94035
(Received 6 January 1976)

The equations of celestial mechanics that govern the time rates of change of the orbital elements are completely derived using elementary dynamics, starting from only Newton's equation and its solution. Two orbital equations and the four most meaningful orbital elements—semimajor axis a, eccentricity e, inclination i, and longitude of pericenter Ω—are written in terms of the orbital energy E and angular momentum $\mathbf{H}$ per unit mass. The six resulting equations are differentiated with respect to time to see the effect on the orbital elements of small changes in E and $\mathbf{H}$. The usual perturbation equations in terms of disturbing force components are then derived by computing the manner in which perturbing forces change E and $\mathbf{H}$. The results are applied in a qualitative discussion of the orbital evolution of particles in nonspherical gravitational fields, through atmospheres, and under the action of tides.

I. INTRODUCTION

Celestial mechanics, once of interest only to erudite applied mathematicians, has been found to be important in today's space age to a wide class of scientists. It is used by the engineer for the precise missions of modern spacecraft, by the geophysicist to learn of the Earth's interior density from variations in satellite orbits, by the observational astronomer for the prediction and explanation of occultation and eclipse phenomena, by the theoretical astrophysicist to build interpretative models for the evolution of binary star systems composed of exotic stellar classes, and by the cosmogonist to reconstruct the solar system's primordial configuration from current data. For all of these scientists, it is the perturbation equations of celestial mechanics—the six equations which describe how an orbit evolves under the action of small disturbing forces—that are especially useful.

Despite numerous applications, the perturbation equations are often poorly understood by those who employ them. It is difficult for many noncelestial mechanicians to completely appreciate the perturbation equations because their derivation historically relies on somewhat sophisticated classical mechanics. These fundamental equations are classically derived (see, e.g., the texts of Moulton,[1] Plummer,[2] or Brown[3]) either through a perturbed Hamilton–Jacobi equation or by an inversion technique, the variation of the elements. While each scheme is mathematically elegant, neither provides much physical insight: the end result is that frequently the nonspecialist applies equations that he has accepted on faith. Even among the post-Sputnik texts concerning celestial mechanics and astrodynamics, most (see, e.g., Deutsch,[4] Roy,[5] or Brouwer and Clemence[6]) give derivations developed more than a century ago. Sterne,[7] on the other hand, while emphasizing the perturbed Hamilton–Jacobi technique, does interpret four of the perturbation equations in a manner similar to our dynamical approach below. Only Danby[8] provides a complete physical basis for the perturbation equations by considering the direct effect of impulses on the parameters which describe an orbit. Most contemporary research which considers the effects of particular disturbing forces does not attempt to develop physical intuition. Blitzer[9] has considered, however, a specific perturbation problem—the effect of the Moon and Sun on the motion of a near Earth satellite—using basic dynamical concepts, such as we will, but does not treat the general perturbation problem. Blitzer[10] has also presented for the general reader a valuable review of reference material on the dynamics of satellite orbits. In the historical past, all pedagogic discussions[11] of the perturbation equations relied heavily on intricate geometrical arguments of the properties of ellipses. These have little appeal today.

Here we wish to present a derivation of the perturbation equations of celestial mechanics that is elementary, yet complete of itself, starting from Newton's equation and its solution, and using only fundamental concepts familiar to most sophomores in physics. Our derivation and the form of the equations will make clear their physical basis; in that way they may improve the understanding of the professional scientist while at the same time providing a significant application of elementary dynamics for the undergraduate. For the first time, to the author's knowledge, the perturbation equations will be written in terms of only the orbital energy and angular momentum, which are constants of the motion in the unperturbed problem, and their time rates of change. Such equations often have a simpler interpretation than the usual perturbation equations, particularly if one is interested in learning of the qualitative effects of specific disturbing forces. From this set of equations, knowing how energy and angular momentum change with time, we will derive without difficulty the classical perturbation equations of celestial mechanics in terms of the disturbing forces. Application of the perturbation equations in terms of energy and angular momentum will then be made to limn the effects of a few of the more common perturbation forces in celestial mechanics.

