SYMBOLS AND DEFINITIONS

a - semi-major axis of ellipse

b - semi-minor axis of ellipse

e - eccentricity of ellipse $e^2 = 1 - b^2/a^2 = 1 - p/a$

E - eccentric anomaly, auxiliary angle defined by $a - r = a\, e \cos E$

i - inclination of orbit. Angle between orbit plane and reference plane, positive when measured counterclockwise, looking from ascending node toward origin.

M - mean anomaly

P - semi-latus rectum or parameter of ellipse $p = b^2/a = a(1-e^2)$

r - radius from focus to body

τ - Epoch - time of passage of perigee

t - time

u - argument of the latitude. Angle in orbit plane from ascending node to body.

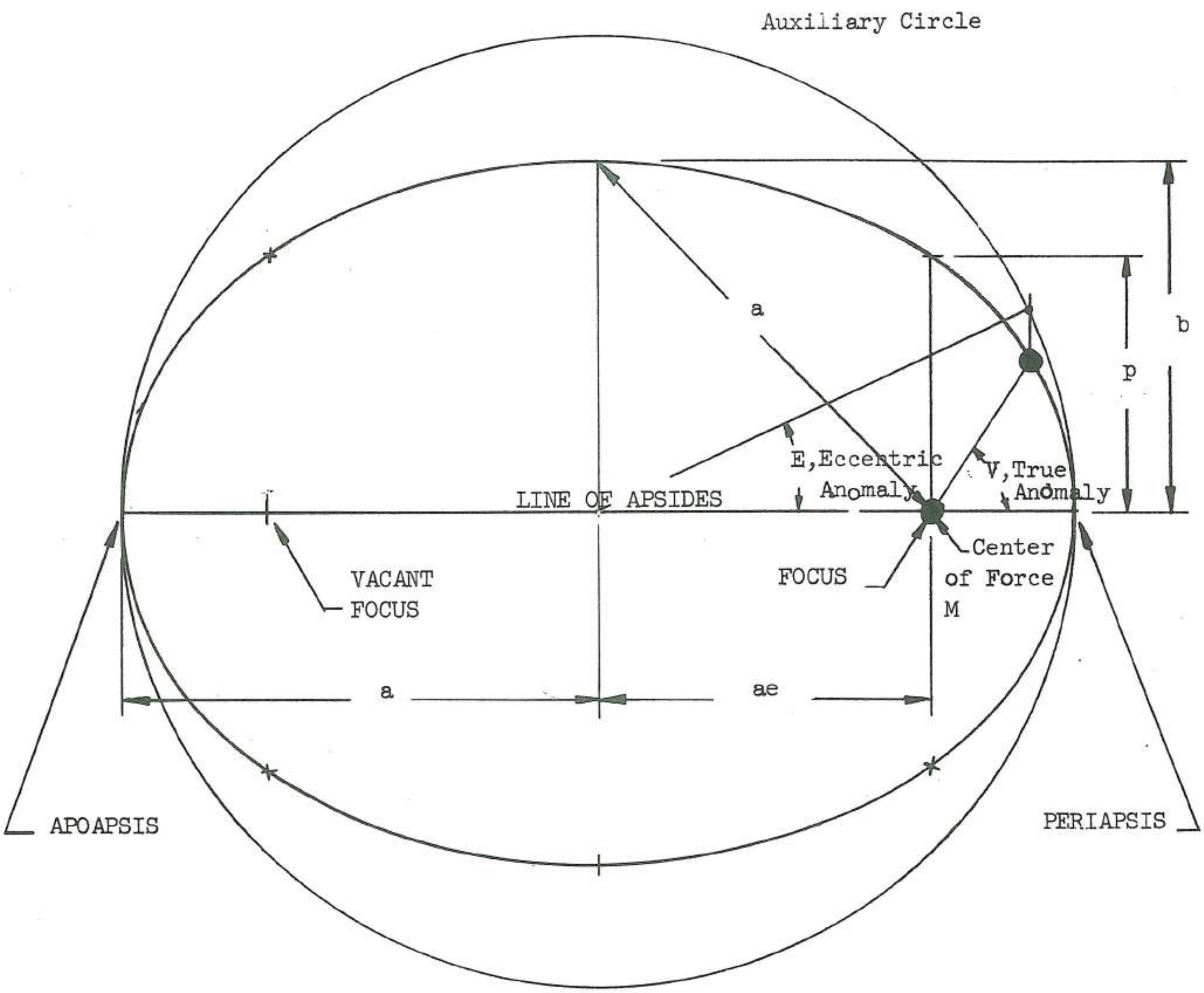
v - true anomaly - angle in orbit plane from perigee to body.

ω - longitude of perigee, angle from ascending node to perigee.

☊ - longitude of the node. Angle from vernal equinox to ascending node.

☋ - angle from vernal equinox to descending node.

♈ - the vernal equinox. (♈ is the symbol for the constellation Aries, the Ram. The vernal equinox is a line toward the First Point of Aries. In heliocentric coordinates, the Earth crosses the extension of the vernal equinox through the Sun at the beginning of Spring. Alternatively, we say that the Sun crosses the vernal equinox at the beginning of Spring.)



a - SEMI-MAJOR AXIS

b - SEMI-MINOR AXIS

p - SEMI-LATUS RECTUM

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

$p = a(2 - e^2) = b^2/a$

Figure 1. Notation for the Ellipse.

INTRODUCTION TO CELESTIAL MECHANICS

(This first section, up to the three body problem, follows very closely the presentation of Professor Wilfred Kaplan at the seminar series, "Trajectory Analysis, Guidance and Control of Space Vehicles," University of Michigan Institute of Science and Technology, 1959.)

We will consider an idealized solar system which consists of n point masses, P_1, P_2, ... P_n of constant mass m_1, m_2, ... m_n attracting each other according to Newton's Law of Gravitation. If $\vec{F}_{12}$ is the force exerted by P_2 on P_1

$$|\vec{F}_{12}| = k^2 \frac{m_1 m_2}{r_{12}^2} \qquad 1.1$$

where k^2 is the constant of gravitation and r_{12} is the distance from P_1 to P_2. Since the force is an attractive force,

$$\vec{F}_{12} = k^2 \frac{m_1 m_2}{r_{12}^2} \frac{\vec{r}_{12}}{r_{12}} \qquad 1.2$$

where $\frac{\vec{r}_{12}}{r_{12}}$ is a unit vector directed from P_1 to P_2.

For a "proper" or "Newtonian" reference frame, Newton's Laws of Motion hold, thus for example,

$$m_1 \vec{a}_1 = \vec{F}_{12} + \vec{F}_{13} + \ldots + \vec{F}_{1n} \qquad 1.3$$

where $\vec{a}_1$ is measured with respect to a proper frame.

In terms of such a coordinate system with origin O, the motion of P_1, ... P_n is governed by the equations

$$m_i \frac{d^2\vec{r}_i}{dt^2} = k^2 \sum{}' \frac{m_i m_j}{r_{ij}^3} (\vec{r}_j - \vec{r}_i) \qquad 1.4$$

where $i = 1,2,\ldots n$

$$\vec{r}_i = \vec{OP}_i$$

and $\sum'$ denotes summation omitting the term for which j = i. For example,

$$m_1 \frac{d^2\vec{r}_1}{dt^2} = k^2 \sum_{j=2}^{n} \frac{m_1 m_j}{r_{1j}^3} (\vec{r}_j - \vec{r}_1) \qquad 1.5$$

This single vector equation represents three equations for the three components of $\vec{r}_1$

$$m_1 \frac{d^2x_1}{dt^2} = k^2 \sum_{j=2}^{n} \frac{m_1 m_j}{r_{1j}^3} (x_j - x_1) \qquad 1.6$$

and similar equations for y_1, z_1, x_2, y_2, z_2, etc.

$$r_{1j} = [(x_1 - x_j)^2 + (y_1 - y_j)^2 + (z_1 - z_j)^2]^{1/2}$$

There are 3n second order equations of motion, hence the n-body problem is of order 6n, non-linear due to the inverse square law of attraction.

If we introduce the potential, U, which we define in this case to be

$$U = -\frac{k^2}{2} \sum_{i=1}^{n}{}' \sum_{j=1}^{n}{}' \frac{m_i m_j}{r_{ij}} \qquad (i \neq j) \qquad 1.7$$

so that $\vec{F}_{ij} = -\,\mathrm{grad}_i U$.

Then 1.6 may be written

$$m_1 \frac{d^2x_1}{dt^2} = -\frac{\partial U}{\partial x_1} \qquad 1.8$$

and the general equations 1.4 may be written:

$$m_i \frac{d^2\vec{r}_i}{dt^2} = -\,\mathrm{grad}_i U \qquad (i=1,\ldots,n) \qquad 1.9$$

where $\text{grad}_i U$ is the vector with components $\partial U/\partial x_i$, $\partial U/\partial y_i$, $\partial U/\partial z_i$. Taking the scalar product of 1.9 with $d\vec{r}_i/dt$ and summing over i

$$\sum_i m_i \frac{d\vec{r}_i}{dt} \cdot \frac{d^2\vec{r}_i}{dt^2} = \sum_i - \frac{d\vec{r}_i}{dt} \cdot \text{grad}_i U \qquad 1.10$$

$$\frac{d}{dt} \frac{1}{2} \sum_i m_i \frac{d\vec{r}_i}{dt} \cdot \frac{d\vec{r}_i}{dt} = - \frac{dU}{dt} \qquad 1.11$$

$$\sum_{i=1}^{n} \frac{1}{2} m_i v_i^2 + U = \text{const} = C_1 \qquad 1.12$$

where

$$v_i^2 = \left(\frac{dx_i}{dt}\right)^2 + \left(\frac{dy_i}{dt}\right)^2 + \left(\frac{dz_i}{dt}\right)^2 \qquad 1.13$$

1.11 is the energy integral--when the motion is due to gravitational forces only, energy is conserved.

