A dynamical picture of the oceanic tides

Eugene I. Butikov[a)]
Department of Physics, St. Petersburg State University, St. Petersburg 198904, Russia

(Received 5 November 2001; accepted 11 June 2002)

A detailed treatment of tide-generating forces is given, followed by a simplified dynamic theory of tidal waves. To clarify the underlying physics, we use a simple model of the ocean that consists of a water shell of uniform depth completely covering the globe. The treatment is appropriate for college and university undergraduate students studying introductory geophysics or astronomy, general physics, or intermediate mechanics. A computer simulation is developed to aid in understanding the properties of sun- or moon-induced tide-generating forces and of the stationary tidal waves created by these forces in the open ocean. © *2002 American Association of Physics Teachers.*
[DOI: 10.1119/1.1498858]

I. INTRODUCTION

All textbooks in introductory astronomy and many in physics and intermediate mechanics mention the existence of oceanic tides as an interesting manifestation of universal gravitation. Pedagogical papers devoted to the tides (see, for example, Refs. 1–9) testify to the fact that many teachers are interested in this topic, but are not satisfied with the clarity and correctness of the commonly accepted explanations of the physics of tidal phenomena. A review of textbooks and related literature shows that the most important aspects of the origin and properties of tides are often treated inaccurately or even erroneously. Much of the confusion over generating tides is related to the roles of the orbital motion of the moon and earth about their common center of mass and of the earth's axial rotation. In discussing the physics behind this phenomenon, authors usually explain (more or less successfully) why two tidal swells appear on the opposite sides of the globe. However, it is difficult to find a plausible explanation of the physical mechanism responsible for the phase shift between the zenith of the moon and the moment of high tide, which at some places approaches 90°. Misunderstandings also occur in discussions about the role of tidal friction in the retardation of axial rotations and in the evolution of orbital motions of the gravitationally coupled celestial bodies.

To clarify the basic physics underlying the tidal phenomena, we suggest a rather simple but rigorous treatment of the tide-generating forces, followed by a theory of the circulating tidal wave produced by these forces. This treatment uses a simplified model of the ocean consisting of a water shell of uniform depth entirely covering the globe. A computer simulation is developed to support the analytical treatment.[10] The simulation gives a dynamical picture of the forces and the tidal wave driven by these forces in the open ocean. This paper and the simulation are intended only to clarify the physical background of this natural phenomenon and do not assume to describe the complete picture. The purely theoretical quantitative description of tides for a given location on the earth, derived solely from first principles, is hardly possible because of the extremely complex structure of the oceans, the actual system that responds with tides and tidal currents to the well-known tide-generating forces.

The paper is organized as follows. First we discuss qualitatively the physical nature of the sun- and moon-induced tide-generating forces in a nonrotating geocentric frame of reference, deriving the mathematical expressions for these forces at an arbitrary point on the earth. Next the static (equilibrium) distortion of the ocean surface under these forces is determined. Then we show that the same expressions for the tidal forces are applicable on the rotating earth, and we discuss how these forces depend on time. We show that a uniform rotation of the system of tidal forces coupled with the apparent motion of the sun (moon) can be represented as a superposition of two oscillating quadrupole systems of forces whose axes make an angle of 45° with respect to one another. Each of these systems of forces generates a steady-state forced oscillation of the ocean (a standing wave). Next we treat the tidal wave circulating around the globe as a superposition of these standing waves. Finally the real-world complications of this simplified picture are discussed briefly, as well as the role of tidal friction in the evolution of the axial rotations and orbital revolutions of celestial bodies.

II. THE TIDE-GENERATING FORCES: AN ELEMENTARY APPROACH

The tides are manifested by alternating vertical displacements of the surface of the sea coupled with horizontal movements of the water that are called the *tidal currents*. It is well known that the tides are caused by the varying gravitational forces that the moon and sun exert on both the earth and its oceans. More exactly, the origin of tidal phenomena is related to the inhomogeneity (nonuniformity) of the lunar and solar gravitational fields across the globe.

The gravitational force the moon exerts on any body on the surface of the earth is much smaller than the gravitational force of the sun. However, because the moon is much closer to the earth than the sun, the inhomogeneity of the lunar gravitational field across the earth is considerably greater than that of the solar field. As a result, moon-induced tides are more than twice as great as sun-induced tides. Nevertheless, to arrive more easily at an understanding of the physical origin of tide-generating forces, we begin our analysis with sun-induced tides. These are somewhat simpler to explain because the center of mass of the sun–earth system very nearly coincides with the center of the sun.

We next divide the problem into two parts: First we discuss the origin and properties of tide-generating forces, after which we investigate qualitatively the much more complicated case of the dynamical effect that these time-varying forces have on the ocean. We note that much of the confu-

sion in the literature is related to the first (rather simple) part of this problem, which can be completely and unambiguously solved using Newtonian mechanics.

The earth as a whole moves with an acceleration relative to an inertial reference frame. This acceleration is produced by the gravitational attraction of the earth to the sun (and also to the moon and to all other celestial bodies). Although the earth travels in an almost circular orbit, its centripetal acceleration $\mathbf{a}_0$ in this orbital motion is generated by the gravitational pull of the sun and hence is just the *acceleration of free fall*, which is independent of the orbital velocity. The earth would move with the same acceleration were it freely falling in the gravitational field of the sun. What is important in this problem is the acceleration, not the orbital velocity, of the earth.

To better understand the tides, we first use a *nonrotating* geocentric reference frame. Although the origin of this frame moves approximately in a circle around the sun (more exactly, around the center of mass of the sun–earth system), the frame itself does not rotate because the directions of its axes are fixed relative to the distant stars. That is, the motion of this frame—revolution without rotation—is a translational (though nearly circular) motion. It reminds us of ''the circular motion of the frying pan'' in the hands of a cook (see Ref. 1). With respect to inertial space, all points of this reference frame move with an acceleration $\mathbf{a}_0$ whose magnitude and direction are the same for all the points. Any body of mass m whose motion is referred to this noninertial geocentric frame (for example, an earth satellite, or a drop of water in the ocean) is subject to the pseudoforce of inertia, $\mathbf{F}_{\text{in}}=-m\mathbf{a}_0$, which is independent of the position of the body relative to the earth. If the body were placed at the center of the earth, this pseudoforce would exactly balance the gravitational attraction of the body to the sun. In other words, if we consider the earth as a giant spaceship orbiting the sun, a body placed at the center of this ship would seem to be weightless with respect to the gravitation of the sun, just as astronauts on an orbital station seem to be weightless in the gravitational field of the earth.

The force of inertia, $\mathbf{F}_{\text{in}}=-m\mathbf{a}_0$, experienced by a body in the freely falling geocentric frame of reference (or in the frame that revolves without axial rotation about the sun–earth center of mass), has the same magnitude and direction everywhere on the earth. On the other hand, the gravitational pull of the sun, $\mathbf{F}_{\text{sun}}$, experienced by the body diminishes with its distance from the sun and is directed to the sun, and hence both the magnitude and direction of $\mathbf{F}_{\text{sun}}$ depend on the position of the body on the earth. Because the earth is an extended body, the pseudoforce $\mathbf{F}_{\text{in}}$ and the force $\mathbf{F}_{\text{sun}}$ are generally unequal and not exactly opposite, except at the center of the earth. The combined actions of the gravitational pull of the sun and the pseudoforce of inertia is the *tidal force*.

In other words, the tidal force at a given position near the earth equals the vector difference of the gravitational pull the sun exerts on an object at this position and the gravitational pull the sun would exert on this object were it at the center of the earth. We may avoid using a noninertial reference frame if we are not inclined to introduce the concept of the pseudoforce of inertia to students. In doing so, we can use a somewhat different language in the subsequent derivation of the tidal force: Instead of discussing the vector addition of the pull of the sun and the corresponding pseudoforce of inertia arising from the noninertial character of the reference frame, we can use instead an inertial frame, in which the tidal force can be found by the vector subtraction of the gravitational force of the sun on the body at its given location with the force of the sun on the body were it located at the center of the earth. Indeed, when viewing the situation on the earth from the inertial frame of reference, we can apply the Galilean law according to which, in the same gravitational field (here the field of the sun), all free bodies experience equal accelerations. Hence the earth as a whole and all free bodies on the earth, being subjected to almost the same solar gravitational field, are very nearly accelerated toward the sun. Consequently we do not particularly notice the influence of solar gravitation on what happens on earth. The small differences between the acceleration of the earth as a whole and of the earthly bodies depend on the distances of the bodies from the center of the earth because these differences are caused by the nonuniformity of the solar gravitational field over the extent of the earth.[11]

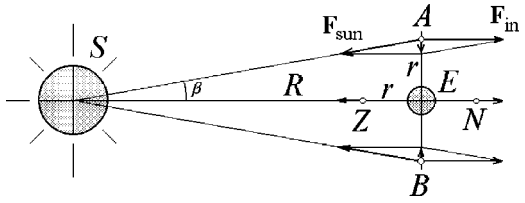

Fig. 1. Sun-induced tide-generating forces at different points A, B, Z, and N.