II. PRELIMINARIES

A. Unperturbed Orbit

In order to define quantities for later use, we first consider a particle m moving in the r^{-2} gravitational field of a fixed point mass M. For a force $\mathbf{F}$ per unit mass, Newton's equation of motion is

$$\mathbf{F} = \ddot{\mathbf{r}} = -\mu r^{-3}\mathbf{r}, \tag{1}$$

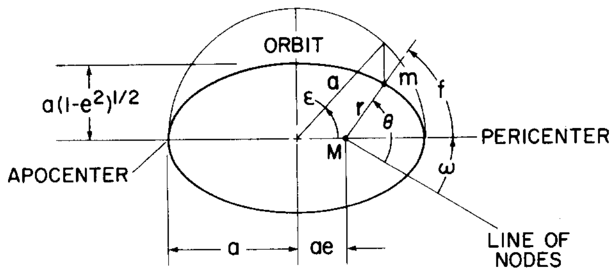

Fig. 1. Diagram of the orbit plane of an elliptic orbit, showing the definition of the orbital elements (a,e,ω), the true anomaly f, the eccentric anomaly ϵ, and pericenter location.

where $\mu = GM$, with G being the universal gravitational constant; $\mathbf{r}$ is the position vector from M to m, and the dot signifies differentiation with respect to time.

Because the force in Eq. (1) is radial, $\mathbf{r} \times \ddot{\mathbf{r}} = \mathbf{d}(\mathbf{r} \times \dot{\mathbf{r}})/dt = 0$. Thus, $\mathbf{H}$, which is defined as $\mathbf{r} \times \dot{\mathbf{r}}$, must be a constant vector; it is the particle's angular momentum per unit mass about M. The constancy of $\mathbf{H}$ implies that the motion takes place in the orbit plane, a fixed plane which is normal to $\mathbf{H}$ and contains M; it also requires that the vector's magnitude,

$$H = r^2\dot{\theta}, \tag{2}$$

be conserved. The angle θ is the position angle measured from some fixed line in the plane; below we will choose this line to be the line of nodes which is later defined in Fig. 2.

Since $\nabla \times \mathbf{F} = 0$, the line integral $\int \mathbf{F} \cdot \mathbf{dr}$ is independent of the path and gives the potential energy per unit mass to be $-\mu/r$. Therefore, the total energy per unit mass is conserved and is

$$E = (1/2)\dot{\mathbf{r}} \cdot \dot{\mathbf{r}} - \mu/r. \tag{3}$$

The orbit $r = r(\theta)$ can be found by solving the radial portion of Eq. (1), making use of Eq. (2). The solution, that of a conic section, is developed in many elementary texts,[12–14] and so we merely quote it:

$$r = p/[1 + e\cos(\theta - \omega)]. \tag{4}$$

The quantities e and ω are constants determined from the initial conditions. The parameter

$$p = H^2/\mu \equiv a(1 - e^2). \tag{5}$$

The right-hand equation defines a and is made to allow a simple geometric description of the orbit; it is not necessary otherwise at this time and will be derived later.

The argument of the cosine term in Eq. (4) is used to introduce the true anomaly,

$$f \equiv \theta - \omega, \tag{6}$$

the particle's angular position measured from pericenter (the point of closest approach to M; see Figs. 1 and 2).

The quantities a, e, and ω are constants which have the simple geometric representation of Fig. 1 for the case of interest in most celestial mechanics problems, namely an elliptical solution (where the eccentricity $e < 1$). The semimajor axis of the ellipse is a, where ω is the argument of pericenter.

One can geometrically show[15] that an equivalent solution is

$$r = a(1 - e\cos\epsilon), \tag{7}$$

where ϵ is the eccentric anomaly, which is the position angle measured from pericenter of a point directly above the actual particle on a circle of radius a (see Fig. 1). Comparing Eq. (7) to Eq. (4) relates the true anomaly to the eccentric anomaly by

$$\cos\epsilon = \frac{e + \cos f}{1 + e\cos f} \tag{8}$$

and, thus,

$$df = \frac{(1 - e^2)^{1/2}}{1 - e\cos\epsilon}\, d\epsilon. \tag{9}$$

Differentiation of Eq. (4) with Eqs. (2) and (6) yields the particle's radial velocity as

