BATTIN R.
An Introduction to the Mathematics and Methods of Astrodynamics, 1999

Chapter 10

Variation of Parameters

ANALYTICAL DEVELOPMENT OF THE VARIATION OF PARAMETERS WAS first given by Leonhard Euler in a series of memoirs on the mutual perturbations of Jupiter and Saturn for which he received the prizes of the French Academy in the years 1748 and 1752. The method is also called *the variation of orbital elements* or *the variation of constants*—the latter referring to integration constants. Euler's treatment of the method of variation of parameters was not entirely general since he did not consider the orbital elements as being simultaneously variable. It is noteworthy, however, that the first steps in the expansion of the disturbing function were made by Euler in those papers.

Joseph-Louis Lagrange wrote his first memoir on the perturbations of Jupiter and Saturn in 1766 in which he made further advances in the variation of parameters method. His final equations were still incorrect because he regarded the major axes and the times of perihelion passage as constants. However, his expressions for the angle of inclination, the longitude of the ascending node, and the argument of perihelion were all perfectly correct. Later, in 1782, he developed completely and for the first time the method of the variation of parameters in a prize memoir on the perturbations of comets moving in elliptical orbits. One of the objectives in this chapter is the derivation of Lagrange's planetary equations.

The most dramatic application of the method was made independently and almost simultaneously by the Englishman John Couch Adams (1819–1892) and the Frenchman Urbain-Jean-Joseph Le Verrier (1811–1877). Each predicted the existence and apparent position of the planet Neptune from the otherwise unexplained irregularities in the motion of Uranus.† The story is one of the most fascinating in the history of astronomy and is an impressive example of the precision which can be achieved using variational methods.

The planet Uranus was discovered on March 13, 1781 by Sir William Herschel shortly before Euler's death in 1783. The other planets had been known since ancient times and Herschel's findings opened the door to an

† A mathematical account of the procedures used by Adams and Le Verrier is given in William Marshall Smart's book *Celestial Mechanics* published in 1953 by Longmans, Green and Co.

era of astronomical discoveries of major importance. Since Uranus is almost visible to the unaided eye, the German astronomer Johann Elbert Bode† (1747–1826) suspected that it might have been mistaken for a star in the past. When the orbital elements had been determined with sufficient accuracy, a search of the old catalogs revealed that Uranus had been observed at least 19 different times—the earliest by the first Astronomer Royal, John Flamsteed, in the year 1690.

When the French astronomer Alexis Bouvard attempted to reconcile the new observations with the old during his preparation of tables for Jupiter, Saturn, and Uranus which were published in 1821, he was forced to abandon the earlier data because of serious and unexplained discrepancies. Even so, the planet began to deviate more and more from Bouvard's predicted positions and, by 1846, the error in longitude was almost two minutes of arc. In 1842, Friedrich Wilhelm Bessel, suspecting the presence of an ultra-Uranian planet, announced his intention of investigating the motion of Uranus. Unfortunately, he died before much could be accomplished.

On July 3, 1841, an undergraduate at St. John's College in Cambridge, England wrote in his journal

> "*Formed a design in the beginning of this week of investigating, as soon as possible after taking my degree, the irregularities in the motion of Uranus … in order to find out whether they may be attributed to the action of an undiscovered planet beyond it …*"

True to his word, by 1845 John Couch Adams had obtained a solution

---

† In 1766 the German astronomer Johann Daniel Titius (1729–1796) of Whittenberg found an empirical formula for the distances of the planets from the sun—a "solution" of the problem to which Kepler devoted so much misplaced energy. According to Titius, the formula for the mean distance

$$a_n = \tfrac{1}{10}\left(4 + 3 \times [2^{n-2}]\right) \text{ a.u.}$$

with $n = 1, 2, 3, 4$ holds for the planets Mercury, Venus, earth, Mars and with $n = 6, 7$ for Jupiter and Saturn. (The symbol $[x]$ denotes the greatest integer contained in x.) The approximation is, indeed, remarkably good. When Uranus was found to conform to the rule for $n = 8$, the formula took on greater significance. The empty space, corresponding to $n = 5$, inspired Johann Bode, director of the Berlin Observatory, to declare

> "*Is it not highly probable that a planet actually revolves in the orbit which the finger of the Almighty has drawn for it? Can we believe the Creator of the world has left this space empty? Certainly not!*"

An association of European astronomers was formed to search for the missing planet. When Ceres was discovered by Giuseppe Piazzi on the first day of January, 1801 at approximately 2.8 a.u., Titius' rule became *Bode's law*. (It is, of course, not a law and, ironically, its association with Titius is almost forgotten.) It is, therefore, not surprising that both men, Adams and Le Verrier, used Bode's law to estimate the mean distance of Neptune as 38.8 a.u.—but it was, in fact, the first planet to violate the rule. (Neptune's mean distance is actually 30.1 astronomical units.)

and in September of that year he gave the results of his computations, on where the new planet could be found, to James Challis, director of the Cambridge observatory. Challis expressed little interest. The next month, Adams contacted the Astronomer Royal, Sir George Biddell Airy, who also reacted with a similar lack of enthusiasm.

Meanwhile, in that same year, the French astronomer Urbain-Jean-Joseph Le Verrier turned his attention to the Uranus problem and *published* his results on June 1, 1846. When Airy saw the close agreement with Adams' calculations, he suggested to Challis on July 9, 1846 that he search for the planet. Indeed, Challis did observe Neptune on August 4 but failed to recognize it as the object of his quest. He had neglected to reconcile his observations with those of the previous night—an unforgivable blunder for a man of his experience.

On September 18, 1846, Le Verrier requested the German astronomer Johann Gottfried Galle to look for the planet with the hope that it could be distinguished from a star by its disk-like appearance. Then on September 23, 1846, after only an hour, Galle found the planet Neptune within one degree of the position computed by Le Verrier.

The reader can imagine the controversy between the English and the French over who deserved the credit. But justice prevailed and when the battle subsided, it was universally agreed that both Adams and Le Verrier would share equally in the glory.

10.1 Variational Methods for Linear Equations

The first application of the method of variation of parameters was made by John Bernoulli† in 1697 to solve the linear differential equation of the first order. The most general such equation is

$$\frac{dy}{dt} + f(t)y = g(t) \tag{10.1}$$

For the solution, consider first the homogeneous linear equation

$$\frac{dy}{dt} + f(t)y = 0 \tag{10.2}$$

† After Newton and Leibnitz, the Bernoulli brothers, James and John, were the two most important founders of the calculus. James Bernoulli (1655–1705) trained for the ministry at the urging of his father but managed to teach himself mathematics. In 1686 he turned to mathematics exclusively and became a professor at the University of Basle. His younger brother John (1667–1748) was steered into business by his father but turned, instead, to medicine and learned mathematics on the side from his brother. Mathematics again won out and he became a professor at Groningen in Holland and, later, succeeded his brother at Basle. His most famous student at the university was Leonhard Euler who completed his studies there at the age of fifteen. It was through the assistance of the younger Bernoullis, Nicholas (1695–1726) and Daniel (1700–1782), both sons of John and both accomplished mathematicians, that Euler in 1733 secured an appointment at the St. Petersburg Academy in Russia.

The Bernoulli family was, indeed, a unique source of mathematical talent.

The variables are separable

$$\frac{dy}{y} = -f(t)\,dt$$

and we have

$$y = ce^{-\int f(t)\,dt} \tag{10.3}$$

where c is a constant.

Suppose now that we allow c to be a function of t and determine the relation that $c(t)$ must satisfy if Eq. (10.3) is to be a solution of the inhomogeneous equation (10.1). By direct substitution we find

$$\frac{dc}{dt}e^{-\int f(t)\,dt} = g(t)$$

Hence

$$c(t) = C + \int g(t)e^{\int f(t)\,dt}\,dt$$

where C is a constant. Thus, the general solution of Eq. (10.1) is

$$y = Ce^{-\int f\,dt} + e^{-\int f\,dt}\int ge^{\int f\,dt}\,dt \tag{10.4}$$

and involves two quadratures.

◇ **Problem 10–1**

Obtain the general solutions of

(1) $$\frac{dy}{dt} - ay = e^{at} \qquad (a = \text{constant})$$

(2) $$\frac{dy}{dt}\cos t + y\sin t = 1$$

using the method of variation of parameters.

Lagrange, in 1774, extended the method to the general n^{th} order linear differential equation

$$L(y) = g(t) \tag{10.5}$$

where the operator $L(y)$ is defined by

$$L(y) \equiv \frac{d^n y}{dt^n} + f_1(t)\frac{d^{n-1}y}{dt^{n-1}} + \cdots + f_{n-1}(t)\frac{dy}{dt} + f_n(t)y$$

It is convenient to convert Eq. (10.5) to a system of n first-order equations written in vector-matrix form. For this purpose, define

$$\mathbf{y}^{\mathrm{T}} = \begin{bmatrix} y & \dfrac{dy}{dt} & \dfrac{d^2y}{dt^2} & \cdots & \dfrac{d^{n-2}y}{dt^{n-2}} & \dfrac{d^{n-1}y}{dt^{n-1}} \end{bmatrix}$$

$$\mathbf{g}^{\mathrm{T}} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & g \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ -f_n & -f_{n-1} & -f_{n-2} & \cdots & -f_2 & -f_1 \end{bmatrix}$$

so that the scalar differential equation (10.5) is equivalent to

$$\frac{d\mathbf{y}}{dt} = \mathbf{F}\mathbf{y} + \mathbf{g} \tag{10.6}$$

Suppose that n linearly independent solutions of the homogeneous equation

$$L(y) = 0 \tag{10.7}$$

are known. Call them $\mathbf{y}_1(t)$, $\mathbf{y}_2(t)$, $\ldots$, $\mathbf{y}_n(t)$ and form a matrix with these vectors as the columns of the array. This is the *Wronskian matrix* $\mathbf{W}$ defined as

$$\mathbf{W} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_n \end{bmatrix}$$

Clearly then, $\mathbf{W}$ satisfies the matrix differential equation

$$\frac{d\mathbf{W}}{dt} = \mathbf{F}\mathbf{W} \tag{10.8}$$

and the general solution of the homogeneous equation is

$$\mathbf{y}_h = \mathbf{W}\mathbf{c} \tag{10.9}$$

where the components of the vector $\mathbf{c}$ are n arbitrary constants.

As before we allow the elements of the vector $\mathbf{c}$ to be functions of t and require that

$$\mathbf{y} = \mathbf{W}\mathbf{c}(t) \tag{10.10}$$

be a solution of Eq. (10.6). That is,

$$\frac{d\mathbf{W}}{dt}\mathbf{c} + \mathbf{W}\frac{d\mathbf{c}}{dt} = \mathbf{F}\mathbf{W}\mathbf{c} + \mathbf{g}$$

But $\mathbf{W}$ is a solution of Eq. (10.8), so that the differential equation for $\mathbf{c}(t)$ is reduced to

$$\mathbf{W}\frac{d\mathbf{c}}{dt} = \mathbf{g} \tag{10.11}$$

Now, the functions $\mathbf{y}_1(t)$, $\mathbf{y}_2(t)$, $\ldots$, $\mathbf{y}_n(t)$ were given as linearly independent so that the Wronskian determinant is not zero. Therefore, the matrix $\mathbf{W}$ is not singular so that

$$\frac{d\mathbf{c}}{dt} = \mathbf{W}^{-1}\mathbf{g} \tag{10.12}$$

which is solved by quadratures for the elements of the vector $\mathbf{c}$.