These differential effects of gravity give rise, in particular, to solar gravitational perturbations of an earth satellite's geocentric orbit. The tide-generating forces slightly distort the earth's gravitational pull that governs the satellite's motion so that after a revolution, the satellite does not return to the same point of the geocentric reference frame. On the surface of the earth, these same forces give rise to the tides. We emphasize that tidal forces are caused not by the sun's gravitational field itself, but rather by the nonuniformity of this field.

Figure 1 illustrates the origin and properties of the tide-generating forces produced by the sun. The free-fall acceleration of the earth E in the gravitational field of the sun S is $a_0=GM_{\text{sun}}/R^2$, where M_{sun} is the mass of the sun, and R is the sun–earth distance. The gravitational pull of the sun $\mathbf{F}_{\text{sun}}$ experienced by the body (for example, a satellite) at point A almost equals the force of inertia $\mathbf{F}_{\text{in}}$ in magnitude because the distances to the sun from the body and from the center of the earth are very nearly equal. However, at point A the direction of the gravitational force $\mathbf{F}_{\text{sun}}$ is not exactly opposite to the force of inertia $\mathbf{F}_{\text{in}}$. Thus their nonzero resultant, the tidal force $\mathbf{F}_A$ at point A, is directed toward the earth. Its magnitude equals $ma_0\beta=ma_0(r/R)=(GmM_{\text{sun}}/R^2)(r/R)$, where $\beta=r/R$ is the angle between the body and the center of the earth as seen from the sun. The tidal force $\mathbf{F}_B$ at the opposite point B equals $\mathbf{F}_A$ in magnitude and is also directed vertically downward to the earth. On the surface of the earth, the tidal force is directed vertically downward at all places (forming a circle) where the sun is in the horizon at that moment.

The distance from the sun to the body at point Z (for which the sun is at the zenith) is smaller than to the center of the earth. Here the gravitational pull of the sun points exactly

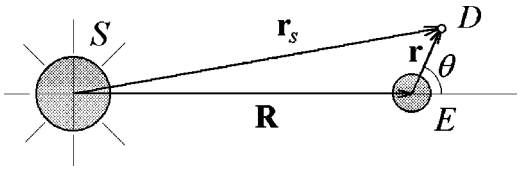

Fig. 2. For calculation of the tide-generating force at arbitrary point D.

opposite to and is somewhat greater than the force of inertia. Hence, the tidal force $\mathbf{F}_Z$ at this point is directed vertically upward, from the earth toward the sun. Its magnitude,

$$F_Z = G\frac{mM_{\text{sun}}}{(R-r)^2} - ma_0 = ma_0\left[\frac{R^2}{(R-r)^2} - 1\right] \approx ma_0\frac{2r}{R} = G\frac{mM_{\text{sun}}}{R^2}\frac{2r}{R}, \qquad (1)$$

is approximately twice the magnitude of the tidal forces at points A and B. Similarly, at the opposite point N (for which the sun is at its nadir) the force of inertia is greater than the gravitational pull of the sun, and so the tidal force $\mathbf{F}_N$ at point N is also directed vertically upward from the earth (and from the sun). In magnitude, $\mathbf{F}_N$ approximately equals $\mathbf{F}_Z$.

The expressions for the tidal forces, $\mathbf{F}_A = (GmM_{\text{sun}}/R^2) \times (r/R)$ and $\mathbf{F}_Z$ given by Eq. (1), are valid also for the tidal forces produced on the earth by the moon if we replace M_{sun} by the mass of the moon and R by the moon–earth distance. There is no intrinsic difference between the sun-induced and moon-induced tide-generating forces. In both cases, the only important factor is the acceleration of the earth under the gravitational pull of the celestial body that causes the tides on the earth, not the orbital velocities of both gravitationally coupled bodies (the earth and the sun, or the earth and the moon).

The tidal force experienced by any object is proportional to its distance r from the center of the earth and inversely proportional to the cube of the distance R to the celestial body that causes the force, and is proportional to the mass of the source body. As noted, lunar tide-generating forces on the earth are more than twice those of the sun (their ratio is approximately 2.2) because the moon is much closer to the earth.

III. TIDAL FORCES AT AN ARBITRARY POINT NEAR THE EARTH

The standard derivation of tidal forces uses the tide-generating potential (see, for example, Refs. 12 and 13) for which the mathematics is somewhat simpler. However, to emphasize the physics underlying the origin of tide-generating forces, we consider the vector addition of the relevant forces, just as in the elementary treatment of Sec. II. To obtain a general mathematical expression for the tide-generating force at an arbitrary point D over the earth (Fig. 2), we introduce the radius vector $\mathbf{r}$ of this point measured from the center of the earth, and also the vector $\mathbf{r}_s = \mathbf{R} + \mathbf{r}$ measured from the center of the sun, S, where $\mathbf{R}$ is the vector of the center of the earth from the center of the sun.

The tidal force $\mathbf{F}_{\text{tid}}$ experienced by a body of mass m at point D in the noninertial, nonrotating geocentric frame is the vector sum of its gravitational attraction to the sun, $\mathbf{F}_{\text{sun}} = -GmM_{\text{sun}}\mathbf{r}_s/r_s^3$, and the force of inertia, $\mathbf{F}_{\text{in}} = -m\mathbf{a}_0 = GmM_{\text{sun}}\mathbf{R}/R^3$:

$$\mathbf{F}_{\text{tid}} = \mathbf{F}_{\text{sun}} + \mathbf{F}_{\text{in}} = -GmM_{\text{sun}}\left(\frac{\mathbf{r}_s}{r_s^3} - \frac{\mathbf{R}}{R^3}\right). \qquad (2)$$

We express $\mathbf{r}_s$ in Eq. (2) as the vector sum $\mathbf{R}+\mathbf{r}$ and calculate the square of $\mathbf{r}_s$. We take into account that $r \ll R$ and write

$$r_s^2 = (\mathbf{R}+\mathbf{r})^2 = R^2 + 2(\mathbf{R}\cdot\mathbf{r}) + r^2 \approx R^2\left(1 + 2\frac{(\mathbf{R}\cdot\mathbf{r})}{R^2}\right). \qquad (3)$$

To find an approximate expression for $1/r_s^3$ in Eq. (2), we raise the right-hand part of Eq. (3) to the power $(-3/2)$. If we substitute the resulting value of $1/r_s^3$ into Eq. (2) for $\mathbf{F}_{\text{tid}}$, we obtain:

$$\mathbf{F}_{\text{tid}} \approx -G\frac{mM_{\text{sun}}}{R^3}\left[(\mathbf{R}+\mathbf{r})\left(1 - \frac{3(\mathbf{R}\cdot\mathbf{r})}{R^2}\right) - \mathbf{R}\right] \approx -G\frac{mM_{\text{sun}}}{R^3}\left[\mathbf{r} - 3\mathbf{R}\frac{(\mathbf{R}\cdot\mathbf{r})}{R^2}\right]. \qquad (4)$$

We note that the main contributions of $\mathbf{F}_{\text{sun}}$ and $\mathbf{F}_{\text{in}}$ to $\mathbf{F}_{\text{tid}}$, whose magnitudes are inversely proportional to R^2, cancel in Eq. (4). This cancellation corresponds to the aforementioned state of weightlessness that we experience on the spaceship Earth with respect to the sun's gravity. For points A and B in Fig. 1, $\mathbf{r}$ is perpendicular to $\mathbf{R}$, and hence the scalar product $(\mathbf{R}\cdot\mathbf{r})$ is zero. Therefore at these points the tidal force is directed opposite to $\mathbf{r}$ (that is, vertically downward), and its magnitude equals $GmM_{\text{sun}}(r/R^3)$. For points Z and N, the tidal force is directed along $\mathbf{r}$ (that is, vertically upward), and its magnitude $2GmM_{\text{sun}}(r/R^3)$ is two times greater than at points A and B. We see that at these four points, the general result given by Eq. (4) agrees with the simpler calculations of Sec. II.