$$\dot{r} = (H/p)e\sin f, \tag{10}$$

while its transverse velocity is

$$r\dot{\theta} = (H/p)(1 + e\cos f). \tag{11}$$

B. Orbital elements

To completely describe the particle's orbit as a function of time, six constants—the orbital elements— are required: they correspond to the six initial conditions $(\mathbf{r}_0, \dot{\mathbf{r}}_0)$ appearing in the general solution of any three-dimensional, second-order ordinary differential equation of motion such as (1). Three orbital elements, a, e, and ω, have already been presented. A fourth is needed to completely describe the two-dimensional motion of the particle in the orbital plane. It, in essence, is used to specify a reference time which locates the particle: frequently, τ, the time of pericenter passage, is chosen. The remaining two orbital elements orient the orbital plane as shown in Fig. 2. The inclination i gives the angle between the orbital plane and some arbitrary fixed plane. The latter usually is taken to be the equatorial plane of the central body for two-body problems or the orbital plane of the central body in cases where the central body itself is in orbit about another mass. The longitude of the ascending node Ω locates the point at which the orbit pierces the arbitrary plane on its upward path. The line of nodes is the intersection of the arbitrary plane and the orbit plane.

The above set of orbital elements $(a,e,i,\omega,\Omega,\tau)$ is selected as the most understandable; however, other choices, which also uniquely define the orbit, are possible and often used. For example, the longitude of pericenter $\tilde{\omega} \equiv \omega + \Omega$, a sum

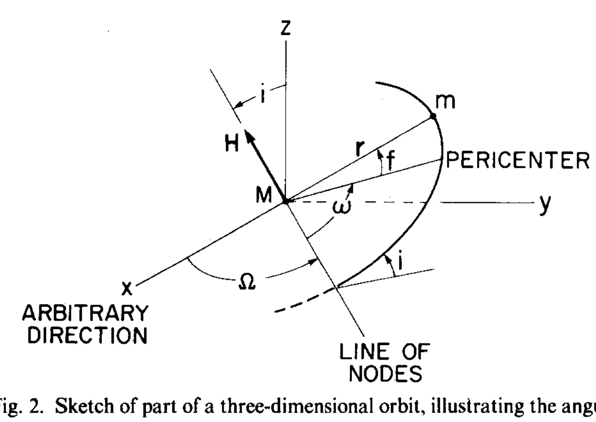

Fig. 2. Sketch of part of a three-dimensional orbit, illustrating the angular momentum vector $\mathbf{H}$ and how it determines the orbital elements Ω and i.

of noncoplanar angles, frequently replaces ω. Also the mean motion n (the time average of the orbital angular velocity); which is given by Kepler's third law as derived immediately below [Eq. (14)], sometimes takes the place of a. A variant of $\chi \equiv -n\tau$, the mean longitude of pericenter, is preferred to τ for reasons which will appear later.

C. Kepler's equation

An orbital equation that explicitly contains the time is needed in order to find the perturbation equation for τ. Kepler's equation is such an equation; it can be derived from the integral of Eq. (2) evaluated from pericenter to a general time t:

$$H\int_{\tau}^{t} dt = \int_{0}^{f} r^2\, df. \tag{12}$$

By means of Eqs. (7), (9), and (16) below, this is

$$t - \tau = a^{3/2}\mu^{-1/2}(\epsilon - e\sin\epsilon). \tag{13}$$

After one orbital period P, Eq. (13) reads

$$P/2\pi \equiv n^{-1} = a^{3/2}\mu^{-1/2}, \tag{14}$$

which is Kepler's observationally discovered third law. Substituting Eq. (14) into Eq. (13), we arrive at Kepler's equation

$$n(t-\tau) = \epsilon - e\sin\epsilon. \tag{15}$$

D. Energy and angular momentum of the orbit

We now wish to express the orbital elements in terms of the orbital energy and angular momentum. Substitution of Eq. (11), combined with Eq. (2), and the time derivative of Eq. (10) into Eq. (1), and the use of the first and last parts of Eq. (5) give

$$H = [\mu a(1-e^2)]^{1/2}, \tag{16}$$

which supports the definition given on the right-hand side of Eq. (5). Squaring Eqs. (10) and (11), and placing them in Eq. (3) with Eqs. (4) and (16), we find

$$E = -\mu/(2a); \tag{17}$$

that is, the semimajor axis is determined solely by the total orbital energy per unit mass. Inverting Eq. (16) and making use of Eq. (17) show that the orbital eccentricity is specified uniquely by a combination of the energy per unit mass and the orbital angular momentum per unit mass:

$$e = (1 + 2H^2E\mu^{-2})^{1/2}. \tag{18}$$

Similarly, as can be seen from Fig. 2, i and Ω are given by components of the angular momentum vector. We define the inclination by

$$\cos i = H_z/H. \tag{19}$$

The projection of H onto the xy plane is normal to the line of nodes, and so the longitude of the ascending node is given by