◇ **Problem 10–2**

Obtain the general solutions of

(1) $$\frac{d^2y}{dt^2} + y = \sec t$$

(2) $$\frac{d^3y}{dt^3} + 4\frac{dy}{dt} = 4\cot 2t$$

(3) $$\frac{d^2y}{dt^2} + 2\frac{dy}{dt} - 3y = te^{-t}$$

◇ **Problem 10–3**

For the second-order linear differential operator

$$L(y) = \frac{d^2y}{dt^2} - \frac{6}{t^2}y$$

show that

$$y_1(t) = t^3 \qquad \text{and} \qquad y_2(t) = \frac{1}{t^2}$$

are linearly independent solutions of $L(y) = 0$. Then use the method of variation of parameters to obtain the general solution of

$$L(y) = t\log t$$

si

10.2 Lagrange's Planetary Equations

The method of the variation of parameters. as originally developed by Lagrange, was to study the disturbed motion of two bodies in the form

$$\frac{d\mathbf{r}}{dt} = \mathbf{v} \qquad \frac{d\mathbf{v}}{dt} + \frac{\mu}{r^3}\mathbf{r} = \left[\frac{\partial R}{\partial \mathbf{r}}\right]^{\mathrm{T}} \tag{10.13}$$

where R is the disturbing function defined in Sect. 8.4. The solution of the undisturbed or two-body motion is known and may be expressed functionally in the form

$$\mathbf{r} = \mathbf{r}(t, \boldsymbol{\alpha}) \qquad \mathbf{v} = \mathbf{v}(t, \boldsymbol{\alpha}) \tag{10.14}$$

where the components of the vector $\boldsymbol{\alpha}$ are the six constants of integration (orbital elements). As in the previous section, we allow $\boldsymbol{\alpha}$ to be a time dependent quantity and require that the two-body solution (10.14) exactly satisfy the equations (10.13) for the disturbed motion.

A set of differential equations for $\boldsymbol{\alpha}(t)$ will result as before; however, they will not be solvable by quadrature. The new set of equations will, in fact, be a transformation of the dependent variables of the problem from the original position and velocity vectors $\mathbf{r}(t)$ and $\mathbf{v}(t)$ to the time-varying orbital elements $\boldsymbol{\alpha}(t)$. Although the differential equations for $\boldsymbol{\alpha}(t)$ will be as complex as the original version, they will have advantages similar to those encountered in Encke's method, i.e., only the disturbing and not the total acceleration will effect changes in $\boldsymbol{\alpha}(t)$. Indeed, one may regard the method of variation of orbital elements as a form of Encke's method in which rectification of the osculating orbit is performed continuously rather than at discrete and widely separated instants of time.

To obtain the variational equations, we substitute Eqs. (10.14) into Eqs. (10.13) and use the fact that

$$\frac{\partial \mathbf{r}}{\partial t} = \mathbf{v} \qquad \frac{\partial \mathbf{v}}{\partial t} + \frac{\mu}{r^3}\mathbf{r} = \mathbf{0} \tag{10.15}$$

Here, the partial derivatives serve to emphasize that when the vector $\boldsymbol{\alpha}$ is considered to be constant, then Eqs. (10.14) are solutions of the equations which describe the undisturbed motion.

For the disturbed motion

$$\frac{d\mathbf{r}}{dt} = \frac{\partial \mathbf{r}}{\partial t} + \frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}}\frac{d\boldsymbol{\alpha}}{dt}$$

and, paralleling the arguments used in the previous section, we have

$$\frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}}\frac{d\boldsymbol{\alpha}}{dt} = \mathbf{0} \tag{10.16}$$

as the condition to be imposed on $\boldsymbol{\alpha}(t)$. Physically, this means we are requiring the velocity vectors of both the disturbed and undisturbed motion to be identical. Similarly,

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{v}}{\partial t} + \frac{\partial \mathbf{v}}{\partial \boldsymbol{\alpha}}\frac{d\boldsymbol{\alpha}}{dt}$$

and, using the second of Eqs. (10.15), we find that

$$\frac{\partial \mathbf{v}}{\partial \boldsymbol{\alpha}}\frac{d\boldsymbol{\alpha}}{dt} = \left[\frac{\partial R}{\partial \mathbf{r}}\right]^{\mathrm{T}} \tag{10.17}$$

must obtain if Eqs. (10.13) are to be satisfied. Equations (10.16) and (10.17) are the required six scalar differential equations to be satisfied by the vector of orbital elements $\boldsymbol{\alpha}(t)$.

The Lagrange Matrix and Lagrangian Brackets

The two matrix-vector variational equations can be combined to produce a more convenient and compact form. For this purpose, we first multiply Eq. (10.16) by $[\partial\mathbf{v}/\partial\boldsymbol{\alpha}]^{\mathrm{T}}$. Then, multiply Eq. (10.17) by $[\partial\mathbf{r}/\partial\boldsymbol{\alpha}]^{\mathrm{T}}$ and subtract the two. The result is expressed as

$$\mathbf{L}\frac{d\boldsymbol{\alpha}}{dt} = \left[\frac{\partial R}{\partial \boldsymbol{\alpha}}\right]^{\mathrm{T}} \tag{10.18}$$

where the matrix

$$\mathbf{L} = \left[\frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}}\right]^{\mathrm{T}} \frac{\partial \mathbf{v}}{\partial \boldsymbol{\alpha}} - \left[\frac{\partial \mathbf{v}}{\partial \boldsymbol{\alpha}}\right]^{\mathrm{T}} \frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}} \tag{10.19}$$

is six-dimensional and skew-symmetric. The form of the right-hand side of Eq. (10.18) follows from the *chain rule* of partial differentiation

$$\frac{\partial R}{\partial \boldsymbol{\alpha}} = \frac{\partial R}{\partial \mathbf{r}} \frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}}$$

The element in the i^{th} row and j^{th} column of the *Lagrange matrix* $\mathbf{L}$ is denoted by $[\alpha_i, \alpha_j]$ and will be referred to as a *Lagrangian bracket*. From Eq. (10.19) we have

$$\begin{aligned}
[\alpha_i, \alpha_j] &= \frac{\partial \mathbf{r}}{\partial \alpha_i} \cdot \frac{\partial \mathbf{v}}{\partial \alpha_j} - \frac{\partial \mathbf{r}}{\partial \alpha_j} \cdot \frac{\partial \mathbf{v}}{\partial \alpha_i} \\
&= \frac{\partial \mathbf{r}^{\mathrm{T}}}{\partial \alpha_i} \frac{\partial \mathbf{v}}{\partial \alpha_j} - \frac{\partial \mathbf{r}^{\mathrm{T}}}{\partial \alpha_j} \frac{\partial \mathbf{v}}{\partial \alpha_i} = \frac{\partial \mathbf{v}^{\mathrm{T}}}{\partial \alpha_j} \frac{\partial \mathbf{r}}{\partial \alpha_i} - \frac{\partial \mathbf{v}^{\mathrm{T}}}{\partial \alpha_i} \frac{\partial \mathbf{r}}{\partial \alpha_j}
\end{aligned} \tag{10.20}$$

An important property of the Lagrange matrix $\mathbf{L}$ is displayed when we calculate the partial derivative of the Lagrangian bracket with respect to t. Thus,

$$\frac{\partial}{\partial t}[\alpha_i, \alpha_j] = \frac{\partial}{\partial \alpha_j}\left(\frac{\partial \mathbf{v}^{\mathrm{T}}}{\partial t}\right) \frac{\partial \mathbf{r}}{\partial \alpha_i} + \frac{\partial \mathbf{v}^{\mathrm{T}}}{\partial \alpha_j} \frac{\partial \mathbf{v}}{\partial \alpha_i} - \frac{\partial}{\partial \alpha_i}\left(\frac{\partial \mathbf{v}^{\mathrm{T}}}{\partial t}\right) \frac{\partial \mathbf{r}}{\partial \alpha_j} - \frac{\partial \mathbf{v}^{\mathrm{T}}}{\partial \alpha_i} \frac{\partial \mathbf{v}}{\partial \alpha_j}$$

and, clearly, the second and fourth terms cancel immediately. Using the gravitational potential function $V = \mu/r$, the second one of Eqs. (10.15) becomes

$$\frac{\partial \mathbf{v}^{\mathrm{T}}}{\partial t} = \frac{\partial V}{\partial \mathbf{r}}$$

so that

$$\begin{aligned}
\frac{\partial}{\partial t}[\alpha_i, \alpha_j] &= \frac{\partial}{\partial \alpha_j}\left(\frac{\partial V}{\partial \mathbf{r}}\right) \frac{\partial \mathbf{r}}{\partial \alpha_i} - \frac{\partial}{\partial \alpha_i}\left(\frac{\partial V}{\partial \mathbf{r}}\right) \frac{\partial \mathbf{r}}{\partial \alpha_j} \\
&= \frac{\partial}{\partial \mathbf{r}}\left(\frac{\partial V}{\partial \alpha_j}\right) \frac{\partial \mathbf{r}}{\partial \alpha_i} - \frac{\partial}{\partial \mathbf{r}}\left(\frac{\partial V}{\partial \alpha_i}\right) \frac{\partial \mathbf{r}}{\partial \alpha_j} \\
&= \frac{\partial^2 V}{\partial \alpha_j \partial \alpha_i} - \frac{\partial^2 V}{\partial \alpha_i \partial \alpha_j} = 0
\end{aligned}$$

In view of this discussion, we can summarize the properties of the Lagrangian brackets as

$$\begin{aligned}
&(1) \qquad [\alpha_i, \alpha_i] = 0 \\
&(2) \qquad [\alpha_i, \alpha_j] = -[\alpha_j, \alpha_i] \\
&(3) \qquad \frac{\partial}{\partial t}[\alpha_i, \alpha_j] = 0
\end{aligned}$$

or, equivalently, for the Lagrange matrix,

$$\mathbf{L}^{\mathrm{T}} = -\mathbf{L} \qquad \text{and} \qquad \frac{\partial \mathbf{L}}{\partial t} = \mathbf{O} \tag{10.21}$$

The fact that the matrix $\mathbf{L}$ is *not an explicit function of* t will be exploited to great advantage in determining the elements of the Lagrange matrix.

Problem 10–4

Consider the case for which the position and velocity vectors $\mathbf{r}_0$ and $\mathbf{v}_0$ at some instant of time t_0 are used as orbital elements, i.e.,

$$\boldsymbol{\alpha}^{\mathrm{T}} = [\mathbf{r}_0^{\mathrm{T}} \quad \mathbf{v}_0^{\mathrm{T}}]$$

Using the fact that the Lagrangian matrix $\mathbf{L}$ is independent of time, show that

$$\mathbf{L} = \mathbf{J}$$

where the matrix $\mathbf{J}$ is defined in Sect. 9.5. Then show that the variational equations are

$$\frac{d\mathbf{r}_0^{\mathrm{T}}}{dt} = -\frac{\partial R}{\partial \mathbf{v}_0} \qquad \frac{d\mathbf{v}_0^{\mathrm{T}}}{dt} = \frac{\partial R}{\partial \mathbf{r}_0}$$

which are in the so-called *canonical form* and should be compared to Hamilton's canonical form of the equations of motion developed in Prob. 2–13.

Computing the Lagrangian Brackets

To compute the Lagrangian brackets we first select an appropriate set of orbital elements. The classical choice is

$$\boldsymbol{\alpha}^{\mathrm{T}} = [\Omega \quad i \quad \omega \quad a \quad e \quad \lambda] \tag{10.22}$$

where Ω, i, ω are the three Euler angles, a is the semimajor axis, e is the eccentricity and

$$\lambda = -n\tau \tag{10.23}$$

where n is the mean motion

$$n = \sqrt{\frac{\mu}{a^3}}$$

and τ is the time of pericenter passage.