IV. HORIZONTAL AND VERTICAL COMPONENTS OF THE TIDAL FORCE

The sun-induced tide-generating forces exerted on the earth have a quadrupole character: They stretch the earth along the sun–earth line, and squeeze the earth in the directions perpendicular to that line. Because of the axial symmetry with respect to the sun–earth line, the vertical and horizontal components of the tidal force depend only on the angle θ shown in Fig. 2 (and on the distance r from the center of the earth). The angle θ determines the position of the mass point m on or near the surface of the earth measured from this line.

Figure 3 shows how the tidal forces are directed at different points near the earth. Because of axial symmetry about the sun–earth line, Fig. 3 applies to any plane passing through the sun–earth line.

The horizontal (tangential to the surface) components of the tidal forces are much more influential on the ocean tides and on the orbits of earth satellites than are the vertical (radial) components, which only modify slightly the earth's gravitational force. For the horizontal component of the tidal

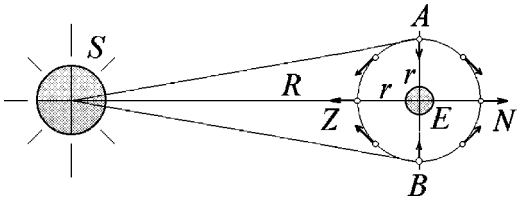

Fig. 3. Directions of the tidal forces at different equatorial points near the earth.

force at an arbitrary point D, whose geocentric position is determined by the two coordinates r and θ (in the plane shown in Fig. 2), Eq. (4) yields:

$$(F_{\text{tid}})_{\text{hor}} = -3G\frac{mM_{\text{sun}}}{R^3} r\cos\theta\sin\theta$$

$$= -3F_{\text{sun}}\frac{r}{R}\cos\theta\sin\theta = -\frac{3}{2}F_{\text{sun}}\frac{r}{R}\sin 2\theta, \quad (5)$$

where $F_{\text{sun}} = GmM_{\text{sun}}/R^2$ is the gravitational pull of the sun on the body. The horizontal component of the tidal force is zero at points A and B and at all other points of the plane orthogonal to the line sun–earth (for which $\theta = 90°$), as well as at points N and Z (for which $\theta = 0°$ and $\theta = 180°$). The horizontal component of the tidal force has its maximum value $(3/2)(r/R)F_{\text{sun}} = (3/2)(r/R)GmM_{\text{sun}}/R^2$ at all points on the earth for which $\theta = 45°$ and $\theta = 135°$. This maximal horizontal component of the solar tide-generating force causes a deviation of the plumb line from the direction of the earth's own gravity only by $0.008''$.

If we take the scalar product of the right-hand side of Eq. (4) for $\mathbf{F}_{\text{tid}}$ with the unit vector $\mathbf{r}/r$, we obtain the dependence of the vertical component $(F_{\text{tid}})_{\text{vert}}$ of the tidal force on the angle θ between $\mathbf{R}$ and $\mathbf{r}$:

$$(F_{\text{tid}})_{\text{vert}} = G\frac{mM_{\text{sun}}}{R^3} r(3\cos^2\theta - 1)$$

$$= \frac{3}{2}G\frac{mM_{\text{sun}}}{R^2}\frac{r}{R}\left(\cos 2\theta + \frac{1}{3}\right). \quad (6)$$

The last term on the right-hand side of Eq. (6) is independent of θ and is thus independent of time on the spinning earth. It can therefore be dropped as far as the tides are concerned. This term in the vertical component of the tidal force is the same everywhere on the earth (for a given value of r) and adds only a tiny constant value to the vertical force of the earth's gravity (about ten million times smaller than mg). Thus, the vertical and horizontal components of the tidal force exerted on a body of mass m located at a position determined by angle θ and radius r are given by:

$$F_{\text{vert}} = (3/2)(r/R)F_{\text{sun}}\cos 2\theta,$$
$$F_{\text{hor}} = -(3/2)(r/R)F_{\text{sun}}\sin 2\theta, \quad (7)$$

where F_{sun} is the total gravitational pull of the sun experienced by the body anywhere on the earth. This representation of the tide-generating force is especially convenient because Eq. (7) defines a tidal force vector whose magnitude $(3/2)\times(r/R)F_{\text{sun}} = (3/2)GmM_{\text{sun}}r/R^3$ is independent of the angle θ. The tidal forces at all points that lie at a given distance r from the earth's center are equal in magnitude and differ only in direction.

Equations (5)–(7) also are valid for the tidal forces produced by the moon, provided we replace the mass of the sun M_{sun} by the mass of the moon M_{moon} and the sun–earth distance R by the moon–earth distance. In this case the angle θ in Eq. (7) determines the position of the body relative to the moon–earth line.

The tide-generating force of the moon, $F_{\text{tidal}} = (3/2)GmM_{\text{moon}}r_0/R^3$, experienced by a body of mass m on the surface of the earth (r_0 is the earth's radius) is very small compared to its weight—the earth's force of gravity $F_{\text{grav}} = mg = GmM_{\text{earth}}/r_0^2$. If we let the ratio $M_{\text{moon}}/M_{\text{earth}} = 1/81$ and the mean earth–moon distance $R = 60r_0$ (actually this distance varies between $57r_0$ and $63.7r_0$ because of the elliptical shape of the moon's orbit), we obtain

$$F_{\text{tidal}}/F_{\text{grav}} = (3/2)(M_{\text{moon}}/M_{\text{earth}})(r_0/R)^3 \approx 8.6\times 10^{-8}. \quad (8)$$

Although the maximal lunar tidal force on the surface of the earth is only about 10^{-7} of the earth's gravitational force, its effect on the ocean water can be considerable because of its horizontal component, which is orthogonal to the earth's gravitational field and varies with time periodically because of the earth's axial rotation. The horizontal component shifts the ocean water around the globe.

V. THE STATIC DISTORTION OF THE WATER SURFACE

To estimate the static (equilibrium) distortion of the ocean's surface due to the tidal forces, we can use the hypothetical situation of a nonrotating planet on which the tide-generating forces are nearly time-independent. From the symmetry of tidal forces, Eq. (7), we can assume that the distorted surface has an ellipsoidal shape given by

$$r(\theta) = r_0 + a\cos 2\theta, \quad (9)$$

where $2a \ll r_0$ is the difference in the static maximal and minimal levels at points Z and A (see Fig. 3). Hence we can write for the small inclination α of the water surface with respect to the horizon:

$$\alpha = \frac{1}{r}\frac{dr(\theta)}{d\theta} \approx -\frac{2a}{r_0}\sin 2\theta. \quad (10)$$

We see that the water surface is horizontal ($\alpha = 0$) at $\theta = 0$ and $\theta = 90°$ (points Z and A). The angle α is maximum and equals $2a/r_0$ at $\theta = 45°$ and at $\theta = 135°$, where the tidal force is directed horizontally. In equilibrium the distorted water surface is orthogonal to the plumb line. The plumb line shows the direction of the vector sum of the earth's gravity and the tidal force. A small departure of the plumb line from the direction of the earth's gravity is caused by the horizontal component of the tidal force. Therefore, the angle α equals the ratio of the horizontal tidal force F_{hor} to the force of the earth's gravity $F_{\text{grav}} = mg$. If we equate $\alpha = 2a/r_0$ at $\theta = 45°$ to $F_{\text{hor}}/F_{\text{grav}}$ and take into account that for sun-induced tides, $F_{\text{hor}}/mg = (3/2)(M_{\text{sun}}/M_{\text{earth}})(r_0^3/R^3)$, we find for the maximal static level difference $2a$ at points Z and A:

$$2a = (3/2)r_0(M_{\text{sun}}/M_{\text{earth}})(r_0^3/R^3). \quad (11)$$

Equation (11) yields $2a = 0.24$ m. A similar expression is valid for the static distortion of the ocean surface due to the lunar tidal force, and yields $2a = 0.54$ m for the moon-induced static distortion. In Sec. VII the equation for this static distortion is also derived from the tide-generating potential.

VI. TIDAL FORCES ON THE ROTATING EARTH

In the above we have used a revolving but nonrotating geocentric reference frame. The origin of this frame moves in a circle around the sun–earth (moon–earth) center of mass, but the frame itself does not rotate because the directions of its axes are fixed relative to the distant stars. That is, the frame moves translationally in a circle. This reference frame is convenient for the analysis of a motion of an artificial satellite. If we ignore the perturbations caused by tidal forces, the earth satellite traces out a closed elliptical orbit relative to this reference frame.