If we multiply each of equations 1.9 vectorially by $\vec{r}_i$ and sum over i

$$\sum_i m_i \vec{r}_i \times \frac{d^2\vec{r}_i}{dt^2} = \sum_i - \vec{r}_i \times \text{grad}_i U \qquad 1.14$$

$$\frac{d}{dt} \sum_i m_i \left(r_i \times \frac{d\vec{r}_i}{dt} \right) = 0$$

Since $\sum_i \vec{r}_i \times \text{grad}_i U = 0$ as can be seen from 1.4. Therefore, from 1.14

$$\sum_{i=1}^{n} m_i \vec{r}_i \times \frac{d\vec{r}_i}{dt} = \text{constant vector} \qquad 1.15$$

1.15 is the expression of conservation of angular momentum and is the equivalent of three scalar equations for the conservation of the individual components of the total angular momentum:

$$\sum_{i=1}^{n} m_i \left(y_i \frac{dz_i}{dt} - z_i \frac{dy_i}{dt} \right) = C_2$$

$$\sum_{i=1}^{n} m_i \left(z_i \frac{dx_i}{dt} - x_i \frac{dz_i}{dt} \right) = C_3$$

$$\sum_{i=1}^{n} m_i \left(x_i \frac{dy_i}{dt} - y_i \frac{dx_i}{dt} \right) = C_4 \qquad 1.16$$

Adding equations 1.9, we see that

$$\sum_{i=1}^{n} m_i \frac{d^2\vec{r}_i}{dt^2} = 0 \qquad 1.17$$

(This is apparent from inspection of equations 1.4, for example.) 1.17 can be integrated to give:

$$\sum_{i=1}^{n} m_i \frac{d\vec{r}_i}{dt} = \text{constant vector} \qquad 1.18$$

and in terms of components

$$\sum_{i=1}^{n} m_i \frac{dx_i}{dt} = C_5$$

$$\sum_{i=1}^{n} m_i \frac{dy_i}{dt} = C_6 \qquad 1.19$$

$$\sum_{i=1}^{n} m_i \frac{dz_i}{dt} = C_7 \qquad 1.19\ (\text{cont'})$$

Each of equations 1.19 can be integrated again to give

$$m\bar{x} = \sum_{i=1}^{n} m_i x_i = C_5 t + C_8$$

$$m\bar{y} = \sum_{i=1}^{n} m_i y_i = C_6 t + C_9$$

$$m\bar{z} = \sum_{i=1}^{n} m_i z_i = C_7 t + C_{10} \qquad 1.20$$

Thus we have accomplished 10 of the 6n integrations required to solve completely the n-body problem. The ten integrals tell us that, in this case, energy is conserved, each of the three components of the angular momentum vector is conserved, each of the three components of the total linear momentum vector is conserved, and each of the three components of the radius vector to the center of mass at most increases linearly with time. We have apparently made remarkable progress--only two more integrations are required to solve completely the two-body problem, only eight more would solve the three-body problem.

The two remaining integrals of the two-body problem are elementary. The greatest mathematicians of all time have spent years attempting to solve the three-body problem and have not advanced the solution considerably beyond the ten elementary integrals we have already given.

THE RESTRICTED THREE-BODY PROBLEM

The restricted three-body problem is that of determining the motion of a body of infinitesimal mass under the gravitational influence of two bodies of finite mass. Although this problem has not been solved, some progress has been made toward the solution and some characteristics of the motion are known. Since this problem closely approximates that of determining the motion of a spaceship in Earth-Moon space, for example, we will very briefly review some of the facts which are known about the motion.

REFERENCES: (1) SPACE TECHNOLOGY, CHAPTER VII
(2) MOULTON, CHAPTER VIII
(3) WHITTAKER, CHAPTER XIII

The system to be studied consists of two finite bodies, the Earth and the Moon, for example, which we will assume to move in circular orbits about their common center of mass, and an infinitesimal body subject to their attraction. Following Moulton, we will choose the unit of mass to be the sum of the Earth's mass and that of the Moon. The unit of distance will be the (assumed constant) distance between the Earth and the Moon (239,073.7 statute miles). The coordinate systems used are shown on the previous page.

In the x_o, y_o, z_o inertial frame, equations 1.4 become:

$$\ddot{x}_o = -\frac{K(1-\mu)(x_o-x_{o1})}{r_1^3} - \frac{K\mu(x_o-x_{o2})}{r_2^3}$$

$$\ddot{y}_o = -\frac{K(1-\mu)(y_o-y_{o1})}{r_1^3} - \frac{K\mu(y_o-y_{o2})}{r_2^3}$$

$$\ddot{z}_o = -\frac{K(1-\mu)z_o}{r_1^3} - \frac{K\mu z_o}{r_2^3} \qquad 1.21$$

where x_o, y_o, z_o are the coordinates of the particle; x_{o1}, y_{o1}, are the coordinates of the earth; and x_{o2}, y_{o2} are the coordinates of the moon. x_{o1}, y_{o1}, x_{o2}, y_{o2} are functions of time. K is the product of the gravitational constant and the total mass of the earth and moon, and μ is the relative mass of the moon.

If we transform equations 1.21 to a rotating coordinate system, x, y, z, in which the earth and moon always lie on the x axis and the origin is at the center of mass of the earth and moon (the angular velocity of the rotating coordinate system is ω), then the equations of motion are:

$$\ddot{x} - 2\omega\dot{y} = \omega^2 x - \frac{K(1-\mu)(x-x_1)}{r_1^3} - \frac{K\mu(x-x_2)}{r_2^3} \qquad 1.22$$

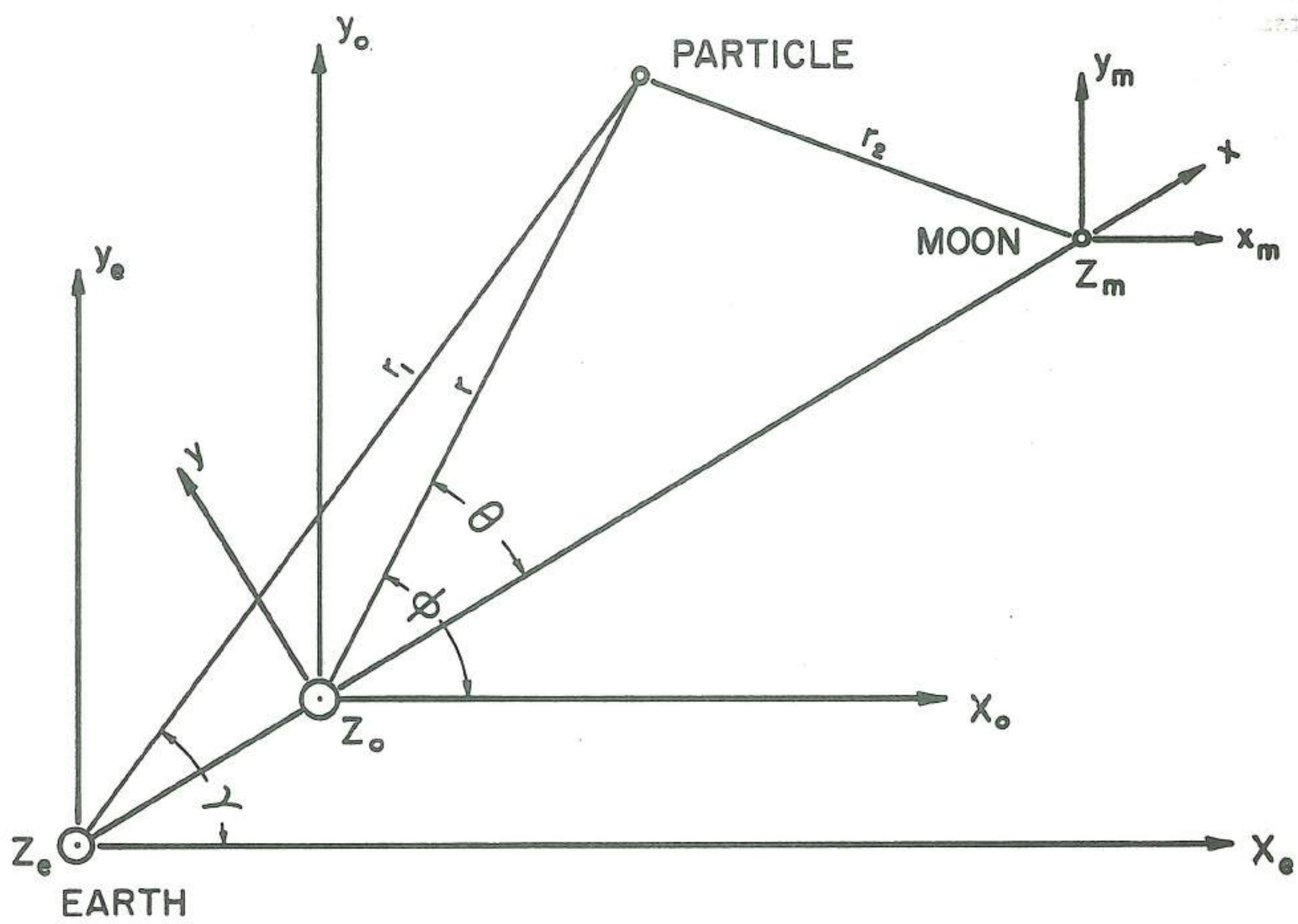


Figure 1.1. Coordinate Systems for Restricted Three Body Problem

$$\ddot{y} + 2\omega\dot{x} = \omega^2 y - \frac{K(1-\mu)y}{r_1^3} - \frac{K\mu y}{r_2^3}$$

$$\ddot{z} = -\frac{K(1-\mu)z}{r_1^3} - \frac{K\mu z}{r_2^3} \qquad 1.22 \text{ (cont')}$$

The earth and moon coordinates x_1 and x_2 are constants in this system.

If we define a function W

$$W(x,y,z) = \frac{1}{2}\omega^2(x^2+y^2) + \frac{K(1-\mu)}{r_1} + \frac{K\mu}{r_2} \qquad 1.23$$

then the equations of motion can be written:

$$\ddot{x} - 2\omega\dot{y} = \frac{\partial W}{\partial x}$$

$$\ddot{y} + 2\omega\dot{x} = \frac{\partial W}{\partial y}$$

$$\ddot{z} = \frac{\partial W}{\partial z} \qquad 1.24$$

Multiplying these by $2\dot{x}$, $2\dot{y}$, $2\dot{z}$, respectively, adding and integrating, we obtain:

$$(\dot{x})^2 + (\dot{y})^2 + (\dot{z})^2 = V^2 = 2W - C \qquad 1.25$$

or:

$$V^2 = \omega^2(x^2+y^2) + \frac{2K(1-\mu)}{r_1} + \frac{2K\mu}{r_2} - C \qquad 1.26$$

1.26 is called Jacobi's Integral and the constant C is the "Jacobian Constant." As reasoned by G. W. Hill, in real motion V^2 must be positive. Thus, for a given C it is possible to draw contours in the x, y plane on which $V = 0$. These contours will be boundaries between regions where V^2 is positive and regions where (formally) V^2 is negative. A real particle cannot enter the region where V^2 is negative, that is, it cannot cross the contour $V^2 = 0$.