The position and velocity vectors, expressed in reference coordinates as functions of the orbital elements, are then

$$\mathbf{r} = \begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{bmatrix} \begin{bmatrix} a(\cos E - e) \\ b \sin E \\ 0 \end{bmatrix} \tag{10.24}$$

$$\mathbf{v} = \begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{bmatrix} \begin{bmatrix} -an \sin E/(1 - e\cos E) \\ bn \cos E/(1 - e\cos E) \\ 0 \end{bmatrix} \tag{10.25}$$

$l_1 = c\Omega c\omega - c i s\omega s\Omega$ $\quad l_2 = -c\Omega s\omega - ci\, c\omega s\Omega$ $\quad l_3 = s\Omega si$

$m_1 = s\Omega c\omega + ci\, s\omega c\Omega$ $\quad m_2 = -s\Omega s\omega + ci\, c\omega c\Omega$ $\quad m_3 = -c\Omega si$

$n_1 = si\, s\omega$ $\quad n_2 = si\, c\omega$ $\quad n_3 = ci$

[The components of **r** and **v** in orbital plane coordinates are obtained from Eqs. (4.40) and the transformation to reference coordinates via the rotation matrix from Eq. (2.3).] Thus, **r** and **v** are functions of the Euler angles through the direction cosine elements of the rotation matrix as given in Eqs. (2.9). The elements a and e enter the relations explicitly through Eqs. (10.24) and (10.25), and implicitly through the mean motion n, the semiminor axis

$$b = a\sqrt{1-e^2}$$

and through the eccentric anomaly E in Kepler's equation

$$E - e\sin E = nt + \lambda \tag{10.26}$$

We begin by calculating the partial derivatives of **r** and **v** with respect to each of the orbital elements. Since the Lagrangian brackets are not explicit functions of time, we may set t equal to any convenient value after the differentiation. The expressions will be simplest at pericenter for which $t = \tau$, $r = q = a(1-e)$, and $E = 0$.

Consider first the partial derivative of **r** with respect to Ω. From Eq. (2.9), we have

$$\begin{aligned}\frac{\partial l_1}{\partial \Omega} &= \frac{\partial}{\partial \Omega}(\cos\Omega\cos\omega - \sin\Omega\sin\omega\cos i)\\ &= -\sin\Omega\cos\omega - \cos\Omega\sin\omega\cos i\\ &= -m_1\end{aligned}$$

and, similarly,

$$\frac{\partial m_1}{\partial \Omega} = l_1 \qquad \frac{\partial n_1}{\partial \Omega} = 0$$

Now since,

$$\begin{bmatrix} a(\cos E - e)\\ b\sin E\\ 0\end{bmatrix} = \begin{bmatrix} q\\ 0\\ 0\end{bmatrix}$$

at pericenter, the derivatives of the other direction cosines do not enter in the calculation of $\partial\mathbf{r}/\partial\Omega$. The result appears in the next equation set.

The derivatives of **r** with respect to i and ω are entirely similar and in exactly the same way we compute the derivatives of **v**. Therefore, at pericenter,

$$\frac{\partial \mathbf{r}}{\partial \Omega} = q\begin{bmatrix}-m_1\\ l_1\\ 0\end{bmatrix} \qquad \frac{\partial \mathbf{r}}{\partial i} = q\sin\omega\begin{bmatrix} l_3\\ m_3\\ n_3\end{bmatrix} \qquad \frac{\partial \mathbf{r}}{\partial \omega} = q\begin{bmatrix} l_2\\ m_2\\ n_2\end{bmatrix}$$

$$\frac{\partial \mathbf{v}}{\partial \Omega} = \frac{nab}{q}\begin{bmatrix}-m_2\\ l_2\\ 0\end{bmatrix} \qquad \frac{\partial \mathbf{v}}{\partial i} = \frac{nab\cos\omega}{q}\begin{bmatrix} l_3\\ m_3\\ n_3\end{bmatrix} \qquad \frac{\partial \mathbf{v}}{\partial \omega} = -\frac{nab}{q}\begin{bmatrix} l_1\\ m_1\\ n_1\end{bmatrix} \tag{10.27}$$

To calculate $\partial\mathbf{r}/\partial a$, we first determine

$$\frac{\partial}{\partial a}[a(\cos E - e)] = \cos E - e - a\sin E\,\frac{\partial E}{\partial a}$$

$$\frac{\partial}{\partial a}(b\sin E) = \sqrt{1-e^2}\sin E + b\cos E\,\frac{\partial E}{\partial a}$$

Then, by differentiating Kepler's equation (10.26)

$$\frac{\partial E}{\partial a} - e\cos E\,\frac{\partial E}{\partial a} = \frac{\partial n}{\partial a}t = -\frac{3n}{2a}t$$

we obtain

$$\frac{\partial E}{\partial a} = -\frac{3nt}{2r}$$

so that at pericenter

$$\frac{\partial}{\partial a}[a(\cos E - e)] = \frac{q}{a} \qquad \text{and} \qquad \frac{\partial}{\partial a}(b\sin E) = -\frac{3bn\tau}{2q}$$

Therefore,

$$\frac{\partial\mathbf{r}}{\partial a} = \frac{q}{a}\begin{bmatrix} l_1 \\ m_1 \\ n_1 \end{bmatrix} - \frac{3bn\tau}{2q}\begin{bmatrix} l_2 \\ m_2 \\ n_2 \end{bmatrix} \tag{10.28}$$

and, similarly,

$$\frac{\partial\mathbf{v}}{\partial a} = \frac{3a^2n^2\tau}{2q^2}\begin{bmatrix} l_1 \\ m_1 \\ n_1 \end{bmatrix} - \frac{bn}{2q}\begin{bmatrix} l_2 \\ m_2 \\ n_2 \end{bmatrix} \tag{10.29}$$

To calculate the derivatives with respect to e and λ, note that

$$\frac{\partial E}{\partial e} - \sin E - e\cos E\,\frac{\partial E}{\partial e} = 0 \qquad \text{and} \qquad \frac{\partial E}{\partial \lambda} - e\cos E\,\frac{\partial E}{\partial \lambda} = 1$$

Hence, for $E = 0$, we have

$$\frac{\partial E}{\partial e} = 0 \qquad \text{and} \qquad \frac{\partial E}{\partial \lambda} = \frac{a}{q}$$

and the rest of the derivation is as before. There obtains

$$\begin{aligned}
\frac{\partial\mathbf{r}}{\partial e} &= -a\begin{bmatrix} l_1 \\ m_1 \\ n_1 \end{bmatrix} \qquad & \frac{\partial\mathbf{r}}{\partial \lambda} &= \frac{ab}{q}\begin{bmatrix} l_2 \\ m_2 \\ n_2 \end{bmatrix} \\
\frac{\partial\mathbf{v}}{\partial e} &= \frac{na^3}{bq}\begin{bmatrix} l_2 \\ m_2 \\ n_2 \end{bmatrix} \qquad & \frac{\partial\mathbf{v}}{\partial \lambda} &= -\frac{na^3}{q^2}\begin{bmatrix} l_1 \\ m_1 \\ n_1 \end{bmatrix}
\end{aligned} \tag{10.30}$$

all of which are, of course, valid only at pericenter.

With all of the derivatives evaluated, it is a simple task to calculate the Lagrangian brackets defined in Eq. (10.20). For example,

$$\begin{aligned}
[i,\Omega] &= \frac{\partial \mathbf{r}}{\partial i}\cdot\frac{\partial \mathbf{v}}{\partial \Omega} - \frac{\partial \mathbf{r}}{\partial \Omega}\cdot\frac{\partial \mathbf{v}}{\partial i} \\
&= nab[(l_2 m_3 - l_3 m_2)\sin\omega - (l_1 m_3 - l_3 m_1)\cos\omega] \\
&= nab(n_1 \sin\omega + n_2 \cos\omega) \\
&= nab \sin i
\end{aligned}$$

and again

$$\begin{aligned}
[\lambda,a] &= \frac{\partial \mathbf{r}}{\partial \lambda}\cdot\frac{\partial \mathbf{v}}{\partial a} - \frac{\partial \mathbf{r}}{\partial a}\cdot\frac{\partial \mathbf{v}}{\partial \lambda} \\
&= -\frac{ab^2 n}{2q^2} + \frac{na^3 q}{q^2 a} = \frac{na}{2}\left(\frac{2aq - b^2}{q^2}\right) \\
&= \tfrac{1}{2}na
\end{aligned}$$

Because of the skew-symmetry of the matrix $\mathbf{L}$, there are just 15 distinct brackets to evaluate and only six of these turn out to be different from zero. The results are summarized as follows:

$$[i,\Omega] = nab\sin i$$

$$[\omega,\Omega] = 0 \qquad [\omega,i] = 0$$

$$[a,\Omega] = -\tfrac{1}{2}nb\cos i \quad [a,i] = 0 \quad [a,\omega] = -\tfrac{1}{2}nb$$

$$[e,\Omega] = \frac{na^3 e}{b}\cos i \quad [e,i] = 0 \quad [e,\omega] = \frac{na^3 e}{b} \quad [e,a] = 0$$

$$[\lambda,\Omega] = 0 \qquad [\lambda,i] = 0 \quad [\lambda,\omega] = 0 \qquad [\lambda,a] = \tfrac{1}{2}na \quad [\lambda,e] = 0$$

With the elements of the Lagrange matrix determined, Eq. (10.18) may be written in component form as

$$\begin{aligned}
-nab\sin i\frac{di}{dt} + \frac{nb}{2}\cos i\frac{da}{dt} - \frac{na^3 e}{b}\cos i\frac{de}{dt} &= \frac{\partial R}{\partial \Omega} \\
nab\sin i\frac{d\Omega}{dt} &= \frac{\partial R}{\partial i} \\
\frac{nb}{2}\frac{da}{dt} - \frac{na^3 e}{b}\frac{de}{dt} &= \frac{\partial R}{\partial \omega} \\
-\frac{nb}{2}\cos i\frac{d\Omega}{dt} - \frac{nb}{2}\frac{d\omega}{dt} - \frac{na}{2}\frac{d\lambda}{dt} &= \frac{\partial R}{\partial a} \\
\frac{na^3 e}{b}\cos i\frac{d\Omega}{dt} + \frac{na^3 e}{b}\frac{d\omega}{dt} &= \frac{\partial R}{\partial e} \\
\frac{na}{2}\frac{da}{dt} &= \frac{\partial R}{\partial \lambda}
\end{aligned}$$

These are easily solved for the derivatives of the orbital elements to produce the classical form of *Lagrange's planetary equations*:

$$
\begin{aligned}
\frac{d\Omega}{dt} &= \frac{1}{nab\sin i}\frac{\partial R}{\partial i} \\
\frac{di}{dt} &= -\frac{1}{nab\sin i}\frac{\partial R}{\partial \Omega} + \frac{\cos i}{nab\sin i}\frac{\partial R}{\partial \omega} \\
\frac{d\omega}{dt} &= -\frac{\cos i}{nab\sin i}\frac{\partial R}{\partial i} + \frac{b}{na^3 e}\frac{\partial R}{\partial e} \\
\frac{da}{dt} &= \frac{2}{na}\frac{\partial R}{\partial \lambda} \\
\frac{de}{dt} &= -\frac{b}{na^3 e}\frac{\partial R}{\partial \omega} + \frac{b^2}{na^4 e}\frac{\partial R}{\partial \lambda} \\
\frac{d\lambda}{dt} &= -\frac{2}{na}\frac{\partial R}{\partial a} - \frac{b^2}{na^4 e}\frac{\partial R}{\partial e}
\end{aligned}
\tag{10.31}
$$

Equations (10.31) demonstrate explicitly that the matrix L is nonsingular so long as the eccentricity e is neither zero nor one and the inclination angle i is not zero. It should be remarked that a different choice of orbital elements will alleviate these annoying singularities as seen in a later section of this chapter.