To introduce tidal forces on the rotating earth, we must use a true geocentric frame of reference that takes part in the daily rotation of the earth. This frame is noninertial, and hence we should be concerned with the acceleration of its different points. We can consider the motion of the earth (and of the geocentric reference frame) as consisting of two components. The first is the component considered above, namely translational motion (revolution without rotation) about the sun–earth (moon–earth) center of mass. The second component is a uniform daily rotation (spin) of the earth about an axis passing through the center of the earth.

Both these motions of the earth are important in the problem of tides, but the roles they play are quite different. The acceleration $\mathbf{a}_0$ related to the translational motion is responsible for the origin of the uniform pseudoforce of inertia $\mathbf{F}_{\text{in}} = -m\mathbf{a}_0$, whose action on a body on the earth, combined with the nonuniform gravitational pull of the sun (moon), is described by the tidal force $\mathbf{F}_{\text{tid}}$ considered previously. We note again that only the acceleration $\mathbf{a}_0$ of this translational motion is important, not the orbital velocity of the earth.[14] To avoid confusion often encountered in the literature (see, for example, Ref. 15), we must be careful with definitions. In discussing tides, we should be concerned only with those gravitational and inertial forces that depend on the apparent position of the celestial body that produces the tide. The axial rotation of the earth is related to the centripetal acceleration and gives rise to centrifugal forces that increase in proportion to the distance from the earth's axis. The centrifugal force of the earth's daily rotation generally is much greater in magnitude than tidal forces. Because of the centrifugal forces, the equilibrium shape of the earth differs slightly from an ideal sphere—it is approximately an ellipsoid of rotation whose equatorial diameter is a bit greater than the polar diameter (see, for example, Ref. 13). The centrifugal effect of the earth's daily rotation causes an equatorial bulge, which is the principal departure of the earth from its spherical shape.[16]

But we are not concerned here with this constant distortion of the earth because this distortion is independent of the apparent position of the celestial body that produces the tides. Therefore, the centripetal acceleration of the axial rotation adds nothing to tidal forces. However, the daily rotation of the earth makes tidal forces time-dependent because the pattern of tidal forces on the earth is coupled to the apparent positions of the sun and moon. A dynamical response of the oceanic waters on the spinning earth to these time-dependent forces is the essence of the phenomenon of tides.

Thus, in the problem of tides, expressions for the tide-generating forces F_{hor} and F_{vert} in Eq. (7) are applicable also to the true geocentric frame of reference, which takes part in the daily axial rotation of the earth. The system of tidal forces shown in Fig. 3, being coupled to the apparent position of the sun (moon), rotates rigidly together with the earth–sun (earth–moon) line. For simplicity, we shall consider the case in which the source celestial body (the sun or moon) occurs in the equatorial plane of the earth. Although the system of tidal forces rotates as a whole with the angular velocity Ω of the earth's axial rotation, that is, with a period of $2\pi/\Omega$, the true period of variation of the tidal forces on the earth equals half this value ($T = \pi/\Omega$) because of the quadrupole symmetry of the system of forces (the *semidiurnal tide*). For the sun-induced tidal forces the period equals 12 h. For the moon-induced tidal forces the period is 12 h 25 min—the difference between the periods is due to the orbital motion of the moon. If we fix a point on the equator of the earth, the local tidal force vector executes a uniform rotation in the vertical plane, making two complete revolutions during a day. The simulation clearly shows how the daily rotation of the whole system of tidal forces produces this doubly fast uniform rotation of the tidal force at a given equatorial point, as seen by an observer on the spinning earth.[10] Because of this periodic dependence on time, the tidal forces, in spite of their small magnitude compared even to the centrifugal force of inertia, produce the oceanic tides.

To find analytical expressions for the time dependence of the tidal forces at a given point in the equatorial plane of the spinning earth, we substitute $\theta = \Omega t$ in Eq. (7). This substitution yields the following expressions for the point of the equator at which the sun culminates (passes through its zenith) at $t = 0$:

$$F_{\text{vert}}(t) = Ar\cos 2\Omega t, \qquad F_{\text{hor}}(t) = -Ar\sin 2\Omega t, \tag{12}$$

where $A = (3/2)F_{\text{sun}}/R = (3/2)GmM_{\text{sun}}/R^3$. At any other equatorial point of the earth, the tidal force vector also rotates in the vertical plane with angular velocity 2Ω. That is, all the vectors at different points rotate synchronously but with different phases.

VII. THE POTENTIAL FUNCTION FOR TIDAL FORCES

An approach often used in deriving an expression for the tidal force is to begin with the potential energy of a body under the influence of tide-generating forces. This approach is simpler than that presented above. However, we have chosen the above approach because it does not obscure the underlying physics and consequently may be considered advantageous to physics instructors. Nevertheless, for completeness, we introduce here the potential function, $U_{\text{tides}}(r,\theta)$, and show how it can be used in calculating the equilibrium shape of the surface of the ocean and the static distortion of the water under tidal forces.

The components of the force that lie in the equatorial plane are given in Eq. (7) and are the negative gradients of the potential function $U_{\text{tides}}(r,\theta)$:

$$F_{\text{vert}} = Ar\cos 2\theta = -\partial U_{\text{tides}}(r,\theta)/\partial r,$$
$$F_{\text{hor}} = -Ar\sin 2\theta = -(1/r)\partial U_{\text{tides}}(r,\theta)/\partial\theta. \tag{13}$$

Therefore, the potential function for the tidal forces can be written as:

$$U_{\text{tides}}(r,\theta) = -(1/2)Ar^2\cos 2\theta$$
$$= -(3/4)(GmM_{\text{sun}}/R^3)r^2\cos 2\theta. \tag{14}$$

The restoring forces that limit the tidal distortion of the water's surface are due to the earth's gravity. If the earth were not rotating relative to the earth–sun line, the static distortion of the water surface covering the globe would be the surface of equal total potential:

$$U(r,\theta) = U_0(r) + U_{\text{tides}}(r,\theta) = \text{const}, \tag{15}$$

where $U_0(r) = mgr$ is the spherically symmetric potential function of the earth's gravity which yields the radial component of the earth's gravitational force $-dU_0(r)/dr = -mg$. Thus,

$$U(r,\theta) = mgr - (1/2)Ar^2\cos 2\theta. \tag{16}$$

In particular, at points Z and A (see Fig. 3) of the water surface, the values of the total potential function, Eq. (16), are equal: $U(r_Z,\pi) = U(r_A,\pi/2)$, from whence we obtain

$$mgr_Z - (1/2)Ar_Z^2 = mgr_A + (1/2)Ar_A^2,$$
$$mg(r_Z - r_A) = (1/2)A(r_A^2 + r_Z^2). \tag{17}$$

We can use this condition to determine the static equilibrium distortion under the tidal forces of the otherwise spherical ocean surface. Let the radii of the distorted water surface at points Z and A be $r_Z = r_0 + a$ and $r_A = r_0 - a$, respectively, where r_0 is the radius of the undistorted surface. Then $2a$ is the static level difference at points Z and A in which the level is maximum and minimum, respectively. Thus, from Eq. (17) we have $2mga = (1/2)A(r_Z^2 + r_A^2) \approx Ar_0^2$, and for $2a$ we obtain:

$$2a = Ar_0^2/(2mg) = (3/2)r_0(F_{\text{sun}}/mg)(r_0/R). \tag{18}$$

We note that $F_{\text{sun}}/mg = (M_{\text{sun}}/M_{\text{earth}})(r_0^2/R^2)$, so that the static distortion of the ocean surface under the sun-induced tidal forces can also be expressed as:

$$2a = (3/2)r_0(M_{\text{sun}}/M_{\text{earth}})(r_0^3/R^3). \tag{19}$$

This expression is the same as Eq. (11) derived by requiring that in equilibrium the surface of the ocean be orthogonal to the vector sum of the earth's gravitational force and the tidal force.

VIII. THE NATURAL WAVE AND THE DRIVING TIDAL FORCES

Most authors oversimplify the problem of tides and consider (after Newton and Bernoulli) only the so-called *static* (or *equilibrium*) theory of tides, which treats the ocean surface as a liquid ellipsoid stretched *along* the earth–moon (earth–sun) line, as if this surface were always in equilibrium under the earth's force of gravity and tidal forces produced by the moon (sun). In this approach, the tidal bulges are aligned with the earth–moon (or earth–sun) axis. Therefore on the spinning earth the moments of high water at a given location should coincide with the upper and lower culminations of the moon (sun), that is, when the moon (sun) passes through its zenith and nadir. Observations do not agree with this prediction. Instead, almost the opposite is usually observed: the moments of low tide occur approximately at the culminations of the moon.