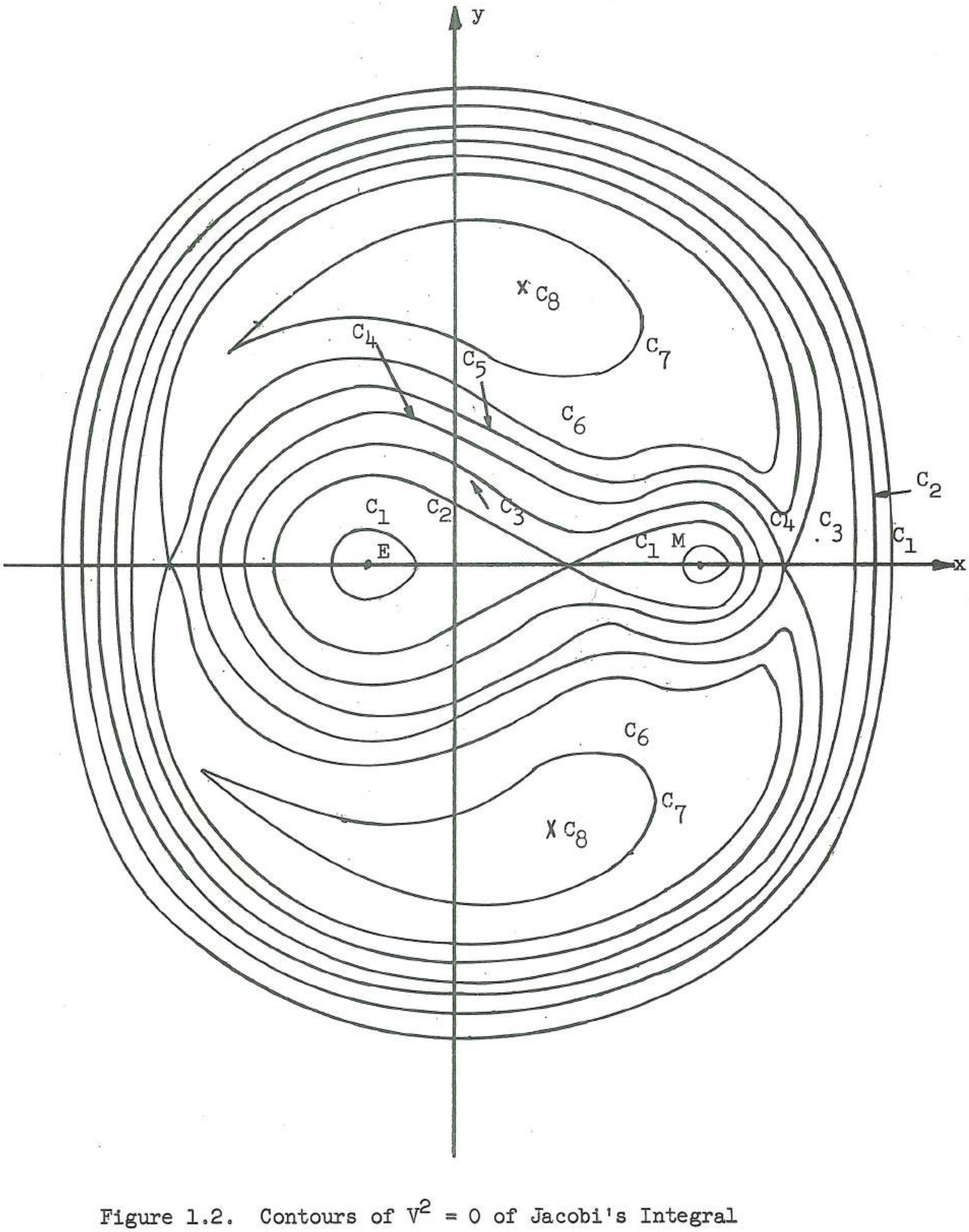


Figure 1.2. Contours of $V^2 = 0$ of Jacobi's Integral

A set of such contours (not to scale) is shown in Figure 1 copied from Figure 7-3, page 7-9, Space Technology. (Note: Chapter VIII of Moulton discusses the three-body problem in some detail. Also Chapter XIII of Whittaker gives a fairly complete account of researches on the three-body problem.)

The C values are in numerical order, so that $C_1 < C_2 < C_3 < C_4$, etc. It is seen that for initial conditions corresponding to C_1 the body can move in a closed path about the earth or the moon but cannot move from earth to moon. For initial conditions corresponding to C between C_2 and C_4 the body can move between the earth and the moon. The C_4 contour is a boundary between contours which enclose the earth and moon and contours which are open behind the moon. Thus, for conditions between C_4 and C_5, for example, the body can escape from the earth-moon system. The two points C_8 are equidistant from the earth and moon and are called libration points. The known analytical solutions of the three-body problem are due to Lagrange, who showed that the equations of motion are satisfied if the body is placed at one of the libration points with the appropriate velocity to remain there. (There are other points, on the x axis, which also fulfill the requirement.) The libration points are of some interest as possible locations for space navigation aids, since for earth-moon space the motion is stable. (If one wishes to use a bell-buoy, it's nice to know where it is.)

It is worthwhile to point out the fallacy of the popular notion that to travel from the earth to the moon one must only project a body toward the moon with velocity sufficient to carry it beyond the "zero gravity" point, the point where earth and moon gravitational fields just balance each other. This point is located about 90% of the distance from the earth to the moon, whereas the contour C_2 (which is the boundary between contours which permit passage to the moon and those which do not) crosses the earth-moon axis about 85% of the distance from earth to moon. The discrepancy is due to the motion of the earth and moon about their common center of mass. For the earth-sun system, the "zero gravity" point is about 165,000 miles from Earth. Our moon, at 240,000 miles from Earth, provides rather strong evidence that the "zero gravity" theory is not very reliable.

Note that for the three-body problem the familiar laws of conservation of energy and of angular momentum do not apply to the bodies individually. Since all three bodies are in motion, they may interchange energy and momentum. Jacobi's integral does furnish a conservation law for the body of negligible mass, however:

$$-C = 2(U - \omega H_z)$$

REFERENCE: Buchheim, R. W., "Motion of a Small Body in Earth-Moon Space," The RAND Corporation, RM 1726.

where U is total energy and H_z is the z component of the angular momentum about the earth-moon axis of rotation.

The three-body problem has been exhaustively investigated by many, including the world's greatest mathematicians. The restricted three-body problem appears to be tantalizingly close to complete solution. Recall that the general three-body problem requires 18 integrals of the equations of motion. Ten of these are furnished by the known ten general integrals, leaving eight to be found. For the restricted three-body problem, only six more are needed. If the motion is confined to a plane, only four more are needed. One of these four is furnished by Jacobi's Integral. Further, Jacobi was able to show how to find the last two integrals if only one more can be found. Thus, only one more integral is needed for the complete solution of the special problem in the plane. However, it was shown by Bruns (in a most remarkably profound, obscure, and complex manner, see Whittaker, "Treatise on the Analytical Dynamics of Particles and Rigid Bodies," Chapter XIV) that no new algebraic integrals exist when rectangular coordinates are used, and Poincaré has proved that when the elements of the orbit are used as variables there are no new uniform transcendental integrals. While these proofs are of great importance, one should not consider that it has been proved that the three-body problem is unsolvable. The simplest of all differential equations, namely the linear differential equation with constant coefficients, in general possesses only transcendental integrals and no algebraic integrals. Before deciding to tackle the problem, however, you are advised to consider that Laplace, Lagrange, Jacobi, Poincaré (and hundreds of others) have shown rather conclusively that it isn't easy.