$$\tan\Omega = -H_x/H_y, \tag{20}$$

where H_x, H_y, and H_z are the components of $\mathbf{H}$. The (x,y,z) inertial coordinate system is fixed to M with the z axis normal to the arbitrary plane discussed above; the x axis points in an arbitrary direction, usually selected to be the direction of the vernal equinox for motions in the solar system. Equations (17)–(20) give four orbital elements in terms of the four pieces of information contained in $\mathbf{H}$ and E. It is perhaps worth noting that $(E,H,H_z;\ \omega,\Omega,\tau)$ are the Jacobi elements, one set of canonical constants for the unperturbed Hamilton–Jacobi equation.[16]

III. PERTURBED PROBLEM

A. Osculating orbital elements

As already noted, the fundamental problem of celestial mechanics is to find how an orbit, which if unperturbed would be a conic section, is modified by a disturbing force $\mathbf{dF}$ taken to be small in comparison to $\mathbf{F}$. This is most commonly done by considering an auxiliary orbit called the *osculating orbit*. This orbit is the (elliptical) path along which the particle would move if $\mathbf{dF}$ suddenly vanished, say at $t = t_1$, and the particle continued along its way under the action of $\mathbf{F}$ alone with $\mathbf{r}(t_1)$ and $\dot{\mathbf{r}}(t_1)$ as initial conditions. Because $\mathbf{dF}$ *is* present, the osculating orbit changes with time but does so slowly since $\mathbf{dF}$ is small. The osculating orbit is so called because it "kisses" the real orbit at time t_1. The orbital elements which specify the osculating orbit are called the *osculating elements*. Once these elements are known as functions of time, the particle's position is determined at any time.

The problem to be solved is then: find the equations governing the time rates of change of the set $(a,e,i,\omega,\Omega,\tau)$ caused by the action of the small disturbing force

$$\mathbf{dF} = \mathbf{R} + \mathbf{T} + \mathbf{N} = R\mathbf{e}_R + T\mathbf{e}_T + N\mathbf{e}_N, \tag{21}$$

where the $\mathbf{e}$'s represent an orthogonal unit vector triad, and the perturbing force is broken into its components; $\mathbf{R}$ is radially outwards along $\mathbf{r}$, $\mathbf{T}$ is transverse to the radial vector in the orbit plane (positive in the direction of motion of the particle), and $\mathbf{N}$ is normal to the orbit plane in the direction $\mathbf{R} \times \mathbf{T}$.

B. Perturbation equations

Once the disturbing force $\mathbf{dF}$ is introduced, Eqs. (17)–(20), which give the orbital elements as functions of E and $\mathbf{H}$, are satisfied only instantaneously; that is, they hold when the orbital elements, as well as E and $\mathbf{H}$, are considered as functions of time. To find the relation between the time rates of change of the orbital elements and the variations of E and $\mathbf{H}$, we merely differentiate the defining equations with respect to time. For example, differentiating Eq. (17), we have

$$\frac{da}{dt} = (2a^2\mu^{-1})\dot{E} \tag{22}$$

as our first perturbation equation. We immediately see that perturbing forces which dissipate energy cause satellite orbits to shrink. To convert Eq. (22) to the usual form of a perturbation equation in which the disturbing forces appear, we recognize that $\dot{E}$ is the work done per unit mass on the body per unit time by the disturbing forces:

$$\frac{dE}{dt} = \dot{\mathbf{r}} \cdot \mathbf{dF} = \dot{r}R + r\dot{\theta}T, \tag{23}$$

according to Eq. (21); N does not appear since the motion

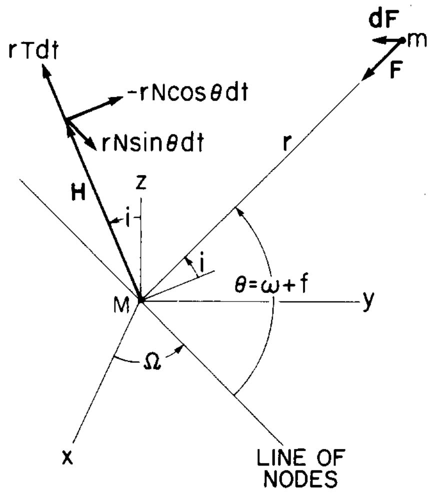

Fig. 3. Force diagram to depict the change of the angular momentum vector under the action of the disturbing force **dF**.

lies in the orbit plane. Substituting Eqs. (10) and (11) into Eq. (23), and that result into Eq. (22) gives

$$\frac{da}{dt} = 2\mu^{-1/2}a^{3/2}(1-e^2)^{-1/2} \times [\mathrm{Re}\sin f + T(1+e\sin f)]; \quad (24)$$

only forces lying in the orbit plane can change the orbit size.