◇ Problem 10–5

The Lagrangian brackets are the sum of three Jacobians

$$[\alpha_i, \alpha_j] = \frac{\partial(x, \dot{x})}{\partial(\alpha_i, \alpha_j)} + \frac{\partial(y, \dot{y})}{\partial(\alpha_i, \alpha_j)} + \frac{\partial(z, \dot{z})}{\partial(\alpha_i, \alpha_j)}$$

where $\dot{x}$, $\dot{y}$, and $\dot{z}$ denote the time derivatives of x, y, and z.

NOTE: The *Jacobian* is a determinant defined by

$$\frac{\partial(u, v)}{\partial(\alpha, \beta)} = \begin{vmatrix} \frac{\partial u}{\partial \alpha} & \frac{\partial u}{\partial \beta} \\ \frac{\partial v}{\partial \alpha} & \frac{\partial v}{\partial \beta} \end{vmatrix}$$

Problem 10–6

Consider a new set of orbital elements $[\boldsymbol{\alpha}^{\mathrm{T}} \quad \boldsymbol{\beta}^{\mathrm{T}}]$ where

$$\boldsymbol{\alpha} = \begin{bmatrix} -\frac{1}{2}n^2a^2 \\ nab \\ nab\cos i \end{bmatrix} \qquad \boldsymbol{\beta} = \begin{bmatrix} -\tau \\ \omega \\ \Omega \end{bmatrix}$$

so that in Eqs. (10.31) the disturbing function $R = R(\Omega, i, \omega, a, e, \lambda)$ is to be replaced by

$$R = R^*(\boldsymbol{\alpha}, \boldsymbol{\beta})$$

First, verify the relations

$$\frac{\partial R}{\partial \Omega} = \frac{\partial R^*}{\partial \beta_3}$$
$$\frac{\partial R}{\partial i} = -nab \sin i \frac{\partial R^*}{\partial \alpha_3}$$
$$\frac{\partial R}{\partial \omega} = \frac{\partial R^*}{\partial \beta_2}$$
$$\frac{\partial R}{\partial a} = \frac{n^2 a}{2}\frac{\partial R^*}{\partial \alpha_1} + \frac{nb}{2}\frac{\partial R^*}{\partial \alpha_2} + \frac{nb}{2}\cos i \frac{\partial R^*}{\partial \alpha_3} + \frac{3\lambda}{2na}\frac{\partial R^*}{\partial \beta_1}$$
$$\frac{\partial R}{\partial e} = -\frac{na^3 e}{b}\frac{\partial R^*}{\partial \alpha_2} - \frac{na^3 e}{b}\cos i \frac{\partial R^*}{\partial \alpha_3}$$
$$\frac{\partial R}{\partial \lambda} = \frac{1}{n}\frac{\partial R^*}{\partial \beta_1}$$

and then show that Lagrange's planetary equations, in terms of the alternate set of orbital elements, are in canonical form; i.e.,

$$\frac{d\boldsymbol{\alpha}}{dt}^{\mathrm{T}} = \frac{\partial R^*}{\partial \boldsymbol{\beta}} \quad \text{and} \quad \frac{d\boldsymbol{\beta}}{dt}^{\mathrm{T}} = -\frac{\partial R^*}{\partial \boldsymbol{\alpha}}$$

The partitioned vector elements $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are said to be *canonically conjugate*. NOTE: The orbital elements α_1, α_2, α_3 are, respectively, the total energy, the angular momentum, and the component of the angular momentum vector along the reference z axis.

sí

10.3 Gauss' Form of the Variational Equations

Although Lagrange's variational equations were derived for the special case in which the disturbing acceleration was represented as the gradient of the disturbing function, this restriction is wholly unnecessary. If the disturbed relative motion of two bodies is formulated as in Sect. 9.4 according to

$$\frac{d^2\mathbf{r}}{dt^2} + \frac{\mu}{r^3}\mathbf{r} = \mathbf{a}_d \tag{10.32}$$

then it is readily seen that the derivation in the previous section leading to Eq. (10.18) is still valid with the result now expressed as

$$\mathbf{L}\frac{d\boldsymbol{\alpha}}{dt} = \left[\frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}}\right]^{\mathrm{T}} \mathbf{a}_d \tag{10.33}$$

The elements of the Lagrange matrix, i.e., the Lagrangian brackets, are as calculated in the previous section. However, the matrix coefficient of the disturbing acceleration vector $\mathbf{a}_d$ in Eq. (10.33) is needed to obtain the appropriate variational equations. (The reader should understand that although $\partial \mathbf{r}/\partial \boldsymbol{\alpha}$ was computed in the previous section as a part of the determination of the Lagrangian brackets, the derivatives so obtained were valid only at the instantaneous pericenter.) We now derive the variational equations appropriate to various choices for component resolutions of the disturbing acceleration vector.

All of the equations derived in the following subsection and attributed to Gauss can also be obtained more simply from the equations of Sect. 10.5. Because of the complexity of the results, it is useful that they be derived using two different methods. It is left for the reader to verify that the two sets of variational equations so obtained are, indeed, equivalent.

Gauss' Equations in Polar Coordinates

The rotation matrix

$$\mathbf{R} = \begin{bmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \\ n_1 & n_2 & n_3 \end{bmatrix}$$

will affect an orthogonal transformation of vector components from osculating orbital plane coordinates $\mathbf{i}_e$, $\mathbf{i}_p$, $\mathbf{i}_h$ to the reference coordinates $\mathbf{i}_x$, $\mathbf{i}_y$, $\mathbf{i}_z$. The direction cosine elements of $\mathbf{R}$ are related to the Euler angles through Eqs. (2.9).

Using the asterisk to distinguish a vector resolved along reference axes from the same vector resolved along osculating axes, we have

$$\mathbf{r}^* = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \qquad \mathbf{r} = \begin{bmatrix} r \\ 0 \\ 0 \end{bmatrix} \qquad \mathbf{a}_d^* = \begin{bmatrix} a_{dx} \\ a_{dy} \\ a_{dz} \end{bmatrix} \qquad \mathbf{a}_d = \begin{bmatrix} a_{dr} \\ a_{d\theta} \\ a_{dh} \end{bmatrix}$$

so that

$$\mathbf{r}^* = \mathbf{R}\mathbf{R}_f\mathbf{r} \qquad \text{and} \qquad \mathbf{a}_d^* = \mathbf{R}\mathbf{R}_f\mathbf{a}_d \tag{10.34}$$

The rotation matrix

$$\mathbf{R}_f = \begin{bmatrix} \cos f & -\sin f & 0 \\ \sin f & \cos f & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} (a/r)(\cos E - e) & -(b/r)\sin E & 0 \\ (b/r)\sin E & (a/r)(\cos E - e) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

provides the necessary transformation from local osculating polar coordinates $\mathbf{i}_r$, $\mathbf{i}_\theta$, $\mathbf{i}_h$ to the orbital plane coordinates $\mathbf{i}_e$, $\mathbf{i}_p$, $\mathbf{i}_h$.

Let α be any one of the six orbital elements. Then, to derive the variational equations in terms of the osculating polar components of the disturbing acceleration, we may calculate

$$\frac{\partial \mathbf{r}^{*\mathrm{T}}}{\partial \alpha}\mathbf{a}_d^* \equiv \mathbf{a}_d^{*\mathrm{T}}\frac{\partial \mathbf{r}^*}{\partial \alpha}$$

and replace the term $\partial R/\partial \alpha$ with this quantity in the Lagrange planetary equations (10.31).

For the three Euler angle elements, we first obtain

$$\frac{\partial \mathbf{R}}{\partial \Omega} = \begin{bmatrix} -m_1 & -m_2 & -m_3 \\ l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\frac{\partial \mathbf{R}}{\partial i} = \begin{bmatrix} l_3 \sin\omega & l_3 \cos\omega & \sin\Omega \cos i \\ m_3 \sin\omega & m_3 \cos\omega & -\cos\Omega \cos i \\ n_3 \sin\omega & n_3 \cos\omega & -\sin i \end{bmatrix}$$

$$\frac{\partial \mathbf{R}}{\partial \omega} = \begin{bmatrix} l_2 & -l_1 & 0 \\ m_2 & -m_1 & 0 \\ n_2 & -n_1 & 0 \end{bmatrix}$$

Then,

$$\frac{\partial R}{\partial \alpha} = \mathbf{a}_d^{*\mathrm{T}} \frac{\partial \mathbf{r}^*}{\partial \alpha} = \mathbf{a}_d^{\mathrm{T}} \mathbf{R}_f^{\mathrm{T}} \mathbf{R}^{\mathrm{T}} \frac{\partial \mathbf{R}}{\partial \alpha} \mathbf{R}_f \mathbf{r} \tag{10.35}$$

is evaluated for $\alpha = \Omega,\ i,\ \omega$. We have

$$\frac{\partial R}{\partial \Omega} = \mathbf{a}_d^{\mathrm{T}} \begin{bmatrix} 0 \\ r\cos i \\ -r\cos\theta \sin i \end{bmatrix}$$

$$\frac{\partial R}{\partial i} = \mathbf{a}_d^{\mathrm{T}} \begin{bmatrix} 0 \\ 0 \\ r\sin\theta \end{bmatrix} \qquad \frac{\partial R}{\partial \omega} = \mathbf{a}_d^{\mathrm{T}} \begin{bmatrix} 0 \\ r \\ 0 \end{bmatrix} \tag{10.36}$$

where

$$\theta = \omega + f$$

is the argument of latitude defined in Sect. 3.4.

The vector $\mathbf{r}^*$ depends on the remaining three elements through the vector

$$\mathbf{R}_f \mathbf{r} = \begin{bmatrix} a(\cos E - e) \\ b\sin E \\ 0 \end{bmatrix}$$

The derivatives with respect to a, e, λ (following the arguments used in computing the Lagrangian brackets) are obtained as follows:

$$\frac{\partial}{\partial a}(\mathbf{R}_f \mathbf{r}) = \begin{bmatrix} \cos E - e + (3ant/2r)\sin E \\ (b/a)\sin E - (3ant/2r)\cos E \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} (r/a)\cos f + (3ant/2b)\sin f \\ (r/a)\sin f - (3ant/2b)(e + \cos f) \\ 0 \end{bmatrix}$$

$$\frac{\partial}{\partial e}(\mathbf{R}_f\mathbf{r}) = \begin{bmatrix} -(a^2/r)\sin^2 E - a \\ -(a^2e/b)\sin E + (ab/r)\sin E\cos E \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} -(a^2r/b^2)\sin^2 f - a \\ (a^2r/b^2)\sin f\cos f \\ 0 \end{bmatrix}$$

$$\frac{\partial}{\partial \lambda}(\mathbf{R}_f\mathbf{r}) = \begin{bmatrix} -(a^2/r)\sin E \\ (ab/r)\cos E \\ 0 \end{bmatrix} = \begin{bmatrix} -(a^2/b)\sin f \\ (a^2/b)(e+\cos f) \\ 0 \end{bmatrix}$$

Then, for $\alpha = a,\ e,\ \lambda$, we evaluate

$$\frac{\partial R}{\partial \alpha} = \mathbf{a}_d^{*\,\mathrm{T}}\frac{\partial \mathbf{r}^*}{\partial \alpha} = \mathbf{a}_d\mathbf{R}_f^{\mathrm{T}}\mathbf{R}^{\mathrm{T}}\mathbf{R}\frac{\partial}{\partial \alpha}(\mathbf{R}_f\mathbf{r}) = \mathbf{a}_d\mathbf{R}_f^{\mathrm{T}}\frac{\partial}{\partial \alpha}(\mathbf{R}_f\mathbf{r}) \tag{10.37}$$

and obtain

$$\frac{\partial R}{\partial a} = \mathbf{a}_d^{\mathrm{T}}\begin{bmatrix} (r/a) - (3ant/2b)e\sin f \\ -(3ant/2b)(1+e\cos f) \\ 0 \end{bmatrix}$$

$$\frac{\partial R}{\partial e} = \mathbf{a}_d^{\mathrm{T}}\begin{bmatrix} -a\cos f \\ [1+(ar/b^2)]a\sin f \\ 0 \end{bmatrix} \qquad \frac{\partial R}{\partial \lambda} = \mathbf{a}_d^{\mathrm{T}}\begin{bmatrix} (a^2/b)e\sin f \\ (a^2/b)(1+e\cos f) \\ 0 \end{bmatrix} \tag{10.38}$$

Eliminating the Secular Term

Before writing out the complete set of variational equations, let us address an undesirable complication caused by the presence of the linear function of time t in the expression for $\partial R/\partial a$. If we examine the Lagrange planetary equations (10.31), we see that $\partial R/\partial a$ appears in, and only in, the equation for the time rate of change of the element λ. An element exhibiting such behavior is clearly inconvenient at best when large values of t are to be considered and should, therefore, be avoided if at all possible. Fortunately, the difficulty can be overcome in the following manner.