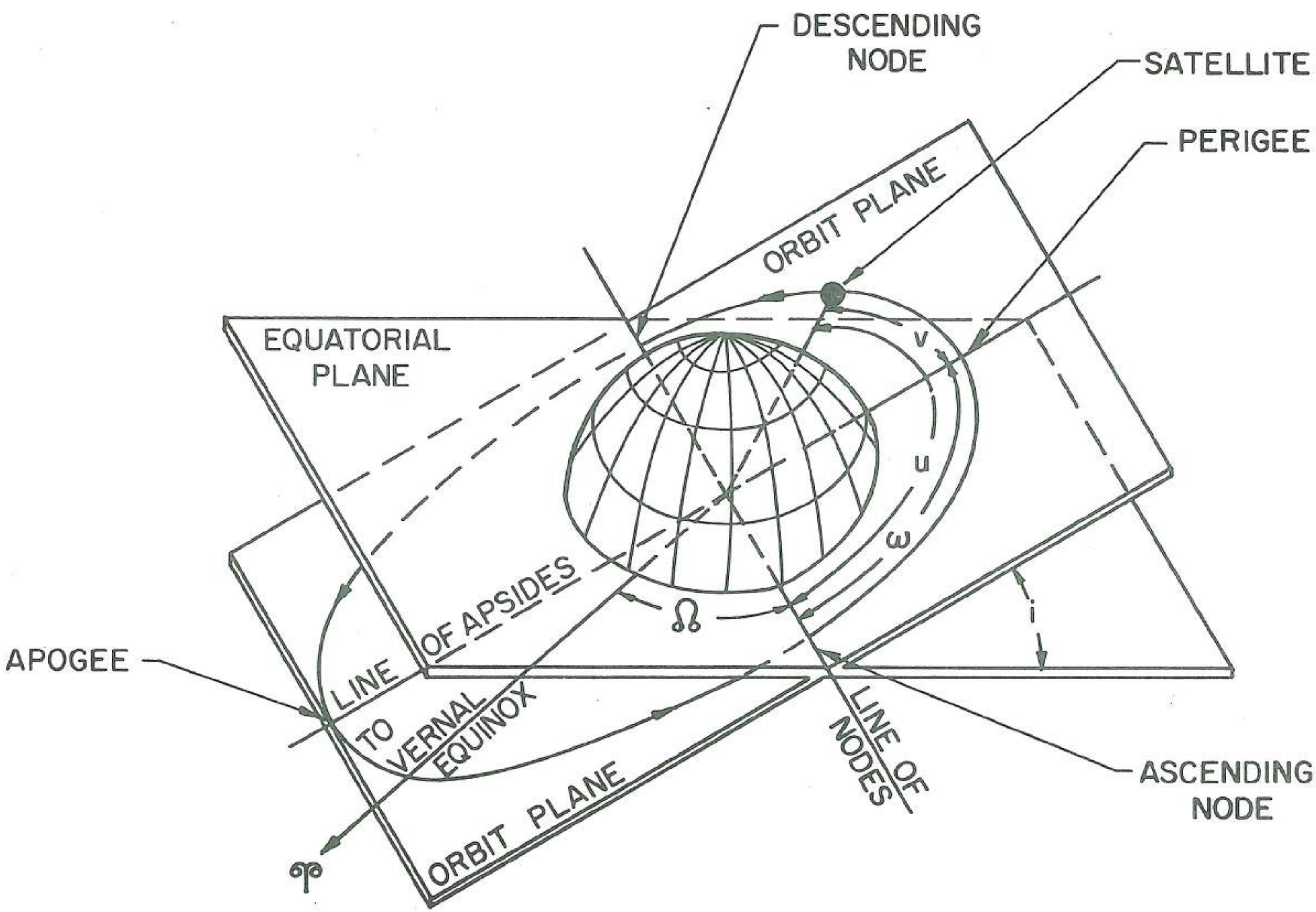


Figure 2. Notation for a Geocentric Orbit

CHAPTER II - THE TWO-BODY PROBLEM

For the two-body problem, equations 1.4 reduce to:

$$m_1 \frac{d^2\vec{r}_1}{dt^2} = k^2 \frac{m_1 m_2}{r_{12}^3} (\vec{r}_2 - \vec{r}_1)$$

2.1

$$m_2 \frac{d^2\vec{r}_2}{dt^2} = k^2 \frac{m_1 m_2}{r_{13}^3} (\vec{r}_1 - \vec{r}_2)$$

If we divide by m_1 and m_2 respectively and subtract the first equation from the second,

$$\frac{d^2}{dt^2} (\vec{r}_2 - \vec{r}_1) = -k^2 \frac{m_1 + m_2}{r_{12}^3} (\vec{r}_2 - \vec{r}_1) \qquad 2.2$$

2.2 is the vector equation of motion of P_2 relative to P_1. Multiplying by m_2, we see that P_2 moves relative to P_1 as though it were attracted by a <u>fixed</u> particle of mass $m_1 + m_2$, located at P_1. P_2 is moving subject to a <u>central force</u> with center at P_1. The problem also may be formulated for motion relative to the center of mass of P_1 and P_2. In this case, P_2 moves as if it were attracted by a mass $\frac{m_1}{(1 + \frac{m_2}{m_1})^2}$ located at the center of mass of P_1 and P_2. In either case, it is readily seen that for a great many problems of interest in astronautics, one of the masses is neglibible when compared with the other and we may consider that the motion of P_2 has no effect on the motion of P_1.

Let $\vec{r} = \vec{r}_2 - \vec{r}_1$

$r = r_{12}$

$\vec{R} = \frac{\vec{r}}{r}$ (R is a unit vector)

$\vec{u} = \frac{d\vec{r}}{dt}$

$u = |\vec{u}|$

Then equation 2.2 may be written

$$\frac{d^2\vec{r}}{dt^2} = \frac{d\vec{u}}{dt} = -k^2 \frac{m_1 + m_2}{r^2} \vec{R} \qquad 2.3$$

Taking the vector product of $\vec{r}$ with 2.3

$$\vec{r} \times \frac{d\vec{u}}{dt} = 0 \qquad 2.4$$

But

$$\vec{r} \times \frac{d\vec{u}}{dt} = \frac{d}{dt} (\vec{r} \times \vec{u}) \qquad 2.5$$

So that

$$\vec{r} \times \vec{u} = \text{const. vector} = \vec{h} \qquad 2.6$$

2.6 is actually the same as 1.14 and expresses the conservation of angular momentum.

Recalling that $\vec{r} \times \vec{u}$ is normal to both $\vec{r}$ and $\vec{u}$,

$$\vec{r} \cdot \vec{h} = \vec{r} \cdot \vec{r} \times \vec{u} = 0 \qquad 2.7$$

Hence the motion remains in a fixed plane normal to $\vec{h}$.

From 2.3 and 2.6,

$$\frac{d\vec{u}}{dt} \times \vec{h} = -k^2 \frac{m_1 + m_2}{r^2} \vec{R} \times (\vec{r} \times \vec{u})$$

$$= -k^2 (m_1 + m_2) \vec{R} \times (\vec{R} \times \frac{\vec{u}}{r}) \qquad 2.8$$

$$= -k^2 (m_1 + m_2) \vec{R} \times (\vec{R} \times \frac{1}{r} \frac{d\vec{r}}{dt})$$

$$= -k^2 (m_1 + m_2) \vec{R} \times (\vec{R} \times \frac{d\vec{R}}{dt})$$

$$\frac{d\vec{u}}{dt} \times \vec{h} = k^2 (m_1 + m_2) \frac{d\vec{R}}{dt}$$

and, since $\vec{h}$ is a constant vector,

$$\frac{d}{dt} (\vec{u} \times \vec{h}) = k^2 (m_1 + m_2) \frac{d}{dt} \vec{R}$$

$$\vec{u} \times \vec{h} = k^2 (m_1 + m_2)(\vec{R} + \vec{e}) \qquad 2.9$$

where $\vec{e}$ is a constant arbitrary vector. Now

$$(\vec{r} \times \vec{u}) \cdot \vec{h} = \vec{h} \cdot (\vec{r} \times \vec{u}) = \vec{r} \cdot (\vec{u} \times \vec{h})$$

Hence, taking the scalar product of 2.9 with $\vec{r}$,

$$(\vec{r} \times \vec{u}) \cdot \vec{h} = k^2 (m_1 + m_2) \vec{r} \cdot (\vec{R} + \vec{e})$$

$$\vec{h} \cdot \vec{h} = h^2 = k^2 (m_1 + m_2)(r + r\, e \cos v) \qquad 2.10$$

where v is the angle between $\vec{r}$ and $\vec{e}$. From 2.10

$$r = \frac{h^2}{k^2 (m_1 + m_2)} \quad \frac{1}{1 + e \cos v} \qquad 2.11$$

If we let

$$\frac{h^2}{k^2 (m_1 + m_2)} = a (1 - e^2) \qquad 2.12$$

$$r = \frac{a (1 - e^2)}{1 + e \cos v} \qquad 2.13$$

2.13 is Kepler's First Law of Planetary Motion. If e is less than 1.0, equation 2.13 is the polar equation of an ellipse with semi-major axis a, eccentricity e. The origin is at one focus of the ellipse and the polar angle "v," called the <u>true anomaly</u> in celestial mechanics, is measured from the point on the ellipse where r is minimum, called the periapsis (perigee if the orbit is geocentric, perihelion if the orbit is heliocentric). 2.13 often is written

$$r = \frac{p}{1 + e \cos v} \qquad 2.14$$

where p is the "parameter" or semi-latus rectum of the conic, $p = a(1 - e^2)$. "e" is called the eccentricity (see Figure 1) and is a measure of the shape of the orbit. (For e = 0, the orbit is a special ellipse, namely, a circle.) When e = 1 the orbit is a parabola, and when $e > 1$ the orbit is a hyperbola.

We have

$$\vec{r} \times \vec{u} = \vec{h} \qquad 2.6$$

where

$$|\vec{h}| = |\vec{r}|\, |\vec{u}| \sin \angle\, \vec{r}, \vec{u}$$

Since

$$|u| \sin \angle\, \vec{r}, \vec{u} = r \frac{dv}{dt}$$

$$r^2 \frac{dv}{dt} = h = 2 \frac{dA}{dt} \qquad 2.15$$

This is Kepler's second law: the radius vector sweeps out area at a constant rate.

Since the radius vector sweeps out the entire area of the ellipse in one period, and the area of an ellipse is $\pi a b$, where b is the semi-minor axis of the ellipse:

$$2 \pi a b = hT \quad 2.16$$

$$T = \frac{2 \pi a b}{h} \quad 2.17$$

but from 2. 12

$$h^2 = a (1-e^2) \left[k^2 (m_1 + m_2)\right] \quad 2.18$$

and from the geometry of the ellipse

$$a (1-e^2) = \frac{b^2}{a} \quad 2.19$$

Therefore:

$$T = \frac{2 \pi a^{3/2}}{k \sqrt{m_1 + m_2}} \quad 2.20$$

2.20 is Kepler's Third Law of Planetary Motion, in corrected form. As originally stated, the law neglected the slight effect of the planetary mass, m_2, on the period. Kepler stated:

"The squares of the periods of the planets vary as the cubes of the semi-major axes."