A complete theory of the tides should take into account the dynamical response of the ocean to the time-dependent generating forces. The dynamical theory of tides (first suggested by Laplace and developed by Airy) treats the tides as a steady-state forced motion (under varying tidal forces) of a dynamical system (the ocean).[17] Such a theory predicts a resonant growth of the steady-state amplitude in cases when the driving period approaches the period of natural oscillations.

To avoid the complications related to the three-dimensional character of the problem and to explain the physical aspect of the dynamical theory using the simplest possible model, we imagine, following Airy, water in a wide canal of uniform depth engirdling the entire earth along the equator. Imagine the water surface in this canal being distorted statically under the tide-generating forces so that two bulges form on opposites sides of the earth, changing the shape of the surface from circular to elliptical. If the forces maintaining this shape suddenly vanish, the earth's gravity would make the distorted surface restore its equilibrium, circular shape. The water would start to flow and the bulges disappear so that after a time, namely a quarter period, the water surface would become circular. But because the water continues to move, after another quarter period the bulges reappear in new positions showing an elliptical distortion of the surface along the line perpendicular to the line of the original distortion. Then the motion repeats itself in reverse. This motion of water in the circular canal is a gravitational standing surface wave whose wavelength equals half-circumference of the globe. Such a mode of oscillation is characterized by a certain natural period.

The superposition of two such standing waves whose phases differ by $\pi/2$ and whose elliptical axes are separated by 45° produces a circulating (traveling) wave of constant elliptical shape and a wavelength equal to half of the earth's circumference. The two opposite bulges in the water surface travel with this wave around the globe preserving their height and shape.[10]

An essential point in explaining the steady-state phase shift between the moments of high tide and the culmination of the moon (sun) is the relation between the natural period T_0 of this circulating wave and the period T of the tide-generating driving forces. It is possible to estimate T_0 as the time taken by the circulating surface wave to travel along half the globe. In the limiting case of very long waves on the surface of shallow water ($\lambda \gg h$) the speed of wave is determined by the earth's gravity g and depth h, and is independent of λ. From hydrodynamics we know that this speed equals $(gh)^{1/2}$ (see, for example, Ref. 18, p. 405). We assume that the mean value h of the ocean depth is 3.5 km. During a period T_0, the wave travels half the circumference of the globe πr_0, and hence $T_0 = \pi r_0/(gh)^{1/2} \approx 30$ h. Thus, the approximately 12-h driving external period T is less than the natural period T_0 of the free oscillation.

We emphasize that it is the shape of the surface (the wave) that circulates around the globe, *not the water itself*. Relative to the earth, points on the surface of the ocean execute os-

cillatory motions in closed paths that are considerably stretched horizontally. On the average, the water is stationary in the geocentric frame.

To obtain the dynamical picture of tides on the rotating earth, we should use the reference frame that rotates with the earth. Relative to this frame, the quadrupole system of tide-generating forces, being coupled to the position of the sun (moon), rotates as a whole while the sun (moon) travels along its apparent daily path around the earth. This rotation of the forces occurs at an angular velocity Ω, the angular velocity of the earth's daily rotation (or the difference between Ω and the angular velocity of the moon in its orbit for moon-induced tides). Such a uniform rigid rotation of the system of mutually fixed vectors can be represented as a superposition of two oscillating quadrupole systems of forces (with a frequency $\omega=2\Omega$) that do not rotate and whose axes make an angle of 45° to one another. At each point one of these forces oscillates along the radial (vertical) direction, while the other force—along the tangential (horizontal) direction. The oscillations of these orthogonal components occur a quarter period out of phase. At any given point in the equatorial plane, the vector sum of these mutually orthogonal oscillating forces produces a force of constant magnitude whose direction rotates uniformly following the apparent motion of the sun (moon), but with angular velocity $\omega=2\Omega$.[10] For different points on the earth, the phases of these rotating vectors differ.

IX. THE TIDES AS FORCED OSCILLATIONS OF THE OCEAN

What is really of interest is the steady-state forced oscillation of the ocean surface due to the time-dependent tidal forces. Each of the two oscillating systems of forces described above excites a mode of forced oscillation of the water in the equatorial canal, specifically the mode of the same symmetry as is characteristic of the corresponding system of driving forces. These modes have elliptical shapes, much like the natural oscillations considered above, namely, the elliptical standing waves whose axes make an angle of 45° with one another. Nevertheless, we can consider these modes to be orthogonal in the sense that their spatial forms are described by eigenfunctions forming an orthogonal basis in the function space. The two forced oscillations in this linear system, each excited by one system of oscillating driving tidal forces, are independent of one another, and the resulting forced motion is a superposition of these forced oscillations.

Any steady-state forced oscillation occurs exactly with the period of the driving force. The amplitude and phase lag of the oscillation depend on the amplitude of the driving force, on the damping factor, and, more importantly, on the relation between the driving and natural periods. The two systems of oscillating driving tidal forces are characterized by equal amplitudes and frequencies. Also the natural frequencies and damping factors of both excited modes are equal. Hence both excited modes also have equal amplitudes and equal phase delays behind the corresponding driving forces. The superposition of these modes produces a forced circulating (traveling) elliptical wave that has the same phase relation with the rotating driving forces as is characteristic of forced oscillations in general.

If we *ignore friction* (dissipation of mechanical energy in the excited wave motion), the forced motion occurs exactly *in phase* with the driving force, provided the driving period is longer than the natural period. Otherwise the forced motion occurs in the *opposite phase* with respect to the driving force. For the simplified model of tides in the equatorial canal of uniform depth (and also for an earth covered everywhere by an ocean of uniform depth), the natural period of free oscillation is longer than the 12-h driving period. Thus the dynamical theory predicts in this case a stationary circulating elliptically shaped wave whose axis (the line of tidal bulges) is *perpendicular* to the earth–sun (earth–moon) line.

On the other hand, the natural period of an elastic wave in the crust of the earth is shorter than the 12-h period of the tidal forces. Hence, in the frictionless model, bulges in the earth's crust are oriented *along* the earth–sun (earth–moon) line. Observations show that the solid body of the earth actually experiences twice-daily tides with maximum amplitude of about 30 cm whose bulges lag approximately 3° behind the earth–moon line.[17]

X. MATHEMATICAL DESCRIPTION OF THE FORCED OSCILLATIONS

Each of the partial forced oscillations can be described by a differential equation of a linear oscillator. Let $q_1(t)$ be the normal coordinate describing the first forced oscillation whose elliptical shape is characterized by a major axis oriented along the earth–sun line (and in the perpendicular direction after a half period), and let $q_2(t)$ be the normal coordinate describing the second oscillation with the axis inclined 45° to the earth–sun line. A disturbance of the water surface caused by the first oscillation can be described by $\Delta r_1(\theta,t)=q_1(t)\cos(2\theta)$, which gives the small vertical displacement of the surface at an arbitrary point (r_0,θ) of the equator. Similarly, the second oscillation causes a distortion of the surface described by $\Delta r_2(\theta,t)=q_2(t)\sin(2\theta)$. The forced oscillations experienced by the normal coordinates $q_1(t)$ and $q_2(t)$ are periodic (steady-state) partial solutions of the two differential equations:

$$\begin{aligned} &\ddot{q}_1+2\gamma\dot{q}_1+\omega_0^2 q_1=\omega_0^2 a\cos\omega t,\\ &\ddot{q}_2+2\gamma\dot{q}_2+\omega_0^2 q_2=\omega_0^2 a\sin\omega t. \end{aligned} \tag{20}$$

Here ω_0 is the natural frequency of the corresponding mode ($\omega_0=2\pi/T_0=2(gh)^{1/2}/r_0$), γ is the damping constant, $\omega=2\Omega$ is the driving frequency, and a is the magnitude of the equilibrium distortion of the ocean surface under the static system of tidal forces (that is, the distortion for the planet whose axial rotation is synchronized with its orbital revolution). The theoretical value of a is given by Eq. (11) or (18). Although the values of ω and a are fairly well known, the situation is quite different regarding the values of ω_0 and γ.