Taking the time derivative of Eq. (18),

$$\frac{de}{dt} = (1/2)e^{-1}(e^2-1)(2\dot{H}/H + \dot{E}/E). \quad (25)$$

Since dH and dE can be of either sign, the two terms may compete with one another to determine whether the orbit is circularized with time or not. The time rate of change of angular momentum equals the applied moment:

$$\frac{\mathbf{dH}}{dt} = \mathbf{r}\times\mathbf{dF} = rT\mathbf{e}_N - rN\mathbf{e}_T. \quad (26)$$

Equation (26) requires that the magnitude of **H** change according to

$$\frac{dH}{dt} = rT, \quad (27)$$

since $-rN\mathbf{e}_T$, being perpendicular to **H**, merely varies the direction of **H**. Substitution of Eqs. (16), (17), (23), and (27) into Eq. (25) with Eqs. (4), (10), and (11) gives

$$\frac{de}{dt} = [a(1-e^2)\mu^{-1}]^{1/2} \times [R\sin f + T(\cos f + \cos\epsilon)], \quad (28)$$

the classical result. Again, only forces in the orbit plane can change the orbit shape.

Differentiating Eq. (19) yields

$$\frac{di}{dt} = [(H/H_z)^2 - 1]^{-1/2}(\dot{H}/H - \dot{H}_z/H_z). \quad (29)$$

It is seen from Fig. 3 that

$$\frac{dH_z}{dt} = rT\cos i - rN\cos\theta\sin i, \quad (30)$$

and so from Eqs. (16), (19), (27), (29), and (30) with Eqs. (4), (10), and (11)

$$\frac{di}{dt} = [a\mu^{-1}(1-e^2)]^{1/2} N\cos\theta/(1+e\cos f). \quad (31)$$

For clarity this can better be written

$$\frac{di}{dt} = \frac{rN\cos\theta}{H}, \quad (32)$$

where the numerator is the component of the torque which rotates **H** about the line of nodes (and which thereby moves the orbit plane). Forces in the orbit plane cannot change the plane's orientation.

The derivative of Eq. (20) provides

$$\frac{d\Omega}{dt} = (H^2 - H_z^2)^{-1}(H_x\dot{H}_y - H_y\dot{H}_x) \quad (33)$$

or

$$\frac{d\Omega}{dt} = (H\sin i)^{-1}(\sin\Omega\,\dot{H}_y + \cos\Omega\,\dot{H}_x). \quad (34)$$

Figure 3 illustrates that

$$\frac{dH_x}{dt} = r(T\sin i\sin\Omega + N\sin\theta\cos\Omega + N\cos\theta\cos i\sin\Omega) \quad (35)$$

and

$$\frac{dH_y}{dt} = r(-T\sin i\cos\Omega + N\sin\theta\sin\Omega - N\cos\theta\cos i\cos\Omega). \quad (36)$$

Equations (35) and (36) can be placed in Eq. (34) to produce, using Eqs. (16) and (4),

$$\frac{d\Omega}{dt} = [a\mu^{-1}(1-e^2)]^{1/2} \times N\sin\theta/[\sin i\,(1+e\cos f)]. \quad (37)$$

The dynamics contained in Eq. (37) are more simply understood when it is written

$$\frac{d\Omega}{dt} = \frac{rN\sin\theta}{H\sin i}, \quad (38)$$

where the denominator is the angular momentum which lies normal to the line of nodes in the xy plane and the numerator is the moment acting to precess the orbit plane.