From 2. 13

$$v = \cos^{-1} \left[\frac{a (1-e^2)}{e r} - \frac{1}{e}\right] \quad 2.21$$

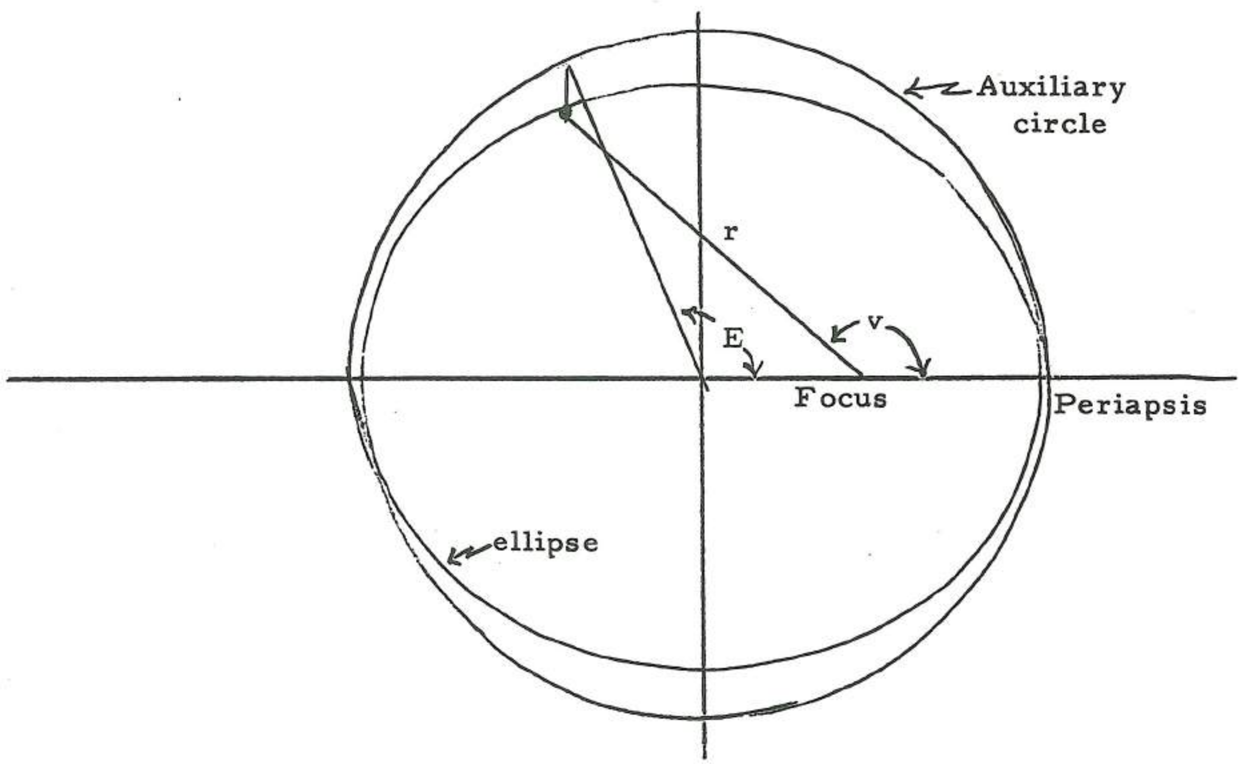
and

$$dv = \frac{a (1-e^2)^{\frac{1}{2}}}{r \sqrt{a^2 e^2 - (a-r)^2}} \; dr \quad 2.22$$

From 2. 15 and 2. 22

$$\frac{a (1-e^2)^{\frac{1}{2}}}{r \sqrt{a^2 e^2 - (a-r)^2}} \; dr = \frac{h \, dt}{r^2} \quad 2.23$$

If we define a new quantity, E (see Figure 3), by



v = true anomaly

E = eccentric anomaly

$$\cos E = \frac{a - r}{a\, e}$$

Fig. 3 The eccentric anomaly.

$$\cos E = \frac{a - r}{a e} \qquad 2.24$$

$$- \sin E \, dE = - \frac{1}{a e} dr$$

but from 2.19

$$\sin E = \frac{\left[a^2 e^2 - (a-r)^2\right]^{\frac{1}{2}}}{a e} \qquad 2.25$$

$$dE = \frac{dr}{\left[a^2 e^2 - (a-r)^2\right]^{\frac{1}{2}}} \qquad 2.26$$

From 2.23:

$$\frac{r}{a} \frac{dr}{\sqrt{a^2 e^2 - (a-r)^2}} = \frac{h \, dt}{a\sqrt{1-e^2}} \qquad 2.27$$

$$(1-e \cos E) \, dE = \frac{h \, dt}{a^2 \sqrt{1-e^2}} \qquad 2.28$$

Which, when integrated gives

$$(E - e \sin E) = n(t - \tau) = M \qquad 2.29$$

where

$$n = \frac{h}{a^2 \sqrt{1-e^2}} = \frac{h}{a b} = \frac{2 \pi}{T} \qquad 2.30$$

τ is a constant

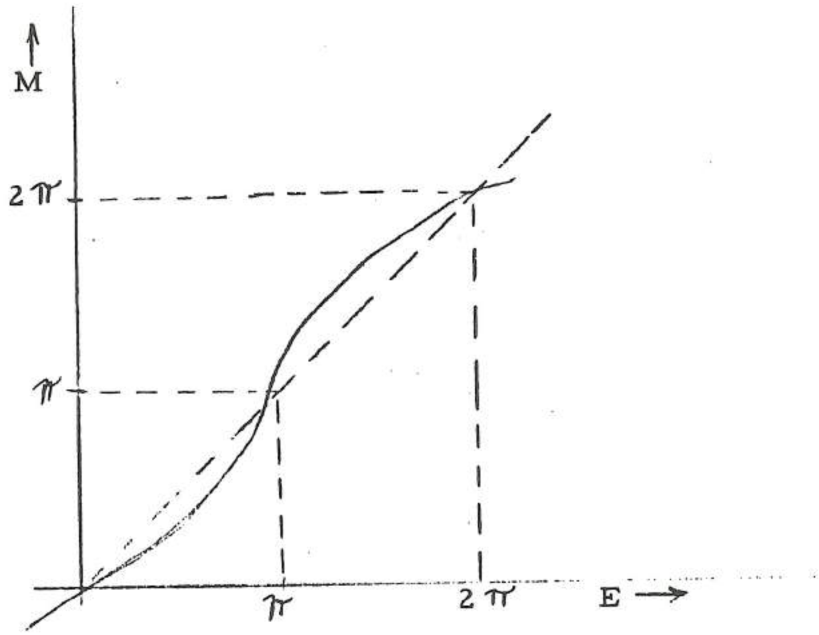
n is called the mean motion--it is the angular rate in a circular orbit of the same period as the actual motion.

E is called the eccentric anomaly. Note that when $E = 0$, $M = 0$, therefore $t = \tau$. τ therefore is the time at perigee, sometimes called the epoch.

M is called the mean anomaly and is the angle which would have been described by the radius vector in time $(t - \tau)$ if the motion had proceeded at the constant rate n.

$$E - e \sin E = M \qquad 2.29$$

is called the Kepler equation. A plot of M vs. E, (for $e \angle 1.0$) is shown on the next page. It is seen that $E - M = e \sin E$ is an odd periodic function of M (hence of t) with period 2 π and so can be expanded in a Fourier sin series.



$$E - M = e \sin E = \sum_{k=1}^{\infty} b_k \sin k M \qquad 2.31$$

where the coefficients b_k are given by

$$b_k = \frac{2}{\pi} \int_0^{\pi} e \sin E \sin k M d M \qquad 2.32$$

$$b_k = \frac{2}{\pi} \int e \sin E \sin k M d M \qquad 2.33$$

Integrating by parts,

$$b_k = \frac{2e}{\pi} \left[\frac{-1}{k} \sin E \cos k M \Big|_0^{\pi} + \frac{1}{k} \int_0^{\pi} \cos k M \cos E d E \right] \qquad 2.34$$

Since $\sin E = 0$ at $M = 0$ and π ,

$$b_k = \frac{e}{\pi k} \int_0^{\pi} 2 \cos k M \cos E d E \qquad 2.35$$

but

$$2 \cos A \cos B = \cos (A + B) + \cos (A-B)$$

$$b_k = \frac{e}{\pi k} \left[\int_0^{\pi} \cos (k M + E) \, dE + \int_0^{\pi} \cos (kM - E) dE \right. \qquad 2.36$$

and since kM = kE - ke sin E

$$b_k = \frac{e}{\pi k} \left\{ \int_0^{\pi} \cos \left[(k + 1) E - k e \sin E \right] dE \right. \qquad 2.37$$
$$\left. + \int_0^{\pi} \cos \left[(k - 1) E - k e \sin E \right] dE \right\}$$

Now the definition of the k^{th} Bessel function of the first kind is:

$$J_k (x) = \frac{1}{\pi} \int_0^{\pi} \cos (k\theta - x \sin \theta) \, d\theta \qquad 2.38$$

hence

$$e \sin E = e \sum_{k=1}^{\infty} \frac{1}{k} \left\{ J_{k+1} (ke) + J_{k-1} (ke) \right\} \sin kM \qquad 2.39$$

or, since

$$n J_n(x) = \frac{x}{2} \left\{ J_{n-1} (x) + J_{n+1} (x) \right\} \qquad 2.40$$

$$e \sin E = \sum_{k=1}^{\infty} \frac{2}{k} J_k (ke) \sin kM \qquad 2.41$$

and

$$E = M + \sum_{k=1}^{\infty} 2 \frac{J_k (ke)}{k} \sin k M \qquad 2.42$$

Thus we finally have "inverted" Kepler's equation and have E - hence r - expressed in terms of M, therefore t. We can obtain similar series for cos E and sin E, and from these can compute rectangular coordinates as functions of time:

$$r \cos v = a (\cos E - e) \qquad 2.43$$
$$r \sin v = a \sqrt{1 - e^2} \sin E$$

Finally, from the above, or from

$$\tan \frac{v}{2} = \sqrt{\frac{1 + e}{1 - e}} \tan \frac{E}{2} \qquad 2.44$$

for example, we can compute v as a function of time. There are other expressions for E as a series in sin kM which are obtainable by expanding the Bessel functions of 2.42 in powers of k. See,for example, Moulton, page 169.