Differentiate the mean anomaly

$$M = nt + \lambda$$

to obtain

$$\frac{dM}{dt} = n - \frac{3nt}{2a}\frac{da}{dt} + \frac{d\lambda}{dt}$$

since the derivative of the mean motion $n = \sqrt{\mu/a^3}$ is simply

$$\frac{dn}{dt} = -\frac{3n}{2a}\frac{da}{dt}$$

Then, using Eqs. (10.31), we have

$$\frac{dM}{dt} = n - \frac{2}{na}\left(\frac{3nt}{2a}\frac{\partial R}{\partial \lambda} + \frac{\partial R}{\partial a}\right) - \frac{b^2}{na^4 e}\frac{\partial R}{\partial e}$$

It is apparent from Eqs. (10.38) that the parenthesized factor in this last equation does not contain t explicitly because of the cancellation. Therefore, an effective artifice for avoiding the difficulty associated with the choice of λ as an orbital element is to replace the variational equation for λ in Lagrange's equations by

$$\frac{dM}{dt} = n + \frac{d\beta}{dt} \tag{10.39}$$

where

$$\frac{d\beta}{dt} = -\frac{3t}{a^2}\frac{\partial R}{\partial \lambda} - \frac{2}{na}\frac{\partial R}{\partial a} - \frac{b^2}{na^4 e}\frac{\partial R}{\partial e} \tag{10.40}$$

The quantity β is then to be regarded as the sixth orbital element instead of $\lambda = -n\tau$.

Summary of Gauss' Equations

Finally, we are ready to summarize the complete set of variational equations. By substituting Eqs. (10.36) and (10.38) into Lagrange's planetary equations (noting that $p = b^2/a$ and $h = nab$), we obtain

$$\begin{aligned}
\frac{d\Omega}{dt} &= \frac{r\sin\theta}{h\sin i}a_{dh} \\
\frac{di}{dt} &= \frac{r\cos\theta}{h}a_{dh} \\
\frac{d\omega}{dt} &= \frac{1}{he}[-p\cos f\, a_{dr} + (p+r)\sin f\, a_{d\theta}] - \frac{r\sin\theta\cos i}{h\sin i}a_{dh} \\
\frac{da}{dt} &= \frac{2a^2}{h}\left(e\sin f\, a_{dr} + \frac{p}{r}a_{d\theta}\right) \\
\frac{de}{dt} &= \frac{1}{h}\{p\sin f\, a_{dr} + [(p+r)\cos f + re]a_{d\theta}\} \\
\frac{dM}{dt} &= n + \frac{b}{ahe}[(p\cos f - 2re)a_{dr} - (p+r)\sin f a_{d\theta}]
\end{aligned} \tag{10.41}$$

(It should be noted that variational equations for either the eccentric or true anomaly may be used in place of the sixth equation above for the mean anomaly. The appropriate equations are the subject of a problem later in this section.)

If initial conditions are specified for Ω, i, ω, a, e, M, these differential equations may be integrated by any convenient numerical method. Needless to say, as a part of the integration process, Kepler's equation must be solved for the osculating eccentric anomaly and the osculating true anomaly determined from an appropriate identity; specifically,

$$M = E - e \sin E \qquad \text{and} \qquad \tan \tfrac{1}{2} f = \sqrt{\frac{1+e}{1-e}} \tan \tfrac{1}{2} E$$

Generally, when the disturbing acceleration is small, a relatively large integration step can be employed. On the other hand, it is necessary to point out that, for this particular choice of orbital elements, the advantage of the variational method is lost for orbits of low inclination or small eccentricity. In these singular cases, the rates of change of Ω, ω and/or β will be large despite the fact that the disturbing acceleration is small. Particular techniques for avoiding these difficulties are treated in later sections.

◇ Problem 10–7

Let a_{dt} and a_{dn} be the components of the disturbing acceleration in the plane of the osculating orbit along the velocity vector and perpendicular to it. Show that

$$\begin{bmatrix} a_{dr} \\ a_{d\theta} \end{bmatrix} = \frac{h}{pv} \begin{bmatrix} e \sin f & -(1 + e\cos f) \\ 1 + e\cos f & e \sin f \end{bmatrix} \begin{bmatrix} a_{dt} \\ a_{dn} \end{bmatrix}$$

and then derive the variational equations in the form

$$\frac{d\Omega}{dt} = \frac{r \sin\theta}{h \sin i} a_{dh}$$

$$\frac{di}{dt} = \frac{r\cos\theta}{h} a_{dh}$$

$$\frac{d\omega}{dt} = \frac{1}{ev}\left[2 \sin f\, a_{dt} + \left(2e + \frac{r}{a}\cos f\right) a_{dn}\right] - \frac{r \sin\theta \cos i}{h \sin i} a_{dh}$$

$$\frac{da}{dt} = \frac{2a^2 v}{\mu} a_{dt}$$

$$\frac{de}{dt} = \frac{1}{v}\left[2(e + \cos f) a_{dt} - \frac{r}{a} \sin f\, a_{dn}\right]$$

$$\frac{dM}{dt} = n - \frac{b}{eav}\left[2\left(1 + \frac{e^2 r}{p}\right) \sin f\, a_{dt} + \frac{r}{a}\cos f\, a_{dn}\right]$$

◇ Problem 10–8

The variational equations for the eccentric and the true anomalies are, in polar coordinates,

$$\frac{dE}{dt} = \frac{na}{r} + \frac{1}{nae}\left[(\cos f - e)\, a_{dr} - \left(1 + \frac{r}{a}\right) \sin f\, a_{d\theta}\right]$$

$$\frac{df}{dt} = \frac{h}{r^2} + \frac{1}{eh}[p \cos f\, a_{dr} - (p + r) \sin f\, a_{d\theta}]$$

and, in tangential–normal coordinates,

$$\frac{dE}{dt} = \frac{na}{r} - \frac{1}{ebv}[2a \sin f\, a_{dt} + r(e + \cos f) a_{dn}]$$

$$\frac{df}{dt} = \frac{h}{r^2} - \frac{1}{ev}\left[2 \sin f\, a_{dt} + \left(2e + \frac{r}{a}\cos f\right) a_{dn}\right]$$

◇ Problem 10–9

The disturbing function for the constant radial thrust acceleration problem of Sect. 8.8 is simply

$$R = r a_{Tr}$$

Derive the variational equations (10.41) for this case directly from the Lagrange planetary equations.

NO

10.4 Nonsingular Elements

For orbits of zero inclination angle, the line of nodes does not exist. For orbits of zero eccentricity the line of apsides is meaningless. Therefore, it is not surprising to find singularities in the variational equations for those elements associated with the node or pericenter. These are the longitude of the node Ω, the argument of pericenter ω, the time of pericenter passage τ (or $\lambda = -n\tau$), and any of the anomalies which are measured from pericenter.

To find variational equations which are nonsingular, we must search for combinations of the usual orbital elements which do not depend on either the line of nodes or the apsidal line. For example, if we add the variational equations for Ω and ω, the resulting equation exhibits no singularity for vanishing inclination angle i. Specifically, from the first and third of Eqs. (10.41), we have

$$\frac{d\varpi}{dt} = \frac{1}{he}[-p\cos f\, a_{dr} + (p+r)\sin f\, a_{d\theta}] + \frac{r}{h}\sin\theta\tan\tfrac{1}{2}i\, a_{dh}$$

where

$$\varpi = \Omega + \omega \tag{10.42}$$

is the longitude of pericenter as defined in Sect. 3.4.

The singularity due to zero eccentricity is still present so that ϖ itself is not a suitable nonsingular orbital element. However, by adding together the variational equations, for ϖ and M, we obtain an equation devoid of either singularity. Since

$$\frac{b}{ahe} - \frac{1}{he} = \frac{b-a}{ahe} = \frac{b^2-a^2}{ahe(a+b)} = -\frac{ae}{h(a+b)}$$

it follows that

$$\frac{dl}{dt} = n - \frac{ae}{h(a+b)}[p\cos f\, a_{dr} - (p+r)\sin f\, a_{d\theta}] - \frac{2br}{ah}a_{dr} + \frac{r\sin\theta\tan\frac{1}{2}i}{h}a_{dh}$$

where

$$l = \varpi + M \tag{10.43}$$

is the mean longitude defined in Sect. 4.3.

Clearly, l should replace M in our set of nonsingular variables, but the equation just obtained is not yet suitable since it involves the true anomaly

f which is referenced to pericenter. To pursue the question further, let us examine the augmented form of Kepler's equation

$$l = \varpi + M = \varpi + E - e\sin E$$

If we define

$$K = \varpi + E \tag{10.44}$$

as the *eccentric longitude*, corresponding to the mean longitude l, then Kepler's equation becomes

$$l = K + e\sin\varpi\cos K - e\cos\varpi\sin K$$

Furthermore, the equation of orbit may be written either in terms of K or in terms of the true longitude

$$L = \varpi + f \tag{10.45}$$

also defined in Sect. 4.3. We have

$$r = a(1 - e\cos E) = a(1 - e\sin\varpi\sin K - e\cos\varpi\cos K)$$

or

$$r = \frac{p}{1 + e\cos f} = \frac{p}{1 + e\sin\varpi\sin L + e\cos\varpi\cos L}$$

Observe that both in the equation of orbit and in Kepler's equation the eccentricity e and the longitude of pericenter ϖ appear only in the combinations $e\sin\varpi$ and $e\cos\varpi$. These functions are, therefore, promising candidates for new elements to replace e and ϖ.

Therefore, define P_1 and P_2 as orbital elements, where

$$P_1 = e\sin\varpi \quad \text{and} \quad P_2 = e\cos\varpi \tag{10.46}$$

and obtain variational equations by differentiating and using the variational equations already obtained for e and ϖ. Hence,

$$\begin{aligned}\frac{dP_1}{dt} &= e\cos\varpi\frac{d\varpi}{dt} + \sin\varpi\frac{de}{dt}\\ &= -\frac{1}{h}[p\cos L\, a_{dr} - (p+r)\sin L\, a_{d\theta} - rP_1 a_{d\theta}] + \frac{r\sin\theta\tan\frac{1}{2}i}{h}P_2 a_{dh}\end{aligned}$$

with a similar expression for P_2.

Although these equations are nonsingular, the argument of latitude θ needs to be expressed in terms of the true longitude L. For this purpose, we write

$$\theta = \omega + f = L - \Omega$$

so that

$$\sin\theta = \sin L\cos\Omega - \cos L\sin\Omega$$

Now, we know that Ω is not itself a nonsingular element. However, $\sin\theta$ appears in the variational equation for P_1 multiplied by $\tan\frac{1}{2}i$ suggesting that the functions $\tan\frac{1}{2}i\sin\Omega$ and $\tan\frac{1}{2}i\cos\Omega$ would be suitable candidates for new elements to replace Ω and i.