In the limiting case of extremely slow rotation of the earth, the steady-state solution of Eq. (20) is $q_1(t)=a\cos\omega t$, $q_2(t)=a\sin\omega t$. This solution describes the quasistatic elliptical distortion whose axis follows adiabatically the slowly rotating earth–sun (earth–moon) line. The major axis of the ellipse at any moment is oriented along this line. The displacement of the water level from its mean position in the equatorial plane in this limiting case is given by

$$\begin{aligned}\Delta r(\theta,t) &= \Delta r_1(\theta,t) + \Delta r_2(\theta,t)\\ &= q_1(t)\cos 2\theta + q_2(t)\sin 2\theta\\ &= a(\cos 2\Omega t \cos 2\theta + \sin 2\Omega t \sin 2\theta)\\ &= a\cos 2(\Omega t - \theta). \end{aligned} \tag{21}$$

To find the distortion of the water surface for an arbitrary value of ω, we can use the relevant well-known steady-state solution to Eq. (20) for the normal coordinates $q_1(t)$ and $q_2(t)$:

$$q_1(t) = q_0\cos(\omega t - \delta), \quad q_2(t) = q_0\sin(\omega t - \delta), \tag{22}$$

where their common amplitude q_0 and phase lag δ are given by

$$\begin{aligned} q_0 &= \frac{\omega_0^2 a}{\sqrt{(\omega_0^2-\omega^2)^2 + 4\gamma^2\omega^2}},\\ \tan\delta &= \frac{2\gamma\omega}{\omega_0^2-\omega^2}. \end{aligned} \tag{23}$$

(See, for example, Ref. 18, pp. 372–373.) Therefore the resulting distortion of the water surface under the tidal forces is given by

$$\begin{aligned}\Delta r(\theta,t) &= \Delta r_1(\theta,t) + \Delta r_2(\theta,t)\\ &= q_1(t)\cos 2\theta + q_2(t)\sin 2\theta\\ &= q_0[\cos(2\Omega t - \delta)\cos 2\theta + \sin(2\Omega t - \delta)\sin 2\theta]\\ &= q_0\cos 2(\Omega t - \delta/2 - \theta). \end{aligned} \tag{24}$$

We see from Eq. (24) that at any time t the maximum (high water) of the tidal wave circulating around the earth is located at the position defined by the angle $\theta_{\max} = \Omega t - \delta/2$. That is, the position of the maximum lags behind the sun (moon) by the angle $\delta/2$. If $\gamma \ll \omega$, it follows from Eq. (23) that this retarding angle is almost zero if $\omega < \omega_0$. In other words, the marine tide would be nearly the equilibrium tide with the high-water time coinciding with culminations of the sun (moon) if the natural period of the circulating wave were less than the 12-h driving period (that is, if $T_0 < T$). However, for our model of the ocean, we estimate the natural period to be close to 30 h. Therefore the situation corresponds to $\omega > \omega_0$, when the steady-state forced oscillations occur nearly in the opposite phase relative to the driving force. In this case the tide should be inverted with respect to the equilibrium one. The retarding angle $\delta/2$ approaches $\pi/2$ according to Eq. (23), which means that for a given equatorial point, the high water occurs when the sun (moon) is almost at the horizon (rather than at zenith or nadir).

At any given place on the equator, it follows from Eq. (24) that the water level (above the average value) varies with t according to $z(t) = q_0\cos(2\Omega t - \delta)$, where $t = 0$ corresponds to the culmination of the sun (moon) at the place in question. We can expect that for the model of a water canal of uniform depth, the value of q_0 given by Eq. (23) is more or less reliable because hydrodynamics allows us to estimate the natural frequency $\omega_0 = 2\pi/T_0 = 2(gh)^{1/2}/r_0$ by using the known speed $\nu = (gh)^{1/2}$ of very long gravitational waves. However, considerable uncertainty is related to the damping factor γ. If we assume that the damping is small ($\gamma \ll \omega_0$), we can conclude that the orientation of the tidal bulges deviates only slightly from the line perpendicular to the sun–earth (moon–earth) line, but the particular value of this deviation remains indefinite.

In the above discussion, we considered only the steady-state oscillation of the ocean surface (the stationary wave), assuming that the transient is already over. For this steady motion to establish itself, some friction (even if very small) is necessary. In the problem under consideration, we are concerned with the water motion caused solely by the eternal tidal forces, and therefore we have had centuries and even millennia to wait for the fading away of the transient. Therefore our use of the steady-state solution is appropriate for tides. We also emphasize that in the dynamical theory of tides, the driving tide-generating forces are perfectly well known, so that most uncertainties originate primarily from a very poor correspondence between the simple model of the dynamical system and the real oceans of the earth.

XI. REAL-WORLD COMPLICATIONS

The pattern of tide-generating forces is coupled to the position of the moon (and the sun) with respect to the earth. For any place on the earth's surface, the relative position of the moon has an average periodicity of 24 h 50 min. The lunar tide-generating force experienced at any location has the same periodicity. When the moon is in the plane of the equator, the force runs through two identical cycles within this time interval because of the quadrupole symmetry of the global pattern of tidal forces. Consequently, the tidal period is 12 h 25 min in this case (the period of the semidiurnal lunar tide). However, the lunar orbit doesn't lie in the plane of the equator, and the moon is alternately to the north and to the south of the equator. The daily rotation of the earth about an axis inclined to the lunar orbital plane introduces an asymmetry in the tides. This asymmetry is apparent as an inequality of the two successive cycles within 24 h 50 min.

Similarly, the sun causes a semidiurnal solar tide with a 12-h period, and a diurnal solar tide with a 24-h period. In a complete description of the local variations of the tidal forces, still other partial tides play a role because of further inequalities in the orbital motions of the moon and the earth. In particular, the elliptical shape of the moon's orbit produces a 40% difference between the lunar tidal forces at the perigee and apogee of the orbit. Also the inclination of the moon's orbit varies periodically in the interval 18.3°–28.6°, causing a partial tide with a period of 18.6 yr. The interference of the sun-induced tidal forces with the moon-induced tidal forces (the lunar forces are about 2.2 times as strong) causes the regular variation of the tidal range between *spring tide*, when the range has its maximum (occurring at a new moon and at a full moon, when the sun and moon are in the same or in the opposite directions), and *neap tide*, when the range has its minimum (which occurs at intermediate phases of the moon). The amplitude of a spring tide may be 2.7 times the amplitude of a neap tide.

Because the earth is not surrounded by an uninterrupted water envelope of equal depth, but rather has a very irregular geographic alternation of land and seas with complex floor geometry, the actual response of the oceans and seas to the tidal forces is extremely complex. In enclosures formed by gulfs and bays, the local tide is generated by an interaction with the tides of the adjacent open ocean. Such a tide often takes the form of a running tidal wave that circulates within the confines of the enclosure. In some nearly enclosed seas,

such as the Mediterranean, Black, and Baltic seas, a steady-state oscillation in the form of a standing wave, or tidal seiche, may be generated by the tidal forces. In these seas, the tidal range of sea level is only on the order of centimeters. In the open ocean, it generally is on the order of decimeters.

In bays and adjacent seas, however, the tidal range may be much greater because the shape of a bay or adjacent sea may favor the enhancement of the tide inside. In particular, there may be a resonance response of the basin concerned with the tide. Tides are most easily observed along seacoasts, where the amplitudes are exaggerated. When tidal currents run into the shallow waters of the continental shelf, their rate of advance is reduced, the energy accumulates in a smaller volume, and the rise and fall are amplified. The details of tidal motions in coastal waters, particularly in channels, gulfs, and estuaries, depend on the details of coastal geometry and water-depth variation over a complex sea floor. Tidal amplitudes and phase lags, the contrast between spring and neap tides, and the variation of times of high and low tide all change widely from place to place.

For the aforementioned reasons, a purely theoretical calculation of the times and heights of tides at a particular location is practically impossible. Nevertheless, for a given place on a coast, the tides can be quite successfully predicted on the basis of accumulated long-term observations of the tides at the place concerned. The analysis of the observations relies on the fact that any tidal pattern in time is a superposition of variations associated with periodicities in the motions of the moon and the sun relative to the earth. The periods involved are the same everywhere on the earth, but the relative amplitudes and phases of their contributions are highly variable from one place to another. Observations over a sufficient time make it possible to calculate which contributions are significant at a particular location and, thus, to forecast tidal times and heights. It is common that 40 harmonic components may be significant for practical calculations at one location.[17]

XII. THE EVOLUTION OF ORBITAL MOTIONS AND SPINS OF CELESTIAL BODIES INDUCED BY TIDAL FORCES

When the forced motion occurs exactly in the same or opposite phase with respect to the driving force, no energy exchange occurs on average between the external source and the oscillatory system. To explain the secular variation (the retardation) of the earth's axial rotation under the tidal forces, we have to take friction into account.