THE VIS-VIVA INTEGRAL

From 2.3 we have:

$$\frac{d^2\vec{r}}{dt^2} = -k^2 \frac{m_1 + m_2}{r^2} \vec{R} \qquad 2.3$$

$$\frac{d\vec{r}}{dt} \cdot \frac{d^2\vec{r}}{dt^2} = -k^2 \frac{m_1 + m_2}{r^2} \frac{d\vec{r}}{dt} \cdot \vec{R} \qquad 2.45$$

but

$$\frac{d\vec{R}}{dt} = \frac{1}{r} \frac{d\vec{r}}{dt} + \vec{r} \frac{d}{dt} \left(\frac{1}{r}\right) \qquad 2.46$$

$$\frac{1}{r^2} \frac{d\vec{r}}{dt} = \frac{1}{r} \frac{d\vec{R}}{dt} - \vec{R} \frac{d}{dt} \left(\frac{1}{r}\right) \qquad 2.47$$

and since $\frac{dR}{dt}$ is normal to $\vec{R}$, 2.45 becomes

$$\frac{d\vec{r}}{dt} \cdot \frac{d^2\vec{r}}{dt^2} = \frac{1}{2} \frac{d}{dt} \left(\frac{d\vec{r}}{dt} \cdot \frac{d\vec{r}}{dt}\right) = k^2 \left(m_1 + m_2\right) \frac{d}{dt} \left(\frac{1}{r}\right) \qquad 2.48$$

$$u^2 = k^2 (m_1 + m_2) \left[\frac{2}{r} + \text{const}\right] \qquad 2.49$$

This is true for any point in the orbit, so must be true, for example, at perigee. Letting the subscript "p" indicate conditions at perigee,

$$u_p^2 = k^2 (m_1 + m_2) \left[\frac{2}{r_p} + \text{const}\right] \qquad 2.50$$

$$\text{const} = \frac{u_p^2}{k^2 (m_1 + m_2)} - \frac{2}{a (1-e)} \qquad 2.51$$

from 2. 15

$$r_p^2 \left(\frac{dv}{dt}\right)_p = h = r_p u_p \tag{2.52}$$

and from 2. 18

$$r_p^2 u_p^2 = h^2 = a (1-e^2) \left[k^2 (m_1 + m_2)\right] \tag{2.53}$$

$$u_p^2 = \frac{a (1-e^2)}{a^2 (1-e)^2} \left[k^2 (m_1 + m_2)\right] \tag{2.54}$$

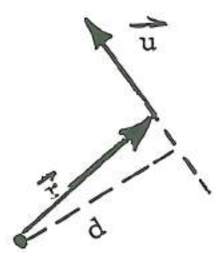
Putting 2.54 in 2.51,

$$\text{const} = \frac{1 + e}{a (1-e)} - \frac{2}{a (1-e)} = - \frac{1}{a} \tag{2.55}$$

2.49 becomes

$$u^2 = k^2 (m_1 + m_2) \left(\frac{2}{r} - \frac{1}{a}\right) \tag{2.56}$$

2.56 is called the "vis viva" integral. (vis-viva-force of motion). One further relationship should be noted: from 2.6 we have $\vec{r} \times \vec{u} = \vec{h}$ (2.6) where $|\vec{h}| = |\vec{r}| |\vec{u}| \sin \measuredangle \vec{r}, \vec{u}$



$|\vec{r}| \sin \measuredangle \vec{r}, \vec{u} = rd$, therefore,

$$u = \frac{h}{d} \tag{2.57}$$

The velocity in the orbit varies inversely with the distance from the center of force to the tangent to the orbit.

CHAPTER III. MECHANICS OF A SYSTEM OF PARTICLES

References: 1. Classical Mechanics, Goldstein, Addison Wesley, 1959.

2. Theoretical Physics, Joos, Hafner, 1935.

- -

Most of us have been taught from the beginnings of our technical education, that Newton's three laws of motion are the most fundamental statement of the laws of motion. Indeed, such a statement is easily defended. However, it can be argued with some justification that Hamilton's Principle is the most fundamental statement of the laws of motion. Quantum mechanics, for example, - in fact most of modern mechanics - developed by way of Hamilton's Principle.

For conservative systems, Hamilton's Principle may be stated as follows: (Ref. 1, p. 30)

"The motion of the system from time t_1 to time t_2 is such that the line integral

$$I = \int_{t_1}^{t_2} L \, dt$$

where $L = T - U$, is an extremum for the path of motion."

Here T is the system kinetic energy and U is the potential. U is a function $U(\dot{q}_j, q_j)$ such that the forces are obtained from U by the prescription

$$Q_j = -\frac{\partial U}{\partial q_j} + \frac{d}{dt}\left(\frac{\partial U}{\partial \dot{q}_j}\right) \qquad 3.1$$

Note that U is a generalization of the function U given in 1.7.

The system can be described in terms of ~~an~~ appropriate consistent set of coordinates and velocities as long as the generalized force can be obtained from 3.1. The product $Q_j q_j$ must have the dimensions of work.

The integral in Hamilton's Principle has just the form which we can handle by variational methods: (See Appendix A)

$$I = \int_{t_1}^{t_2} L\,(q_i, \dot{q}_i, t)\, dt \qquad 3.2$$

and therefore a necessary condition for Hamilton's Principle to hold is

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0 \qquad 3.3$$

These are Lagrange's equations of motion -- obtained immediately from Hamilton's Principle by the variational technique.

Now let us obtain these equations by a different method which clearly shows the fundamental character of Hamilton's Principle.

Consider a system of n particles, a typical particle having mass m_r and coordinates x_{r_i} $(r = 1, 2, \ldots n; i = 1, 2, 3)$. The components of force acting on it are X_{r_i}.

Then the equations of motion, 1.3, for each r and each i are:

$$m_r \ddot{x}_{r_i} = X_{r_i} \qquad 3.4$$

If we multiply these equations by a set of small displacements δx_{r_i}, which are arbitrary functions of time and sum over both r and i, we have a single equation:

$$\sum_{r,i} m_r \ddot{x}_{r_i} \delta x_{r_i} = \sum_{r,i} X_{r_i} \delta x_{r_i} \qquad 3.5$$

This single equation is completely equivalent to the original 3n equations, because, since the δx_{r_i} are arbitrary, we can always equate coefficients of δx_{r_i} on each side of the equation and recover the original 3n equations.

Now integrate between times t_o and t_1 and require that the δx_{r_i} vanish at t_o and t_1

$$\int_{t_o}^{t_1} \sum_{r,i} m_r \ddot{x}_{r_i} \delta x_{r_i} \, dt = \int_{t_o}^{t_1} \sum_{r,i} X_{r_i} \delta x_{r_i} \, dt \qquad 3.6$$

The left-hand side of 3.6 can be integrated by parts:

$$\int_{t_o}^{t_1} \sum_{r,i} m_r \ddot{x}_{r_i} \delta x_{r_i} \, dt = \sum_{r,i} m_r \dot{x}_{r_i} \delta x_{r_i} \Big]_{t_o}^{t_1}$$

$$- \int_{t_o}^{t_1} \sum m_r \dot{x}_{r_i} \frac{d}{dt} \delta x_{r_i} \, dt$$

$$= \sum_{r,i} m_r \dot{x}_{r_i} \delta x_{r_i} \Big]_{t_o}^{t_1} - \int_{t_o}^{t_1} \sum_{r,i} m_r \dot{x}_{r_i} \delta \dot{x}_{r_i} \, dt \qquad 3.7$$

And, therefore, since the δx_{r_i} vanish at t_o and t_1 and $\delta\left(\frac{\dot{x}_{r_i}^2}{2}\right)$ is the same as $\dot{x}_{r_i} \delta \dot{x}_{r_i}$

$$\int_{t_o}^{t_1} \left\{ \delta\left(\tfrac{1}{2} \sum_{r,i} m_r \dot{x}_{r_i}^2\right) + \sum_{r,i} X_{r_i} \delta x_{r_i} \right\} dt = 0 \qquad 3.8$$

Equation 3.8 is the most general form of Hamilton's Principle in classical dynamics. $\sum_{r,i} m_r \frac{\dot{x}_{r_i}^2}{2}$ is the kinetic energy. If the system is conservative, a potential function exists such that

$$- \frac{\partial U}{\partial x_{r_i}} = X_{r_i} \qquad 3.9$$

Then 3.8 can be written as

$$\delta \int_{t_o}^{t_1} (T-U) \, dt = 0 \qquad 3.10$$

Thus, we have started with Newton's Laws of Motion and have derived Hamilton's Principle in a very general form (3.8) and in the more familiar form (3.10) valid for a conservative system. The process is quite reversible and we could as well have started with Hamilton's Principle and derived Newton's Laws. As mentioned earlier, this latter process is today considered to be somewhat more natural than the one we followed.

Generalized Coordinates

In deriving Equation 3.8 we assumed a set of rectangular Cartesian coordinates. (Equation 3.4) Use of Cartesian coordinates may be inconvenient, and certainly Equation 3.10 implies that Hamilton's Principle is valid in any consistent coordinate system.

Let us assume that our 3n Cartesian coordinates are related to 3n generalized coordinates by 3n transformation equations. The generalized coordinates may not have the dimensions of length - for example, in spherical coordinates the two angles are dimensionless. It may even be convenient to use coordinates with dimensions of energy or momentum, and quantities such as the coefficients of a Fourier expansion have been used to good advantage.

The transformation equations are given in the general form:

$$x_{r_i} = x_{r_i}(q_1, q_2, \ldots q_{3n}) \qquad 3.11$$

From 3.11, we can express the velocity as

$$\dot{x}_{r_i} = \sum_j \frac{\partial x_{r_i}}{\partial q_j} \dot{q}_j \qquad 3.12$$

Similarly, the arbitrary displacement δx_{r_i} can be expressed as:

$$\delta \dot{x}_{r_i} = \sum_j \frac{\partial x_{r_i}}{\partial q_j} \delta q_j \qquad 3.13$$

Substituting 3.13 into 3.6

$$\int_{t_o}^{t_1} \left(\sum_{r,i} m_r \ddot{x}_{r_i} \sum_j \frac{\partial x_{r_i}}{\partial q_j} \delta q_j \right) dt =$$

$$\int_{t_o}^{t_1} \left(\sum_{r,i} X_{r_i} \sum_j \frac{\partial x_{r_i}}{\partial q_j} \delta q_j \right) dt \qquad 3.14$$

If we define
$$Q_j = \sum_{r,i} X_{r_i} \frac{\partial x_{r_i}}{\partial q_j} \qquad 3.15$$

Then the right hand side of 3.14 can be written

$$\int_{t_o}^{t_1} \sum_j Q_j \, \delta q_j \qquad 3.16$$

Turning now to the left hand side of 3.14, consider the following relation:

$$\sum_{r,i} m_r \ddot{x}_{r_i} \frac{\partial x_{r_i}}{\partial q_j} = \sum_{r,i} \left\{ \frac{d}{dt} \left(m_r \dot{x}_{r_i} \frac{\partial x_{r_i}}{\partial q_j} \right) - m_r \dot{x}_{r_i} \frac{d}{dt} \frac{\partial x_{r_i}}{\partial q_j} \right\} \qquad 3.17$$

In the last term of 3.17, the order of differentiation can be interchanged. (You can check this from Equation 3.12.)