Again, we are led to define Q_1 and Q_2 as orbital elements, where

$$Q_1 = \tan\tfrac{1}{2}i\sin\Omega \qquad \text{and} \qquad Q_2 = \tan\tfrac{1}{2}i\cos\Omega \tag{10.47}$$

and obtain

$$\begin{aligned}\frac{dQ_1}{dt} &= \tan\tfrac{1}{2}i\cos\Omega\frac{d\Omega}{dt} + \frac{1}{2}\sec^2\tfrac{1}{2}i\sin\Omega\frac{di}{dt}\\ &= \frac{r}{2h}\sec^2\tfrac{1}{2}i(\sin\theta\cos\Omega + \cos\theta\sin\Omega)a_{dh}\\ &= \frac{r}{2h}(1 + Q_1^2 + Q_2^2)\sin L\, a_{dh}\end{aligned}$$

with a similar result for Q_2. The element set is now complete.

Finally, we note that the classical elements are easily recoverable from the new elements. For example,

$$e^2 = P_1^2 + P_2^2 \qquad \tan^2\tfrac{1}{2}i = Q_1^2 + Q_2^2$$

$$\tan\varpi = \frac{P_1}{P_2} \qquad \tan\Omega = \frac{Q_1}{Q_2}$$

provided, of course, that P_2 and Q_2 are not zero.

We now summarize the variational equations for the elements a, P_1, P_2, Q_1, Q_2, l which have recently been named the *equinoctial variables* by Professor Roger A. Broucke of the University of Texas. They are, indeed, nonsingular except for the rectilinear orbit $h = 0$ and for the orbit whose inclination angle $i = \pi$. (These singularities can also be eliminated but we will not pursue the question further.)

With P_1, P_2, Q_1, and Q_2 chosen to replace the classical elements e, Ω, i, and ω and defined as†

$$\begin{aligned} P_1 &= e\sin\varpi & Q_1 &= \tan\tfrac{1}{2}i\sin\Omega\\ P_2 &= e\cos\varpi & Q_2 &= \tan\tfrac{1}{2}i\cos\Omega\end{aligned}$$

1. The equation for the semimajor axis is

$$\frac{da}{dt} = \frac{2a^2}{h}\left[(P_2\sin L - P_1\cos L)a_{dr} + \frac{p}{r}a_{d\theta}\right] \tag{10.48}$$

2. The equations for P_1, P_2, Q_1, and Q_2 are

$$\frac{dP_1}{dt} = \frac{r}{h}\left\{-\frac{p}{r}\cos L\, a_{dr} + \left[P_1 + \left(1 + \frac{p}{r}\right)\sin L\right]a_{d\theta} - P_2(Q_1\cos L - Q_2\sin L)a_{dh}\right\} \tag{10.49}$$

† Lagrange first introduced this element set (using i instead of $\frac{1}{2}i$) in 1774 for his study of secular variations. His notation for the four elements was h, l, p, and q.

$$\frac{dP_2}{dt} = \frac{r}{h}\left\{\frac{p}{r}\sin L\, a_{dr} + \left[P_2 + \left(1+\frac{p}{r}\right)\cos L\right] a_{d\theta} + P_1(Q_1\cos L - Q_2\sin L)a_{dh}\right\} \tag{10.50}$$

$$\frac{dQ_1}{dt} = \frac{r}{2h}(1+Q_1^2+Q_2^2)\sin L\, a_{dh} \tag{10.51}$$

$$\frac{dQ_2}{dt} = \frac{r}{2h}(1+Q_1^2+Q_2^2)\cos L\, a_{dh} \tag{10.52}$$

3. The equation for the mean longitude is

$$\frac{dl}{dt} = n - \frac{r}{h}\left\{\left[\frac{a}{a+b}\left(\frac{p}{r}\right)(P_1\sin L + P_2\cos L) + \frac{2b}{a}\right] a_{dr} + \frac{a}{a+b}\left(1+\frac{p}{r}\right)(P_1\cos L - P_2\sin L)a_{d\theta} + (Q_1\cos L - Q_2\sin L)a_{dh}\right\} \tag{10.53}$$

where

$$b = a\sqrt{1-P_1^2-P_2^2} \qquad h = nab$$

$$\frac{p}{r} = 1 + P_1\sin L + P_2\cos L \qquad \frac{r}{h} = \frac{h}{\mu(1+P_1\sin L + P_2\cos L)}$$

4. The true longitude L is obtained from the mean longitude l by first solving Kepler's equation

$$l = K + P_1\cos K - P_2\sin K$$

for the eccentric longitude K and determining r from the equation of orbit

$$r = a(1 - P_1\sin K - P_2\cos K)$$

Then, L is calculated from the eccentric longitude according to the easily derived relations

$$\sin L = \frac{a}{r}\left[\left(1 - \frac{a}{a+b}P_2^2\right)\sin K + \frac{a}{a+b}P_1P_2\cos K - P_1\right]$$

$$\cos L = \frac{a}{r}\left[\left(1 - \frac{a}{a+b}P_1^2\right)\cos K + \frac{a}{a+b}P_1P_2\sin K - P_2\right]$$

where

$$\frac{a}{a+b} = \frac{1}{1+\sqrt{1-e^2}} = \frac{1}{1+\sqrt{1-P_1^2-P_2^2}}$$

or alternately expressed as

$$\frac{a}{a+b} = \frac{\beta}{e}$$

in terms of the parameter β defined in Prob. 4–7.

Verification of the validity of these equations is left as an exercise for the reader.

Problem 10–10

The *equinoctial coordinate axes* are defined with respect to the reference axes as follows:

(1) a positive rotation about the vector $\mathbf{i}_z$ through an angle Ω to establish the direction of the ascending node $\mathbf{i}_n$,

(2) a positive rotation about the vector $\mathbf{i}_n$ through an angle i to establish the direction of $\mathbf{i}_h$, and

(3) a negative rotation about the vector $\mathbf{i}_h$ through an angle Ω.

The position and velocity vectors are expressed in components along the equinoctial axes as

$$\mathbf{r} = r\begin{bmatrix} \cos L \\ \sin L \\ 0 \end{bmatrix} \qquad \text{and} \qquad \mathbf{v} = \frac{h}{p}\begin{bmatrix} -P_1 - \sin L \\ P_2 + \cos L \\ 0 \end{bmatrix}$$

and the rotation matrix, to transform from equinoctial coordinates to reference coordinates, in terms of the equinoctial elements is

$$\mathbf{R} = \frac{1}{1+Q_1^2+Q_2^2}\begin{bmatrix} 1-Q_1^2+Q_2^2 & 2Q_1Q_2 & 2Q_1 \\ 2Q_1Q_2 & 1+Q_1^2-Q_2^2 & -2Q_2 \\ -2Q_1 & 2Q_2 & 1-Q_1^2-Q_2^2 \end{bmatrix}$$

◇ **Problem 10–11**

The equations of motion for the constant radial thrust problem of Sect. 8.8 can be written as the following set of nonsingular variational equations:

$$\frac{dP_1}{dt} = -\frac{h}{\mu}\cos\theta\, a_{T_r}$$

$$\frac{dP_2}{dt} = \frac{h}{\mu}\sin\theta\, a_{T_r}$$

$$\frac{da}{dt} = \frac{2a^2}{h}(P_2\sin\theta - P_1\cos\theta)a_{T_r}$$

$$\frac{d\theta}{dt} = \frac{\mu^2}{h^3}(1 + P_1\sin\theta + P_2\cos\theta)^2$$

where

$$h = \sqrt{\mu a(1 - P_1^2 - P_2^2)}$$

with the initial conditions at $t = t_0$ obtained from

$$P_1 = P_2 = \theta = 0$$

and

$$a = r_0$$

NO

10.5 The Poisson Matrix and Vector Variations

The Lagrange matrix $\mathbf{L}$ defined in Eq. (10.19) can be written in a more compact form which renders obvious the proper expression for its inverse. Define a six-dimensional state vector $\mathbf{s}$ whose partitions are the position and velocity vectors $\mathbf{r}$ and $\mathbf{v}$:

$$\mathbf{s} = \begin{bmatrix} \mathbf{r} \\ \mathbf{v} \end{bmatrix} \tag{10.54}$$

Then, we can readily show that

$$\mathbf{L} = \left[\frac{\partial \mathbf{s}}{\partial \boldsymbol{\alpha}}\right]^{\mathrm{T}} \mathbf{J} \frac{\partial \mathbf{s}}{\partial \boldsymbol{\alpha}} \tag{10.55}$$

where the 6×6 matrix $\mathbf{J}$, introduced in Sect. 9.5, is defined by

$$\mathbf{J} = \begin{bmatrix} \mathbf{O} & \mathbf{I} \\ -\mathbf{I} & \mathbf{O} \end{bmatrix} \tag{10.56}$$

The vector $\mathbf{s}$ is, of course, a function of both the time t and the orbital elements $\boldsymbol{\alpha}$. Therefore,

$$\mathbf{s} = \mathbf{s}(t, \boldsymbol{\alpha}) \tag{10.57}$$

may be thought of as a transformation from element space to state space. The matrix $\partial \mathbf{s}/\partial \boldsymbol{\alpha}$ is called the *Jacobian matrix* of the transformation.

The inverse transformation

$$\boldsymbol{\alpha} = \boldsymbol{\alpha}(t, \mathbf{s}) \tag{10.58}$$

certainly exists and the associated Jacobian matrix $\partial \boldsymbol{\alpha}/\partial \mathbf{s}$ is the inverse of $\partial \mathbf{s}/\partial \boldsymbol{\alpha}$. Indeed, the matrix $\partial \boldsymbol{\alpha}/\partial \mathbf{s}$ is frequently called the *matrizant* of the two-body problem. Therefore,

$$\frac{\partial \mathbf{s}}{\partial \boldsymbol{\alpha}} \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{s}} = \mathbf{I} \qquad \text{and} \qquad \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{s}} \frac{\partial \mathbf{s}}{\partial \boldsymbol{\alpha}} = \mathbf{I}$$

Since $\mathbf{J}^2 = -\mathbf{I}$, it is now trivial to construct the inverse of the Lagrange matrix from the form given in Eq. (10.55).

The Poisson Matrix

The matrix

$$\mathbf{P} = \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{s}} \mathbf{J} \left[\frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{s}}\right]^{\mathrm{T}} \tag{10.59}$$

is called the *Poisson matrix*.† Clearly, we have

$$\mathbf{LP} = \mathbf{PL} = -\mathbf{I}$$

† Siméon-Denis Poisson (1781–1840) was one of the greatest of the nineteenth century analysts and mathematical physicists. Although he was urged by his father to study medicine, he entered the Ecole Polytechnique first as a student and then as a professor of mathematics. He was one of the founders of the mathematical theory of elasticity and a major contributor to the theories of heat conduction and water waves.

so that

$$\mathbf{P} = -\mathbf{L}^{-1}$$

Since $\mathbf{L}$ is skew-symmetric, so also is $\mathbf{P}$ and, further, the transpose of the Poisson matrix is the inverse of the Lagrange matrix

$$\mathbf{P}^{\mathrm{T}} = \mathbf{L}^{-1} \tag{10.60}$$

The element in the i^{th} row and j^{th} column of the matrix $\mathbf{P}$ is denoted by (α_i, α_j) and called a *Poisson bracket*. Now, from the expanded form of Eq. (10.59)

$$\mathbf{P} = \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{r}} \left[\frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{v}}\right]^{\mathrm{T}} - \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{v}} \left[\frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{r}}\right]^{\mathrm{T}} \tag{10.61}$$

it follows that the Poisson brackets are obtained from

$$(\alpha_i, \alpha_j) = \frac{\partial \alpha_i}{\partial \mathbf{r}} \left[\frac{\partial \alpha_j}{\partial \mathbf{v}}\right]^{\mathrm{T}} - \frac{\partial \alpha_i}{\partial \mathbf{v}} \left[\frac{\partial \alpha_j}{\partial \mathbf{r}}\right]^{\mathrm{T}} \tag{10.62}$$

Furthermore, they have properties identical to those demonstrated for the Lagrangian brackets in Sect. 10.2.