The development of the final two perturbation equations is not as straightforward as those above since the orbital elements ω and τ are not explicit functions of E and **H**. We can, however, find their variations by returning to the equations which define the orbit. Rewriting Eq. (4) after substituting Eqs. (5) and (18),

$$H^2 = \mu r[1 + (1 + 2EH^2\mu^{-2})^{1/2}\cos(\theta-\omega)]. \quad (39)$$

If a disturbing force **dF** is considered to be applied for an instant, E, **H**, and ω change, but r does not since the particle location instantaneously remains the same:

$$\frac{d\omega}{dt} = \dot{\theta} + [r^{-1} - E(e\mu)^{-1}\cos(\theta-\omega)] \times 2H\dot{H}/[e\mu\sin(\theta-\omega)] - H^2(e\mu)^{-2}\dot{E}\cot(\theta-\omega). \quad (40)$$

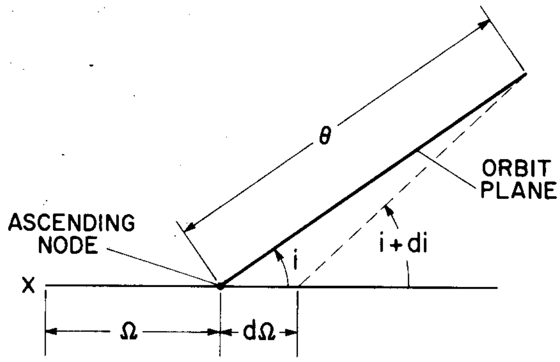

Fig. 4. View along the edge of the orbit plane to illustrate that changes in Ω affect θ.

The angular position θ of the particle changes instantaneously only because, as can be seen in Fig. 4, a change in the longitude of node Ω affects the evaluation of θ, which is the particle location measured from the line of nodes. For algebraic simplicity we have not replaced e in Eq. (40) by its form in E and H [cf. Eq. (18)]. Substituting Eqs. (23) and (27) in Eq. (40) gives

$$\frac{d\omega}{dt} = e^{-1}[a\mu^{-1}(1 - e^2)]^{1/2} \times [-R\cos f + T\sin f(2 + e\cos f)] - \cos i\,\dot{\Omega} \quad (41)$$

in terms of the disturbing forces. Forces in the orbit plane affect ω by changing a and e, as is apparent from the above derivation.

The same approach is taken with τ, except we instead differentiate Kepler's equation (15) with respect to time, substituting $\dot{\epsilon}$ as evaluated from the derivative of Eq. (7), to obtain

$$\frac{d\chi}{dt} = \left(-\frac{3}{2}ntE^{-1} + \frac{(1 - e^2)^{3/2}(2e - \cos f - e\cos^2 f)}{2e^2 \sin f(1 + e\cos f)E}\right)\dot{E} - (1 - e^2)^{3/2}H^{-1}e^{-2}\dot{H}\cot f, \quad (42)$$

recalling that $\chi \equiv n\tau$. Again for relative algebraic simplicity we leave e rather than substituting its value in terms of E and H [Eq. (18)]. In terms of forces, Eq. (42) is

$$\frac{d\tau}{dt} = \left[3(\tau - t)[a\mu^{-1}(1 - e^2)^{-1}]^{1/2} e\sin f + (1 - e^2)a^2\mu^{-1}\left(-\frac{\cos f}{e} + \frac{2}{1 + e\cos f}\right)\right]R + \left(3(\tau - t)[a\mu^{-1}(1 - e^2)^{-1}]^{1/2}(1 + e\cos f) + (1 - e^2)a^2\mu^{-1}e^{-1}\frac{\sin f(2 + e\cos f)}{1 + e\cos f}\right)T. \quad (43)$$

It is the presence of time on the right-hand side of Eqs. (42) and (43) which makes both χ and τ undesirable as the final orbital element since their time derivatives grow with time. As alluded to earlier, schemes using a variant of χ are used to eliminate this difficulty.[17-20] Briefly, what is done is that a new orbital element is defined which varies with time in such a manner that its time derivative does not contain time. However, since Eq. (43) was derived here strictly for the sake of developing a complete set of perturbation equations and since it will not be used in later discussions, we do not consider this complication further.

Equations (22), (25), (29), (33), (40), and (42) are the perturbation equations written in terms of energy per unit mass and angular momentum per unit mass, and their time rates of change. By knowing the latter quantities in terms of forces, we have transformed these perturbation equations into Eqs. (24), (28), (31), (37), (41), and (43), the usual perturbation equations of celestial mechanics written in terms of perturbing forces, which are classically derived by other means.[21-24] The perturbation equations can as well be expressed in terms of a disturbing function whose partial derivatives equal the force components we have used.