One may wonder why the dissipation of mechanical energy in the tides has a scale that seems very modest. The point is that only the *wave* circulates around the globe, not the water itself. The phase lag δ of the steady-state forced oscillation behind the periodic driving force is determined by Eq. (23). For the mode of oscillations in which we are interested, this phase-frequency characteristic is almost a step function (zero for $\omega < \omega_0$, that is, for $T > T_0$, and $-\pi$ otherwise). Only near resonance ($\omega \approx \omega_0$) is this step slightly smoothed over. Therefore the displacement of the tidal water bulges from the line perpendicular to the sun–earth (moon–earth) axis is very small.

However, this displacement, which destroys the symmetry of the system (Fig. 4), is absolutely necessary in principle in order that the driving tidal forces be capable of maintaining the circulating tidal wave (that is, of preventing it from damping out). If the earth is taken as the reference frame, we can see that by virtue of this phase shift and the corresponding displacement of bulges, the tidal forces exert a retarding torque relative to the earth's axis and thus do nonzero net work on the system. This work compensates for the frictional losses experienced by the tidal traveling surface wave and measures the gradual reduction of the mechanical energy of the system. The energy is provided by the axial rotation (spin) of the earth. Hence the spin secularly slows down and the angular momentum of the axial rotation diminishes.

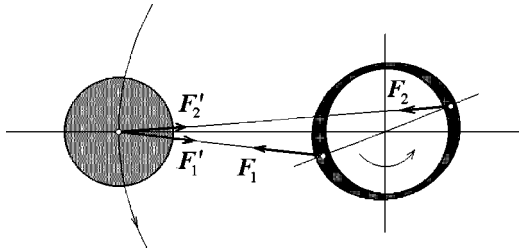

Fig. 4. Gravitational interaction between the moon and the tidal bulges.

Looking at the whole system from the inertial reference frame, we should remember that the sun (moon) interacts with the earth only by its central gravitational force. If the bulges were oriented exactly along or perpendicularly to the sun earth (moon–earth) axis, this gravitational force would not exert a torque on the earth. If we consider the gravitational forces $\mathbf{F}_1$ and $\mathbf{F}_2$ (Fig. 4) exerted on the bulges, we conclude that the retarding torque about the earth's axis, which slows down the axial rotation, is due to the above-mentioned displacement of the bulges which destroys the symmetry of the system with respect to the earth–sun (earth–moon) line.

However, the total torque of the central gravitational field of the sun (moon) exerted on the earth and the bulges of its liquid shell, measured relative to the sun (or to the moon for moon-induced tides), is *zero*. Hence the total angular momentum of the system is conserved, as it should be in any closed system. The diminishing of the earth's spin due to tidal friction means that the orbital momentum of the system slowly increases during the tidal evolution. The earth's orbit gradually expands. The lack of symmetry (produced by tidal friction) does not influence the conservation of total angular momentum, although it causes a slow secular redistribution of the angular momentum between the spin and the orbital motion. As the orbit expands, the mechanical energy of the orbital motion also increases. This additional mechanical energy, as well as the dissipated energy, is borrowed from the energy of axial rotation.[19]

This conclusion about expanding the moon's orbit, derived from the conservation of angular momentum, is often encountered in the literature (see, for example, Ref. 20). Although quite convincing, it nevertheless leaves the actual mechanism unexplained. To understand the physical reason for this phenomenon, it helps to take the forces into account. If we consider the properties of the gravitational forces $\mathbf{F}_1'$ and $\mathbf{F}_2'$ (see Fig. 4) that are exerted on the moon by the earth's tidal bulges and their influence on the orbital motion, we draw attention to a subtle peculiarity that deserves discussion. While the orbit expands, the orbital velocity of the moon diminishes. However, from the asymmetry in the configuration that is responsible for the evolution, we can conclude that the resultant gravitational force exerted on the

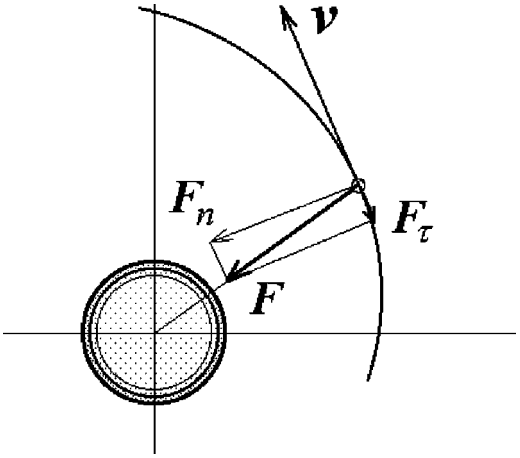

Fig. 5. The main (central) gravitational pull of the earth exerted on the moon.

moon by the tidal bulges is directed *forward*, in the direction of the orbital motion. How can this accelerating force slow down the orbital motion? All authors who write about tidal evolution leave this question unanswered.

This situation is similar to the widely known paradox of an earth satellite in a circular orbit that gradually descends in the rarified upper atmosphere: Intuitively we expect that the weak atmospheric drag should slow down the satellite, but instead, the satellite gains speed as its orbit gradually decreases. Because of air resistance, the satellite is accelerated in the direction of its motion, as if the retarding force of air resistance were pushing the satellite forward. An explanation of this so-called aerodynamical paradox of the satellite can be found in Ref. 21.

To understand the slowing down of the moon during tidal evolution, we must take into account that the moon gradually spirals away from the earth and its orbit spreads out, so that the actual motion of the moon occurs along an expanding spiral. A portion of this trajectory (with a strongly exaggerated expansion) is shown schematically in Fig. 5. Because of this expansion, the perpendicular to the trajectory is directed not to the center of the earth but rather slightly in front of the center. Therefore the main gravitational pull **F** exerted on the moon by the earth has a retarding tangential component $\mathbf{F}_\tau$ directed back along the trajectory. This component is greater in magnitude than the forward-directed tangential component of $\mathbf{F}_1'$ and $\mathbf{F}_2'$ (see Fig. 4) that are exerted on the moon by the tidal bulges (this component is not shown in Fig. 5). Hence the total tangential acceleration of the moon is directed against the velocity.

Generally, in order to explain tidal evolution, that is, the reduction of spin and the secular variation of the orbits of gravitationally coupled celestial bodies, it is necessary to take into account both the dynamic distortion of the spherical shape of the body (and of its liquid shell, if any) under the tidal forces, and the additional displacement of the bulges caused by tidal friction. The nonuniform gravitational field of one body in an orbit about another distorts the shape of the second. The dissipation of energy stored in the resultant tidal distortions leads to a coupling that causes secular changes in the orbit and in the spins of both bodies. Retardation of the axial rotation and evolution of the orbit will continue until the axial rotation is synchronous with the mean orbital revolution.

This effect is vital to an understanding of the history of the earth and moon. That the moon always keeps the same face turned toward the earth is attributed to the past effects of tidal friction in the moon. The dissipation of tidal energy on the earth results in a slowing of the earth's axial rotation while the moon's orbit is gradually expanding. Both the currently observed increase in the length of the day of 0.0016 s/century and the recession of the moon of 3 to 4 cm/yr are understood as consequences of the tides raised by the moon on the earth. Billions of years from now the moon will be so far from the earth that the duration of the month will be equal to the duration of the day. The tidal evolution of the system ends with synchronization of the axial rotation of both orbiting bodies with their orbital revolution. The length of both the day and month in this final state of coherent rotation will be approximately 50 present days, as can be calculated on the basis of angular-momentum conservation (see, for example, Ref. 13). Similarly, tidal effects on the earth influence its axial rotation and its orbital revolution around the sun.[22]

Tidal dissipation accounts for the current states of axial rotation of several planets, the spin states of most of the planetary satellites, and the spins and orbits of close binary stars. For example, all the major and close planetary satellites in the solar system (with the exception of Saturn's satellite Hyperon) are observed to be rotating synchronously with their orbital motion. The distant planet Pluto and its satellite Charon are the pair in the solar system that has almost certainly reached the end point where further tidal evolution has ceased. In this state the orbit is circular, with both bodies rotating synchronously with the orbital motion and both spin axes perpendicular to the orbital plane. Similarly, many close binary stars are observed to have circular orbits and synchronized spins, providing numerous examples of evolution under tidal forces elsewhere in the Milky Way. The role of tides in the cosmogony was first recognized by the astronomer George Darwin, who developed a theory of the heavenly evolution under tidal friction.[23]