Using the Poisson matrix, we may write Lagrange's variational equation (10.18) as

$$\frac{d\boldsymbol{\alpha}}{dt} = \mathbf{P}^{\mathrm{T}} \left[\frac{\partial R}{\partial \boldsymbol{\alpha}}\right]^{\mathrm{T}} \tag{10.63}$$

in terms of the disturbing function R or as

$$\frac{d\boldsymbol{\alpha}}{dt} = \mathbf{P}^{\mathrm{T}} \left[\frac{\partial \mathbf{r}}{\partial \boldsymbol{\alpha}}\right]^{\mathrm{T}} \mathbf{a}_d \tag{10.64}$$

in terms of the disturbing acceleration vector $\mathbf{a}_d$. Now, substitute for $\mathbf{P}^{\mathrm{T}}$ from Eq. (10.61), and in the first case,

$$\frac{d\boldsymbol{\alpha}}{dt} = \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{v}} \left[\frac{\partial R}{\partial \mathbf{r}}\right]^{\mathrm{T}} - \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{r}} \left[\frac{\partial R}{\partial \mathbf{v}}\right]^{\mathrm{T}}$$

But R is a function only of position, so that the result is simply

$$\frac{d\boldsymbol{\alpha}}{dt} = \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{v}} \left[\frac{\partial R}{\partial \mathbf{r}}\right]^{\mathrm{T}} \tag{10.65}$$

In the second case,

$$\frac{d\boldsymbol{\alpha}}{dt} = \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{v}} \left[\frac{\partial \mathbf{r}}{\partial \mathbf{r}}\right]^{\mathrm{T}} \mathbf{a}_d - \frac{\partial \boldsymbol{\alpha}}{\partial \mathbf{r}} \left[\frac{\partial \mathbf{r}}{\partial \mathbf{v}}\right]^{\mathrm{T}} \mathbf{a}_d$$

and since the state-vector components are to be regarded as independent variables, then

$$\frac{\partial \mathbf{r}}{\partial \mathbf{r}} = \mathbf{I} \qquad \text{and} \qquad \frac{\partial \mathbf{r}}{\partial \mathbf{v}} = \mathbf{O}$$

Hence, the result

$$\frac{d\alpha}{dt} = \frac{\partial \alpha}{\partial \mathbf{v}} \mathbf{a}_d \tag{10.66}$$

This last equation is particularly useful in that it provides a direct method for determining variational equations of vector orbital elements as well as scalar elements in vector form—as such they will be independent of the coordinate system in which the components of the disturbing acceleration vector $\mathbf{a}_d$ might be expressed.

Variation of the Semimajor Axis

We begin with the energy or vis-viva integral, defined in Eq. (3.17), which is written as

$$\mu\left(\frac{2}{r} - \frac{1}{a}\right) = v^2 = \mathbf{v}\cdot\mathbf{v} = \mathbf{v}^{\mathrm{T}}\mathbf{v}$$

Then, we calculate the partial derivative with respect to the vector $\mathbf{v}$ and obtain

$$\frac{\mu}{a^2}\frac{\partial a}{\partial \mathbf{v}} = 2\mathbf{v}^{\mathrm{T}}$$

According to Eq. (10.66), we have

$$\frac{da}{dt} = \frac{\partial a}{\partial \mathbf{v}} \mathbf{a}_d$$

so that the variational equation for the semimajor axis a is simply

$$\frac{da}{dt} = \frac{2a^2}{\mu}\mathbf{v}\cdot\mathbf{a}_d \tag{10.67}$$

Variation of the Angular Momentum Vector

According to Eq. (3.12), the angular momentum vector is defined as

$$\mathbf{h} = \mathbf{r} \times \mathbf{v}$$

Paralleling the arguments used in Sect. 2.2, we replace the vector product by the matrix-vector product

$$\mathbf{h} = \mathbf{S}_r \mathbf{v} \tag{10.68}$$

where the skew-symmetric matrix $\mathbf{S}_r$ is defined as

$$\mathbf{S}_r = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix}$$

Then, we calculate

$$\frac{\partial \mathbf{h}}{\partial \mathbf{v}} = \mathbf{S}_r \frac{\partial \mathbf{v}}{\partial \mathbf{v}} = \mathbf{S}_r \mathbf{I} = \mathbf{S}_r$$

so that Eq. (10.66) gives

$$\frac{d\mathbf{h}}{dt} = \frac{\partial \mathbf{h}}{\partial \mathbf{v}}\mathbf{a}_d = \mathbf{S}_r\mathbf{a}_d$$

Thus, the variational equation for the vector angular momentum is

$$\frac{d\mathbf{h}}{dt} = \mathbf{r} \times \mathbf{a}_d \tag{10.69}$$

There are two possible vector forms for the variation of the scalar angular momentum h. On the one hand, if we write

$$h^2 = \mathbf{h} \cdot \mathbf{h} = \mathbf{h}^{\mathrm{T}}\mathbf{h}$$

then we have

$$2h\frac{\partial h}{\partial \mathbf{v}} = 2\mathbf{h}^{\mathrm{T}}\frac{\partial \mathbf{h}}{\partial \mathbf{v}} = 2\mathbf{h}^{\mathrm{T}}\mathbf{S}_r$$

so that

$$\frac{dh}{dt} = \mathbf{i}_h \cdot \mathbf{r} \times \mathbf{a}_d = \mathbf{i}_h \times \mathbf{r} \cdot \mathbf{a}_d \tag{10.70}$$

or, alternately,

$$\frac{dh}{dt} = r\,\mathbf{i}_\theta \cdot \mathbf{a}_d \tag{10.71}$$

On the other hand,

$$\begin{aligned} h^2 &= (\mathbf{r} \times \mathbf{v}) \cdot (\mathbf{r} \times \mathbf{v}) \\ &= (\mathbf{r} \cdot \mathbf{r})(\mathbf{v} \cdot \mathbf{v}) - (\mathbf{r} \cdot \mathbf{v})(\mathbf{r} \cdot \mathbf{v}) \\ &= \mathbf{r}^{\mathrm{T}}\mathbf{r}\mathbf{v}^{\mathrm{T}}\mathbf{v} - \mathbf{r}^{\mathrm{T}}\mathbf{v}\mathbf{r}^{\mathrm{T}}\mathbf{v} \end{aligned}$$

so that

$$2h\frac{\partial h}{\partial \mathbf{v}} = 2\mathbf{r}^{\mathrm{T}}\mathbf{r}\mathbf{v}^{\mathrm{T}} - 2\mathbf{r}^{\mathrm{T}}\mathbf{v}\mathbf{r}^{\mathrm{T}}$$

Hence,

$$\frac{dh}{dt} = \frac{1}{h}\mathbf{r}^{\mathrm{T}}(\mathbf{r}\mathbf{v}^{\mathrm{T}} - \mathbf{v}\mathbf{r}^{\mathrm{T}})\mathbf{a}_d \tag{10.72}$$

or, alternately,

$$\frac{dh}{dt} = \frac{1}{h}[r^2(\mathbf{v} \cdot \mathbf{a}_d) - (\mathbf{r} \cdot \mathbf{v})(\mathbf{r} \cdot \mathbf{a}_d)] \tag{10.73}$$

Variation of the Eccentricity Vector

The eccentricity or Laplace vector was defined in Eq. (3.14) and can be written in any of the possible forms

$$\mu \mathbf{e} = \mathbf{v} \times \mathbf{h} - \mu \, \mathbf{i}_r$$
$$= -\mathbf{S}_h \mathbf{v} - \mu \, \mathbf{i}_r = \mathbf{S}_v \mathbf{h} - \mu \, \mathbf{i}_r$$

where the matrices $\mathbf{S}_h$ and $\mathbf{S}_v$ are constructed in the same manner as the matrix $\mathbf{S}_r$ used for the angular momentum derivation. Again we have

$$\mu \frac{\partial \mathbf{e}}{\partial \mathbf{v}} = -\mathbf{S}_h \frac{\partial \mathbf{v}}{\partial \mathbf{v}} + \mathbf{S}_v \frac{\partial \mathbf{h}}{\partial \mathbf{v}}$$

so that

$$\mu \frac{d\mathbf{e}}{dt} = -\mathbf{h} \times \mathbf{a}_d + \mathbf{v} \times \frac{d\mathbf{h}}{dt}$$

Thus, the variation of the eccentricity vector can be expressed in any of the following forms:

$$\mu \frac{d\mathbf{e}}{dt} = \mathbf{a}_d \times (\mathbf{r} \times \mathbf{v}) + (\mathbf{a}_d \times \mathbf{r}) \times \mathbf{v} \tag{10.74}$$

$$\mu \frac{d\mathbf{e}}{dt} = 2(\mathbf{v} \cdot \mathbf{a}_d)\mathbf{r} - (\mathbf{r} \cdot \mathbf{a}_d)\mathbf{v} - (\mathbf{r} \cdot \mathbf{v})\mathbf{a}_d \tag{10.75}$$

$$\mu \frac{d\mathbf{e}}{dt} = (2\mathbf{r}\mathbf{v}^{\mathrm{T}} - \mathbf{v}\mathbf{r}^{\mathrm{T}} - \mathbf{r}^{\mathrm{T}}\mathbf{v}\mathbf{I})\mathbf{a}_d \tag{10.76}$$

By now it should be apparent that each of the variational equations derived thus far can be obtained formally according to the rule:

Apply the usual rules of differentiation to any two-body identity. Treat $\mathbf{r}$ *as constant, orbital elements as variables, and replace the time rate of change of* $\mathbf{v}$ *by* $\mathbf{a}_d$.

This convenient rule has general validity.

For example, to obtain the variational equation for the eccentricity, begin with the expression

$$p = h^2/\mu = a(1 - e^2)$$

defining the parameter. Then,

$$2h\frac{dh}{dt} = \mu(1 - e^2)\frac{da}{dt} - 2\mu a e \frac{de}{dt}$$

and, substituting from Eqs. (10.67) and (10.73), yields

$$\frac{de}{dt} = \frac{1}{\mu a e}[(\mathbf{r} \cdot \mathbf{v})(\mathbf{r} \cdot \mathbf{a}_d) + (pa - r^2)(\mathbf{v} \cdot \mathbf{a}_d)] \tag{10.77}$$

Variation of the Inclination and Longitude of the Node

The angular momentum vector $\mathbf{h}$ is, of course, normal to the plane of the osculating orbit and may be expressed in terms of components along reference axes as

$$\mathbf{h} = h\,\mathbf{i}_h = h(\sin\Omega\sin i\,\mathbf{i}_x - \cos\Omega\sin i\,\mathbf{i}_y + \cos i\,\mathbf{i}_z)$$

since the unit vector $\mathbf{i}_h$ is identical to the vector $\mathbf{i}_\zeta$ of Eq. (2.6). Applying the formal rule, i.e., calculating the ordinary time derivative of this two-body identity, results in

$$\frac{d\mathbf{h}}{dt} = h\sin i\frac{d\Omega}{dt}\,\mathbf{i}_n - h\frac{di}{dt}\,\mathbf{i}_m + \frac{dh}{dt}\,\mathbf{i}_h$$

where $\mathbf{i}_n$ is a unit vector in the direction of the ascending node and $\mathbf{i}_m$ is in the plane of the osculating orbit and normal to $\mathbf{i}_n$ such that $\mathbf{i}_n$, $\mathbf{i}_m$, $\mathbf{i}_h$ form an orthogonal triad. Expressions for $\mathbf{i}_n$ and $\mathbf{i}_m$ in terms of components along the reference axes are given in Eqs. (2.5) and (2.8).