Further, from Equation 3.12, it can be seen that

$$\frac{\partial \dot{x}_{r_i}}{\partial \dot{q}_j} = \frac{\partial x_{r_i}}{\partial q_j} \qquad 3.18$$

Putting these changes in 3.17

$$\sum_{r,i} m_r \ddot{x}_{r_i} \frac{\partial x_{r_i}}{\partial q_j} = \sum_{r,i} \left\{ \frac{d}{dt} \left(m_r \dot{x}_{r_i} \frac{\partial \dot{x}_{r_i}}{\partial \dot{q}_j} \right) - m_r \dot{x}_{r_i} \frac{\partial \dot{x}_{r_i}}{\partial q_j} \right\} \qquad 3.19$$

$$= \frac{d}{dt} \sum_{r,i} \left\{ \tfrac{1}{2} m_r \frac{\partial}{\partial \dot{q}_j} \dot{x}_{r_i}^2 \right\} - \sum_{r,i} \frac{\partial}{\partial q_j} \tfrac{1}{2} m_r \dot{x}_{r_i}^2$$

Thus Equation 3.14 can be written

$$\int_{t_o}^{t_1} \sum_j \left\{ \frac{d}{dt} \sum_{r,i} \frac{\partial}{\partial \dot{q}_j} \tfrac{1}{2} m_r \dot{x}_{r_i}^2 - \frac{\partial}{\partial q_j} \tfrac{1}{2} m_r \dot{x}_{r_i}^2 \right\} \delta q_j \, dt$$

$$= \int_{t_o}^{t_1} \sum_j Q_j \, \delta q_j \, dt \qquad 3.20$$

If we identify $\sum_{r,i} \frac{1}{2} m_r \dot{x}_{r_i}^2$ as the system kinetic energy, T, 3.20 can be written

$$\int_{t_o}^{t_1} \sum_j \left(\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} - Q_j \right) \delta q_j \, dt = 0 \qquad 3.21$$

and since the δq_j are arbitrary, we can conclude

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_j} - \frac{\partial T}{\partial q_j} = Q_j \qquad j = 1, 2, 3, \ldots, 3n \qquad 3.22$$

Equation 3.22 is very general since we placed few restrictions on the transformation equations which relate the x_{r_i} and the q_j. It was assumed implicitly that the derivatives $\frac{\partial x_{r_i}}{\partial q_j}$ exist and that the transformation can be inverted. Note that the Q_j need not have the dimensions of force - from 3.14 it can be seen that $Q_j \, \delta q_j$ has the dimensions of work. If, for example, we choose generalized coordinates which have the dimensions of work, then the generalized forces will be dimensionless.

Equations 3.22 are sometimes referred to as Lagrange's equations. Equations 3.21 are a generalized expression for 3.8, Hamilton's Principle, in generalized coordinates. If a generalized potential exists such that

$$Q_j = -\frac{\partial U}{\partial q_j} \qquad 3.23$$

Then 3.22 can be written

$$\frac{d}{dt} \frac{\partial (T-U)}{\partial \dot{q}_j} - \frac{\partial}{\partial q_j} (T-U) = 0 \qquad 3.24$$

We can write 3.24 in the form

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}_j} - \frac{\partial L}{\partial q_j} = 0 \qquad 3.25$$

where L = T - U.

Actually 3.23 is unnecessarily restrictive. As indicated by 3.1, U can be a function of q_j and $\dot{q}_j$; if the generalized forces can be obtained from U by the rule

$$Q_j = -\frac{\partial U}{\partial q_j} + \frac{d}{dt}\frac{\partial U}{\partial \dot{q}_j} \qquad 3.1$$

U is then called a "velocity dependent potential", (Reference 1, Section 1-5). This applies to the very important case of electromagnetic forces on moving charges, as in magnetohydrodynamics, for example.

Equations 3.25 are 3n second order differential equations which describe the motion of the system of n particles. Usually, they are not integrable in terms of known functions, but even when a complete solution can not be obtained, it is usually possible to tell a great deal about the motion of the system.

For example, if the independent variable, t, does not appear explicitly in the Lagrangian, L, then by Equation A.32, a first integral of the equations of motion is furnished immediately.

$$\sum_j \dot{q}_j \frac{\partial L}{\partial \dot{q}_j} - L = \text{const.} \qquad 3.26$$

This quantity will be identified shortly as the Hamiltonian, and 3.26 states that if U is a function of the q_j only, the total energy is a constant of the motion. 3.26 is an exceedingly valuable result, valid even for very complex systems.

As another example, consider the case where the Lagrangian does not depend explicitly on one of the coordinates, say q_r, even though it may depend on the corresponding velocity, $\dot{q}_r$. Then the Lagrange equation of motion,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_r} - \frac{\partial L}{\partial q_r} = 0 \qquad 3.25$$

reduces to

$$\frac{\partial L}{\partial \dot{q}_r} = \text{const} \qquad 3.27$$

$\frac{\partial L}{\partial \dot{q}_r}$ is called the generalized momentum conjugate to q_r, or simply the conjugate or canonical momentum. The coordinate, q_r, in this case is called a cyclic coordinate, or an ignorable coordinate, and accordingly, 3.27 states that the generalized momentum, conjugate to a cyclic coordinate, is conserved.

Hamilton's Equations

Lagrange's equations are 3n second order differential equations. Much knowledge about the system is conveniently obtained by examination of an equivalent system of 6n first order equations. These first order equations can be obtained by a "Legendre transformation" as follows: Consider a function of two variables, $f(x, y)$, then

$$df = \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy \tag{3.28}$$

$$= u\, dx + v\, dy \tag{3.29}$$

We wish to transform from the independent variables x, y to independent variables u, y. Then our function $f(x, y)$ will be transformed into some new function $g(u, y)$. If we let

$$g = f - u x \tag{3.30}$$

$$dg = df - u\, dx - x\, du \tag{3.31}$$

$$= v\, dy - x\, du \tag{3.32}$$

which is in the desired form.

We will illustrate the procedure for the case of motion of a single particle in two dimensions, x and y, then consider the general case.

In this case, the Lagrangian is

$$L = L(x, y, \dot{x}, \dot{y}, t) \tag{3.33}$$

$$dL = \frac{\partial L}{\partial x}\, dx + \frac{\partial L}{\partial y}\, dy + \frac{\partial L}{\partial \dot{x}}\, d\dot{x} + \frac{\partial L}{\partial \dot{y}}\, d\dot{y} + \frac{\partial L}{\partial t}\, dt \tag{3.34}$$

$$= p\, dx + q\, dy + r\, d\dot{x} + s\, d\dot{y} + w\, dt \tag{3.35}$$

We wish to transform from these five independent variables to the variables x, y, r, s, t, where

$$r = \frac{\partial L}{\partial \dot{x}} \qquad s = \frac{\partial L}{\partial \dot{y}} \tag{3.36}$$

It is desired to express the equations of motion in terms of these new variables, which will be recognized as the coordinates and canonical momenta conjugate to the coordinates. (See 3.27) We define a new function, H.

$$H = r\dot{x} + s\dot{y} - L \tag{3.37}$$

$$dH = r\, d\dot{x} + \dot{x}\, dr + s\, d\dot{y} + \dot{y}\, ds - dL \tag{3.38}$$

and from 3.35

$$dH = \dot{x}\, dr + \dot{y}\, ds - p\, dx - q\, dy - w dt \qquad 3.39$$

Thus, H = H (r, s, x, y, t) as required and

$$dH = \frac{\partial H}{\partial r}\, dr + \frac{\partial H}{\partial s}\, ds + \frac{\partial H}{\partial x}\, dx + \frac{\partial H}{\partial y}\, dy + \frac{\partial H}{\partial t}\, dt \qquad 3.40$$

Comparing coefficients in 3.39 and 3.40

$$\frac{\partial H}{\partial r} = \dot{x} \qquad \frac{\partial H}{\partial s} = \dot{y}$$

$$\frac{\partial H}{\partial x} = -p \qquad \frac{\partial H}{\partial y} = -q \qquad 3.41$$

$$\frac{\partial H}{\partial t} = -w$$

Remembering the definitions of p and q from 3.34 and 3.35,

$$p = \frac{\partial L}{\partial x} \qquad q = \frac{\partial L}{\partial y} \qquad w = \frac{\partial L}{\partial t} \qquad 3.42$$

From Lagrange's equations of motion, 3.25,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{\partial L}{\partial x} = p$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{y}} = \frac{\partial L}{\partial y} = q \qquad 3.43$$

Putting 3.36 and 3.43 into 3.41,

$$\frac{\partial H}{\partial r} = \dot{x} \qquad \frac{\partial H}{\partial s} = \dot{y}$$

$$\frac{\partial H}{\partial x} = -\dot{r} \qquad \frac{\partial H}{\partial y} = -\dot{s} \qquad 3.44$$

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}$$

Equations 3.44 together with the definitions 3.36 and 3.37 are the canonical equations of Hamilton for this case.

In general, our mechanics problems are to be formulated using as independent variables the generalized coordinates, q_i, and the generalized momenta,

p_i, where

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad (q_j \, , \, \dot{q}_j) \qquad 3.45$$

Thus we will define a new function, H (p, q, t)

$$H(p, q, t) = \sum_i \dot{q}_i p_i - L(q_i, \dot{q}_i, t) \qquad 3.46$$

Equation 3.46 is, except for a minus sign, similar to equation 3.30, where f and L correspond, p_i and u correspond, and x and q_i correspond.