Another interesting manifestation of the tidal forces is the *Roche limit*, the minimum distance to which a large (natural) satellite can approach its primary body without being torn apart by tidal forces. To evaluate this critical distance R_c, we can equate the vertical tidal force, Eq. (6), exerted on a mass point located at $\theta=0$ or $\theta=\pi$ on the surface of a satellite of radius r_{sat} and mass m_{sat} by its primary of mass M, and the force of self-gravitation of the satellite (that is, the force of gravitational attraction of this mass point m to the satellite):

$$2\frac{GmM}{R_c^3}r_{\text{sat}}=\frac{Gmm_{\text{sat}}}{r_{\text{sat}}^2},$$

whence

$$R_c=r_{\text{sat}}\sqrt[3]{\frac{2M}{m_{\text{sat}}}}=r_{\text{planet}}\sqrt[3]{\frac{2\rho}{\rho_{\text{sat}}}}. \tag{25}$$

In Eq. (25) r_{planet} is the radius of the primary, ρ is its mean density, and ρ_{sat} is the satellite's mean density. If the satellite and its primary are of similar composition ($\rho\approx\rho_{\text{sat}}$), the theoretical limit is about $2^{1/3}=1.26$ times the radius of the larger body. The famous rings of Saturn lie inside Saturn's Roche limit and may be the debris of a demolished moon.

ACKNOWLEDGMENTS

The author is grateful to Professor Hermann Härtel of Kiel University for initiating an interesting discussion regarding the common understanding of the physics behind the tidal phenomena. This discussion stimulated the development of the simulation program[10] that visualizes and clarifies the dynamical approach to the oceanic tides. I also appreciate the generous helpful assistance of Professor Stefan Machlup of Case Western Reserve University and Professor Robert Brehme of Wake Forest University.

[a)]Electronic mail: butikov@spb.runnet.ru

[1]E. Tsantes, ''Note on the tides,'' Am. J. Phys. **42** (4), 330–333 (1974).

[2]P. Seligmann and M. Steinberg, ''Simple hydrodynamic treatment of ocean tides,'' Am. J. Phys. **43** (12), 1106–1108 (1975).

[3]E. Horsfield, ''Cause of the earth tides,'' Am. J. Phys. **44** (8), 793–794 (1976).

[4]A. B. Arons, ''Basic physics of the semidiurnal lunar tide,'' Am. J. Phys. **47** (11), 934–937 (1979).

[5]G. M. Kapoulitsas, ''On the generation of tides,'' Eur. J. Phys. **6** (3), 20–207 (1985).

[6]L. M. Celnikier, ''Tidal braking of the Earth's rotation—a study in cubism,'' Eur. J. Phys. **11** (1), 60–62 (1990).

[7]J. M. Woosley, ''Satellites and tides,'' Phys. Educ. **29** (3), 177–179 (1994).

[8]H. Härtel, ''The tides–a neglected topic,'' Phys. Educ. **35** (1), 40–45 (2000).

[9]J. Viiri, ''Student's understanding of tides,'' Phys. Educ. **35** (2), 105–110 (2000).

[10]E. Butikov, *The Ocean Tides*, a set of Java applets at http://www.ifmo.ru/butikov/Projects/Tides0.html.

[11]A very clear physical explanation of tidal forces in which the concept of pseudoforces is avoided can be found in the textbook *Physics*, S. Machlup (Wiley, New York, 1988). See the essay on pp. 125–128 on ''Are we lighter at noon than at midnight?'' A similar approach to the effects of solar gravity on the orbiting earth is related to Einstein's principle of equivalence: The situation in a noninertial reference frame that is falling freely in a gravitational field is equivalent to what would happen in an inertial frame in the absence of this gravitational field. A freely falling closed cabin of an elevator is often used in gedanken experiments that appeal to the principle of equivalence. Because the earth in its orbital motion is falling freely in the gravitational field of the sun, our planet is similar to such a falling cabin in the sense that the sun doesn't interfere much due its gravity in our earthly dealings. However, the freely falling cabin is exactly equivalent to an inertial reference frame only when the gravitational field is uniform. In a nonuniform gravitational field, such as the field of the sun, the equivalence holds approximately only for a small enough cabin. Finite dimensions of the earth (the freely falling ''cabin'') cause departures from the exact equivalence, which are revealed as tidal forces. In other words, we can treat tidal forces as a manifestation of the local nature of the equivalence principle.

[12]D. E. Cartwright, ''Oceanic tides,'' Rep. Prog. Phys. **40** (6), 665–708 (1977).

[13]F. D. Stacey, *Physics of the Earth* (Wiley, New York, 1969), 2nd ed.

[14]As we have emphasized, only the translational acceleration of the earth is essential for the explanation of tides: The tidal forces would have been the same in an imaginary scenario of the earth and sun (earth and moon) freely falling toward their common center of mass under mutual gravitation, without the orbital motion or revolution about this center. The acceleration of the earth caused by its rotation (related both to the monthly revolution about the earth–moon center of mass or to the daily spinning motion) adds nothing to the tidal forces. A misunderstanding of the role of rotation can also be a cause of confusion. For example, in Ref. 15 it is written that ''The planet and the satellite orbit around their mutual center of mass. The centrifugal acceleration that results from this whirling varies with distance from the center of mass across the planet. This effect, together with the differential force of the satellite's gravity across the planet, is responsible for the rising of two tidal bulges.'' However, only the second effect is responsible for producing the tides, while the variation of centripetal acceleration across the planet influences only the nonuniform but constant (time independent) centrifugal force, which only insignificantly modifies the earth's gravity (modifies only the static equilibrium shape of the earth) and thus cannot be responsible for the origin of moving tidal bulges.

[15]P. Goldreich, ''Tides and the earth-moon system,'' Sci. Am. **226** (4), 42–52 (1972).

[16]The equatorial bulge of the earth produced by the time-independent centrifugal force of inertia is responsible for the gradual cyclic change in the direction of the earth's axis of rotation (with a period of almost 26,000 yr). This precession was the third-discovered motion of the earth, after the far more obvious daily rotation and annual revolution. Precession is caused by the gravitational influence of the sun and the moon acting on the earth's equatorial bulge. For tides, the time-independent oblateness of the earth (caused by the daily axial rotation of the earth) is inessential, and the earth can be taken as ideally spherical.

[17]Encyclopedia Britannica CD, see http://www.eb.com.

[18]A. G. Alenitsyn, E. I. Butikov, and A. S. Kondratyev, *Concise Handbook of Mathematics and Physics* (CRC Press, Boca Raton, FL, 1997).

[19]The redistribution of energy during tidal evolution is also a subject of some confusion in the literature. In Ref. 15 the motion of the moon in its orbit is erroneously indicated as one of the sources of the mechanical energy dissipated by tidal friction. However, as the moon's orbit expands, the mechanical energy of the orbital motion increases. This mechanical energy, as well as the energy dissipated by tidal friction, also has its source in the energy of axial rotation of the earth.

[20]M. Zeilik, *Astronomy: The Evolving Universe* (Harper and Row, New York, 1982), 3rd ed.

[21]E. Butikov, *Planets and Satellites* (computer simulation programs), User's Manual (American Institute of Physics, Physics Academic Software, New York, 1998), pp. 39–40.

[22]The variation of the orbital motion during the tidal evolution is also interesting in the sense that it disproves a common belief that internal forces cannot influence the motion of the center of mass of a system. This statement is true only for a system moving in a homogeneous external field. The aforementioned expanding of the earth's orbit during the tidal evolution is actually caused (though indirectly) by internal forces, namely by gravitational forces between the ocean water and the hard body of the earth (earth's self-gravity), and frictional forces. The internal forces change the configuration of the system (consisting of the earth and the ocean water) that moves in the nonhomogeneous external gravitational field. Therefore these internal forces change (indirectly) the resultant external gravitational force exerted on the system. These variations of the external gravitational force modify the orbital motion of the system. During tidal evolution, the internal forces of mutual gravitation together with friction cause the redistribution of masses (of the earth and tidal bulges) that move in the external nonhomogeneous central field.

[23]George H. Darwin, *The Tides and Kindred Phenomena in the Solar System* (Freeman, San Francisco, 1962).