The appropriate variational equations for the longitude of the node Ω and the inclination angle i are obtained by calculating the scalar product of the last equation with $\mathbf{i}_n$ and $\mathbf{i}_m$, respectively. We have

$$\frac{d\Omega}{dt} = \frac{1}{h\sin i}\,\mathbf{i}_n \times \mathbf{r}\cdot\mathbf{a}_d = \frac{r\sin\theta}{h\sin i}\,\mathbf{i}_h\cdot\mathbf{a}_d \tag{10.78}$$

$$\frac{di}{dt} = -\frac{1}{h}\,\mathbf{i}_m \times \mathbf{r}\cdot\mathbf{a}_d = \frac{r\cos\theta}{h}\,\mathbf{i}_h\cdot\mathbf{a}_d \tag{10.79}$$

where $\theta = \omega + f$ is the argument of latitude. Note that a third scalar product with $\mathbf{i}_h$ produces the same variational equation for h as obtained previously in Eq. (10.70).

Variation of the Argument of Pericenter

The argument of latitude θ is defined as the angle between the position vector and the ascending node. Thus, from

$$\mathbf{i}_n = \cos\Omega\,\mathbf{i}_x + \sin\Omega\,\mathbf{i}_y$$

it follows that

$$\cos\theta = \mathbf{i}_n\cdot\mathbf{i}_r = \cos\Omega\,(\mathbf{i}_x\cdot\mathbf{i}_r) + \sin\Omega\,(\mathbf{i}_y\cdot\mathbf{i}_r)$$

Hence,

$$-\sin\theta\,\frac{\partial\theta}{\partial\mathbf{v}} = [-\sin\Omega\,(\mathbf{i}_x\cdot\mathbf{i}_r) + \cos\Omega\,(\mathbf{i}_y\cdot\mathbf{i}_r)]\frac{\partial\Omega}{\partial\mathbf{v}}$$

Next, from the results of Prob. 3–21,

$$\mathbf{i}_x\cdot\mathbf{i}_r = \cos\Omega\cos\theta - \sin\Omega\sin\theta\cos i$$

$$\mathbf{i}_y\cdot\mathbf{i}_r = \sin\Omega\cos\theta + \cos\Omega\sin\theta\cos i$$

so that, after substitution and cancellation, we obtain

$$\frac{\partial \theta}{\partial \mathbf{v}} = -\cos i \, \frac{\partial \Omega}{\partial \mathbf{v}}$$

or

$$\frac{\partial \theta}{\partial \mathbf{v}} \mathbf{a}_d = -\cos i \, \frac{d\Omega}{dt} \tag{10.80}$$

This last expression is the perturbative derivative of θ, i.e., the change in θ due to the change in $\mathbf{i}_n$ from which the angle θ is measured. The total time rate of change of θ is the sum

$$\frac{d\theta}{dt} = \frac{\partial \theta}{\partial t} + \frac{\partial \theta}{\partial \mathbf{v}} \mathbf{a}_d$$

where $\partial\theta/\partial t$ represents the change in θ due to ordinary two-body motion with constant orbital elements as specified by Kepler's second law. Thus,

$$\frac{d\theta}{dt} = \frac{h}{r^2} - \cos i \, \frac{d\Omega}{dt} \tag{10.81}$$

with $\partial\Omega/dt$ obtained from Eq. (10.78).

Finally, since $\theta = \omega + f$, we can use Eq. (10.80) to write the variational equation for the argument of pericenter as

$$\frac{d\omega}{dt} = -\frac{\partial f}{\partial \mathbf{v}} \mathbf{a}_d - \cos i \, \frac{d\Omega}{dt} \tag{10.82}$$

which involves the perturbative derivative of the true anomaly f. This we calculate in the next subsection.

Variations of the Anomalies

By differentiating the equation of orbit

$$r(1 + e\cos f) = \frac{h^2}{\mu}$$

obtain

$$re\sin f \frac{\partial f}{\partial \mathbf{v}} = r\cos f \frac{\partial e}{\partial \mathbf{v}} - \frac{2h}{\mu}\frac{\partial h}{\partial \mathbf{v}} \tag{10.83}$$

Also, from Eq. (3.31), we establish

$$\frac{\mu}{h} re\sin f = \mathbf{r} \cdot \mathbf{v}$$

which, when differentiated, yields

$$re\cos f \frac{\partial f}{\partial \mathbf{v}} = -r\sin f \frac{\partial e}{\partial \mathbf{v}} + \frac{\mathbf{r}\cdot\mathbf{v}}{\mu}\frac{\partial h}{\partial \mathbf{v}} + \frac{h}{\mu}\mathbf{r}^{\mathrm{T}} \tag{10.84}$$

Now multiply Eq. (10.83) by $\sin f$, Eq. (10.84) by $\cos f$, and add the two. After some fairly straightforward reduction, we obtain

$$reh\frac{\partial f}{\partial \mathbf{v}} = p\cos f\,\mathbf{r}^{\mathrm{T}} - (p+r)\sin f\frac{\partial h}{\partial \mathbf{v}}$$

Then, substitution for $\partial h/\partial \mathbf{v}$ produces

$$\frac{\partial f}{\partial \mathbf{v}} = \frac{r}{h^2 e}\left\{\left[\frac{h}{p}(\cos f + e) + \frac{eh}{r}\right]\mathbf{r}^{\mathrm{T}} - (p+r)\sin f\,\mathbf{v}^{\mathrm{T}}\right\} \tag{10.85}$$

as the perturbative derivative of the true anomaly.

This last expression for $\partial f/\partial \mathbf{v}$ can be used in Eq. (10.82) to complete the variational equation for the argument of pericenter ω. It may also be used to obtain the total time rate of change of the true anomaly as

$$\frac{df}{dt} = \frac{h}{r^2} + \frac{\partial f}{\partial \mathbf{v}}\mathbf{a}_d \tag{10.86}$$

Thus far, the formulas in this section are equally valid for hyperbolic as well elliptic osculating orbits. In the remainder of this subsection, we consider the eccentric and mean anomalies which, of course, apply only to the ellipse and leave as an exercise the parallel arguments for the hyperbola.

From the identity relating the eccentric and true anomalies

$$\cos E = \frac{\cos f + e}{1 + e\cos f}$$

we obtain, in the usual manner,

$$b\frac{\partial E}{\partial \mathbf{v}} = r\frac{\partial f}{\partial \mathbf{v}} - \frac{ra}{p}\sin f\frac{\partial e}{\partial \mathbf{v}}$$

which, after substitution and reduction, results in

$$\frac{\partial E}{\partial \mathbf{v}} = \frac{r}{\mu b e}\left[\frac{h}{p}(\cos f + e)\mathbf{r}^{\mathrm{T}} - (r+a)\sin f\,\mathbf{v}^{\mathrm{T}}\right] \tag{10.87}$$

Similarly, from Kepler's equation

$$M = E - e\sin E$$

we obtain

$$\frac{\partial M}{\partial \mathbf{v}} = \frac{r}{a}\frac{\partial E}{\partial \mathbf{v}} - \sin E\frac{\partial e}{\partial \mathbf{v}}$$

or, in reduced form,

$$\frac{\partial M}{\partial \mathbf{v}} = \frac{rb}{ha^2 e}\left[\cos f\,\mathbf{r}^{\mathrm{T}} - \frac{a}{h}(r+p)\sin f\,\mathbf{v}^{\mathrm{T}}\right] \tag{10.88}$$

The total time derivatives of the eccentric and mean anomalies are then

$$\frac{dE}{dt} = \frac{na}{r} + \frac{\partial E}{\partial \mathbf{v}}\mathbf{a}_d$$
$$\frac{dM}{dt} = n + \frac{\partial M}{\partial \mathbf{v}}\mathbf{a}_d \tag{10.89}$$

10.6 Applications of the Variational Method

In this section we consider several interesting and important applications of the concepts thus far developed in this chapter. The first example utilizes the Lagrange planetary equations to study the average effect of the J_2 term in the earth's gravitational potential on the motion of an earth orbiting satellite. The second example is an application of Gauss' form of the variational equations to analyze the effect of atmospheric drag on the orbital elements of a satellite in earth orbit.

Effect of J_2 on Satellite Orbits

The disturbing function associated with the J_2 term in the earth's gravitational field

$$R = -\frac{Gm}{r} J_2 \left(\frac{r_{eq}}{r}\right)^2 P_2(\cos\phi) \tag{10.90}$$

is obtained from Eq. (8.92). The colatitude angle ϕ is related to the orbital elements and calculated from

$$\cos\phi = \mathbf{i}_r \cdot \mathbf{i}_z = \sin(\omega + f)\sin i$$

using the results of Prob. 3–21. Hence, the Legendre polynomial $P_2(\cos\phi)$ is expressed as

$$P_2(\cos\phi) = \tfrac{1}{2}[3\sin^2(\omega + f)\sin^2 i - 1]$$

so that the disturbing function assumes the form

$$R = -\frac{GmJ_2r_{eq}^2}{2p^3}(1 + e\cos f)^3[3\sin^2(\omega + f)\sin^2 i - 1] \tag{10.91}$$

where r has been replaced by the equation of orbit.

The disturbing function can be expanded as a Fourier series in the mean anomaly M using the technique of Sect. 5.3. The constant term in the series is simply the average value of R over one orbit, i.e.,

$$\overline{R} = \frac{1}{2\pi}\int_0^{2\pi} R\,dM$$

Since $dM = n\,dt$ and $r^2 df = h\,dt$, then clearly,

$$\overline{R} = \frac{1}{2\pi}\int_0^{2\pi} \frac{n}{h} R r^2 \, df$$

Substituting from Eq. (10.91) and performing the integration yields

$$\overline{R} = \frac{n^2 J_2 r_{eq}^2}{4(1-e^2)^{\frac{3}{2}}}(2 - 3\sin^2 i) \qquad (10.92)$$

Thus, the average value of the disturbing function depends only on the three orbital elements a, e, and i.

When $\overline{R}$ is used for R in Lagrange's planetary equations (10.31), we have, immediately, expressions for the average rates of change of the satellite orbital elements during a single revolution. For example, since $\overline{R}$ is not a function of Ω, ω, or λ, we see that

$$\overline{\frac{da}{dt}} = 0 \qquad \overline{\frac{de}{dt}} = 0 \qquad (10.93)$$

On the other hand, we obtain for the longitude of the node

$$\overline{\frac{d\Omega}{dt}} = -\frac{3}{2} J_2 \left(\frac{r_{eq}}{p}\right)^2 n \cos i \qquad (10.94)$$

Thus, the plane of the orbit rotates about the earth's polar axis in a direction opposite to that of the motion of the satellite with a mean rate of rotation given by Eq. (10.94). This phenomenon is referred to as the *regression of the node*.

In a similar manner, we obtain for the mean rate of rotation of the line of apsides

$$\overline{\frac{d\omega}{dt}} = \frac{3}{4} J_2 \left(\frac{r_{eq}}{p}\right)^2 n(5\cos^2 i - 1) \qquad (10.95)$$

It is apparent that there exists a *critical inclination angle*

$$i_{crit} = 63^\circ 26'.1$$

such that, if i exceeds i_{crit}, the line of apsides will regress while, if i is smaller than i_{crit}, the apsidal line will advance.

◇ Problem 10–12

For an earth orbiting satellite, show that

$$\overline{\frac{d\Omega}{dt}} = -9.96 \left(\frac{r_{eq}}{a}\right)^{3.5} (1-e^2)^{-2} \cos i \quad \text{degrees/day}$$

$$\overline{\frac{d\omega}{dt}} = 5.0 \left(\frac{r_{eq}}{a}\right)^{3.5} (1-e^2)^{-2} (5\cos^2 i - 1) \quad \text{degrees/day}$$

using appropriate values for the physical data of the earth.