IV. APPLICATIONS

We now wish to apply the equations derived above to understand the gross features of several typical orbital evolution problems in celestial mechanics.[25] The primary concern in most astronomical applications is with the variations of (a,e,i) because these orbital elements approximately characterize the orbit and are those which appear in the perturbation equations themselves. Stress will be placed in our discussion on understanding very long time orbital changes as caused by changes in energy and angular momentum; however, the perturbation equations in terms of forces will be used when a clearer understanding is developed from them.

A. Motion in nonspherical gravitational fields

Since any gravitational field can be derived from a potential function, energy is conserved. This implies by Eq. (22) that the semimajor axis a does not vary for motion in any static, gravitational field. An axially symmetric gravitational field, such as that of an oblate planet, cannot change H since $T = 0$. It can produce changes in H_z; however, these will be short periodic effects when N is an odd function of θ as it is for an oblate planet. Thus, according to Eqs. (25) and (29), e and i are constant, ignoring short period effects. The line of nodes *does* rotate in the positive direction at a constant average rate for an oblate planet as can be seen from Eq. (38). Similar arguments apply for an orbit perturbed by a third body.

B. Atmospheric drag

A particle moving through an atmosphere dissipates energy and so a shrinks, according to Eq. (22). The eccentricity also decreases because the right-hand term in Eq. (25) is positive, considering Eqs. (23), (27), (10), and (11), since

$$\frac{(\dot{E}/E)}{|2\dot{H}/H|} = \frac{(1 + e\cos f)^2}{1 - e^2} > 1 \quad (44)$$

for any reasonable atmospheric drag law with a solely transverse force; this is true only until $e = 0$ when $|2\dot{H}/H| = \dot{E}/E$: the orbit thereafter remains circular. Since the drag lies in the orbit plane, $N = 0$ and the orbit plane stays fixed in space. Thus, the orbital inclination is constant [see Eq. (32)] and Ω is unchanged [see Eq. (38)].

C. Planetary tides

The tidal forces produced by a secondary body distort the shape of the primary body. To first order, the distortion is an elongation of the primary along the line connecting their

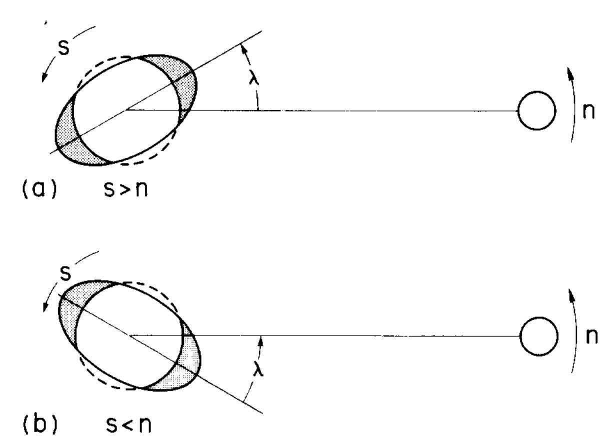

Fig. 5. Tidal effects. (a) Leading tides ($s > n$). A secondary body distorts the primary which then rotates from beneath it. The secondary body is pulled forward more by the near bulge than it is slowed by the far bulge. (b) Lagging tides ($s < n$). The satellite is pulled in by tidal forces.

centers (see Fig. 5). However, because energy dissipation delays the tidal response of the primary, the tidal bulge rotates from beneath the secondary that caused it. As shown in Fig. 5(a), for the usual case in the solar system (where the secondary's orbital angular velocity n is less than the primary's rotational angular velocity s), the tidal bulge leads the secondary body by λ and applies a force on the satellite in the positive **T** direction. Hence, energy and angular momentum are added to the satellite's orbit. The orbit grows. The eccentricity e usually increases in accordance with the discussions presented immediately above, [cf. Eq. (44)] recalling that $\dot{E}$ and $\dot{H}$ here are of opposite sign to the atmospheric drag case; however, this does not remain true for close satellites where higher order tidal terms can become important and reverse the signs of $\dot{E}$ and $\dot{H}$. This approach has been graphically depicted by Counselman[26] and Greenberg.[27] The orbital inclination and longitude of node undergo changes (unless $i = 0$) because the primary's rotation moves the delayed tidal bulge out of the particle's orbit plane along a constant longitude, and so $N \neq 0$; since N is an odd function of $(\theta - \lambda)$ and λ is small, the integrated change of i is much greater than the change of Ω [cf. Eqs. (32) and (38)].

When $n > s$, as it is for Phobos, the innermost satellite of Mars, and for Neptune's retrograde Triton, the situation in Fig. 5(b) obtains. The tidal forces then withdraw energy and angular momentum from the orbit: the orbit collapses and usually is circularized.

D. Satellite tides

In light of the growth of e for most satellites in the solar system under the action of planetary tides, it might appear surprising that virtually all close satellites have nearly circular orbits. The resolution of this quandary is contained in Eq. (25). All close satellites are known to be in synchronous rotation (i.e., their rotational angular velocities equal n) and so tides *in* the satellites produce only radial perturbations. Thus, H is conserved, but E decreases due to frictional losses during the tidal flexing of the satellite. To put it in terms of Eq. (23), work is done by the particle because the maximum tidal force occurs slightly after pericenter at the time of greatest tidal distortion. This means that more work is done *by* the particle in going from pericenter to apocenter than is done *on* the particle in returning from apocenter to pericenter. Since $N = 0$, the elements i and Ω remain constant. The orbit size does not significantly change since the energy loss is generally very small; a comparison of Eq. (22) with Eq. (25) shows that for nearly circular orbits $\dot{e} = e^{-1}[\dot{a}/(2a)]$, indicating that substantial variations in e can occur with little effect on a.

V. CONCLUSION

We have developed the perturbation equations of celestial mechanics, which might appear to be complicated at first glance, from a very elementary dynamical basis. By applying the insight gained from this new approach we have been able to predict simply the orbital evolution of particles acted on by some disturbing forces that are common in celestial mechanics.

ACKNOWLEDGMENTS

The author wishes to thank NASA for support under the Intergovernmental Personnel Act during his stay at Ames Research Center. The comments of Richard Durisen, Patrick Hamill, Richard Greenberg, and Kenneth Lebsock helped to significantly clarify this manuscript.

*Permanent address: Theoretical and Applied Mechanics, Center for Radiophysics and Space Research, Cornell University, Ithaca, NY 14853.

[1]E. W. Brown, *Introductory Treatise on Lunar Theory* (Cambridge U.P., London, 1896; Dover, New York, 1960).
[2]F. R. Moulton, *An Introduction to Celestial Mechanics* (MacMillan, New York, 1914, 2nd ed.; Dover, New York, 1970).
[3]H. C. Plummer, *An Introductory Treatise on Dynamical Astronomy* (Cambridge U.P., London, 1918; Dover, New York, 1960).
[4]R. Deutsch, *Orbital Dynamics of Space Vehicles* (Prentice-Hall, Englewood Cliffs, NJ, 1963).
[5]A. E. Roy, *The Foundations of Astrodynamics* (MacMillan, New York, 1965).
[6]D. Brouwer and G. M. Clemence, *Methods of Celestial Mechanics* (Academic, New York, 1961).
[7]T. E. Sterne, *An Introduction to Celestial Mechanics* (Interscience, New York, 1960).
[8]J. M. A. Danby, *Fundamentals of Celestial Mechanics* (MacMillan, New York, 1962).
[9]L. Blitzer, Am. J. Phys. **27**, 634 (1959).
[10]L. Blitzer, Am. J. Phys. **31**, 233 (1963).
[11]G. B. Airy, *Gravitation: An Elementary Explanation of the Principal Perturbations in the Solar System, 2nd ed.* (MacMillan, London, 1884; Neo, Peaks Island, ME, 1969).
[12]Reference 4, pp. 62–63.
[13]R. A. Becker, *Introduction to Theoretical Mechanics* (McGraw-Hill, New York, 1954), pp. 227–229.
[14]J. L. Meriam, *Dynamics* (Wiley, New York, 1966), pp. 121–123.
[15]Reference 8, pp. 126–127.
[16]Reference 4, p. 156.
[17]Reference 4, pp. 167–169.
[18]Reference 5, pp. 166–167.
[19]Reference 3, pp. 102–103.
[20]Reference 4, pp. 242–243.
[21]Reference 3, p. 151.
[22]Reference 4, p. 175.
[23]Reference 3, p. 108.
[24]Reference 4, p. 242.
[25]See the discussion by J. A. Burns, in *Planetary Satellites,* edited by J. A. Burns (University of Arizona Press, Tucson, 1976), which is concerned with the orbital evolution of natural satellite orbits.
[26]C. C. Counselman, III, Astrophys. J. **180**, 307 (1973).
[27]R. J. Greenberg, Icarus **23**, 51 (1974).