Considering H as a function of p_i, q_i and t we have

$$dH = \sum_i \frac{\partial H}{\partial p_i} dp_i + \sum_i \frac{\partial H}{\partial q_i} dq_i + \frac{\partial H}{\partial t} dt \qquad 3.47$$

but from 3.46 we have

$$dH = \sum_i \dot{q}_i \, dp_i + \sum_i p_i \, d\dot{q}_i$$

$$- \sum_i \frac{\partial L}{\partial q_i} dq_i - \sum_i \frac{\partial L}{\partial \dot{q}_i} d\dot{q}_i - \frac{\partial L}{\partial t} dt \qquad 3.48$$

Equating 3.47 and 3.48 and collecting coefficients of the differentials,

$$\sum_i \left(\frac{\partial H}{\partial p_i} - \dot{q}_i \right) dp_i + \sum_i \left(\frac{\partial H}{\partial q_i} + \frac{\partial L}{\partial q_i} \right) dq_i$$

$$+ \left(\frac{\partial H}{\partial t} + \frac{\partial L}{\partial t} \right) dt - \sum_i \left(p_i - \frac{\partial L}{\partial \dot{q}_i} \right) d\dot{q}_i = 0 \qquad 3.49$$

by definition of p_i, ($p_i = \frac{\partial L}{\partial \dot{q}_i}$) the last term vanishes. By Lagrange's equations,

$$\frac{\partial L}{\partial q_i} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = \dot{p}_i \qquad 3.50$$

Since the differentials dp_i and dq_i are arbitrary, we can conclude that:

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \qquad \frac{\partial H}{\partial t} = - \frac{\partial L}{\partial t}$$

$$\frac{\partial H}{\partial q_i} = - \dot{p}_i \qquad 3.51$$

Equations 3.51 are known as the canonical equations of Hamilton. They are 6n first order equations in the 6n variables, p_i, q_i, which are the equivalent of the 3n second order Lagrange equations. To obtain them, we

1) Set up the Lagrangian

$$L = T - U = L(q, \dot{q}, t)$$

2) Determine the conjugate momentum for each coordinate. The momentum conjugate to the coordinate q_i is defined as:

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \qquad 3.45$$

3) Using the momenta from (2) set up the Hamiltonian,

$$H = \sum_i \dot{q}_i p_i - L \qquad 3.46$$

4) Using 3.46 in 3.51, the equations of motion follow.

Significance of the Hamiltonian

From equations 3.51 we have

$$\frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}$$

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i \qquad 3.51$$

$$\frac{\partial H}{\partial p_i} = \dot{q}_i$$

but

$$\frac{dH}{dt} = \sum_i \left(\frac{\partial H}{\partial q_i} \dot{q}_i + \frac{\partial H}{\partial p_i} \dot{p}_i \right) + \frac{\partial H}{\partial t} \qquad 3.52$$

and from 3.51 above

$$\frac{dH}{dt} = (-\dot{p}_i\dot{q}_i + \dot{p}_i\dot{q}_i) + \frac{\partial H}{\partial t} \qquad 3.53$$

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}$$

Thus, if the Lagrangian does not contain the time explicitly, neither does the Hamiltonian and in this case the Hamiltonian itself is a constant of the motion.

$$H = \alpha_1 \qquad 3.54$$

If the transformation equations from rectangular to generalized coordinates do not contain the time explicitly, i.e.

$$x_i = f_i\,(q_1, \ldots q_n)$$

$$\dot{x}_i = \sum_i \frac{\partial f_i}{\partial q_j}\;\dot{q}_j$$

then it is readily seen that the kinetic energy, T, is a homogeneous function of the q_i of degree 2.

A homogeneous function of degree n is one for which

$$f(hx_1,\ hx_2) = h^n\, f(x_1,\ x_2) \qquad 3.55$$

For such a function, Euler's theorem on homogeneous functions is

$$\sum_i x_i \frac{\partial f}{\partial x_i} = n\, f(x_i) \qquad 3.56$$

This is easily seen by computing $\frac{\partial f}{\partial h}$, then setting h = 1. Since the kinetic energy is homogeneous of degree 2, in q_i, then from 3.56

$$\sum \dot{q}_i \frac{\partial T}{\partial \dot{q}_i} = 2T \qquad 3.57$$

If the potential is not velocity dependent, $U = U(q_i)$, then

$$\frac{\partial T}{\partial \dot{q}_i} = \frac{\partial L}{\partial \dot{q}_i} = p_i \qquad 3.58$$

Therefore

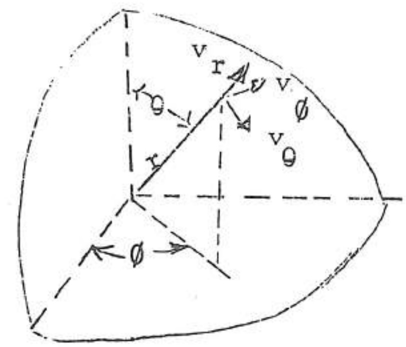
$$H = \sum_i p_i \dot{q}_i - L = 2T - (T-U)$$
$$H = T + U = E \qquad 3.59$$

where E is the total energy. Thus, if the potential is not velocity dependent and the transformation equations do not involve the time explicitly, then

$$H = E = \text{const.} \qquad 3.60$$

Energy is conserved.

Example - Planetary Motion



From the figure, we see that the velocity components in the r, θ, and ϕ directions are.

$$v_r = \dot{r}$$
$$v_\theta = r\,\dot{\theta} \qquad 3.61$$
$$v_\phi = r \sin\theta\,\dot{\phi}$$

We assume an attractive inverse - square central force field. For such a force field, $F = -\frac{k}{r^2}$ and $U = -\frac{k}{r}$. (The negative sign is used for an attractive force.) Then the Lagrangian is

$$L = T - U = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2 + r^2\sin^2\theta\,\dot{\phi}^2) + \frac{k}{r} \qquad 3.62$$

Lagrange's equations of motion are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \qquad 3.25$$

$$m\ddot{r} - mr\dot{\theta}^2 - mr\sin^2\theta\,\dot{\phi}^2 + \frac{k}{r^2} = 0$$

$$mr^2\ddot{\theta} + 2\,mr\dot{r}\dot{\theta} - mr^2\sin\theta\cos\theta\,\dot{\phi}^2 = 0 \qquad 3.63$$

$$mr^2\sin^2\theta\,\dot{\phi} = \text{const.}$$

Equations 3.63 will not be solved now since motion in the central force field has already been studied by another method. Two results - which are obtained without even writing down the equations of motion - are of some interest, however. First, since the independent variable, time, does not occur in the Lagrangian, the Hamiltonian is a constant of the motion (Equations 3.54 and 3.59). Second, since the variable ϕ does not appear in the

Lagrangian, ϕ is a cyclic coordinate and the momentum conjugate to ϕ is a constant of the motion (Equation 3.27).

To continue the example, let us now follow the "recipe" and write down the canonical equations of motion. The Hamiltonian, H, is given by

$$H = \sum_i \dot{q}_i p_i - L \qquad 3.46$$

where

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \qquad 3.45$$

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}$$

$$p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta}$$

$$p_\phi = \frac{\partial L}{\partial \dot{\phi}} = mr^2 \sin^2 \theta \, \dot{\phi} \qquad 3.64$$

$$H = \frac{p_r^2}{m} + \frac{p_\theta^2}{mr^2} + \frac{p_\phi^2}{mr^2 \sin\theta}$$

$$- \left\{ \frac{m}{2} \frac{p_r^2}{m^2} + \frac{mr^2}{2} \frac{p_\theta^2}{m^2 r^4} + \frac{mr^2 \sin^2\theta}{2} \frac{p_\phi^2}{m^2 r^4 \sin^4\theta} + \frac{k}{r} \right\} \qquad 3.65$$

$$H = \frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2 \sin^2\theta} - \frac{k}{r}$$

The equations of motion follow from

$$\frac{\partial H}{\partial p_i} = \dot{q}_i$$

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i \qquad 3.51$$

The six canonical equations of motion are:

$$\frac{p_r}{m} = \dot{r}$$

$$\frac{p_\theta}{mr^2} = \dot{\theta} \qquad 3.66$$

$$\frac{p_\phi}{mr^2\sin^2\theta} = \dot{\phi}$$

$$\frac{p_\theta^2}{mr^3} + \frac{p_\phi^2}{mr^3\sin^2\theta} - \frac{k}{r^2} = \dot{p}_r$$

$$\frac{p_\phi^2 \cos\theta}{mr^2\sin^3\theta} = \dot{p}_\theta \qquad 3.66$$

$$0 = \dot{p}_\phi$$

also $\frac{\partial H}{\partial t} = 0$, $H = \text{const} = \alpha_1$

$$\frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2\sin^2\theta} - \frac{k}{r} = \alpha_1 \qquad 3.67$$

3.67 and the last of equations 3.66 furnish us with two integrals of the equations of motion which may be used to eliminate two of the variables, r, θ, ϕ, p_r, p_θ, p_ϕ.

The Hamilton-Jacobi Theory

The two integrals of the equations of motion given by 3.67 and the last of 3.66 were obtained immediately because of a coordinate missing from the Hamiltonian. (Incidentally, time can be considered a coordinate - in this case the corresponding canonical momentum is -H.) The fact that ϕ was missing from the Hamiltonian was a peculiarity of the spherical coordinate system. Had we used rectangular cartesian coordinates, for example, all of the coordinates would have been present in the Hamiltonian (except time) and none of the conjugate momenta would have been constant. This fact, that by proper selection of the coordinates we can reduce the solution of the canonical equations of motion to the trivial problem of integrating equations of the form

$$\dot{p}_i = 0 \qquad 3.68$$

sets the goal of the Hamilton-Jacobi theory. We will try to find a coordinate system in which all of the canonical equations are of the form 3.68.

We will be interested in a transformation of coordinates from the old variables, p_i, q_i to new variables P_i, Q_i. The equations of motion should be canonical, i.e., Hamilton's Principle should hold when the motion is described in terms of both the old coordinates and the new ones.

Hamilton's Principle is:

$$\delta \int_{t_o}^{t_1} L\, dt = \delta \int_{t_o}^{t_1} \left(\sum_i p_i q_i - H \right) dt = \delta \int_{t_o}^{t_1} \left(\sum P_i Q_i - K \right) dt = 0 \qquad 3.69$$

where the new Hamiltonian is denoted by K.

From 3.69, we cannot conclude that the two integrands are equal, since 3.69 is only a statement about the variation of the two integrals. This is analogous to a statement about differentials. If, for example, we know that

$$dG = dF \qquad 3.70$$

we can only conclude that

$$G = F + \text{const.} \qquad 3.71$$

Thus, we could write 3.69 as

$$\delta \int_{t_o}^{t_1} \left(\sum p_i \dot{q}_i - H \right) dt = \delta \left[\int_{t_o}^{t_1} \left(\sum P_i \dot{Q}_i - K \right) dt + C \right] \qquad 3.72$$

If we write

$$C = \int_{t_o}^{t_1} \frac{dF}{dt}\, dt = F(t_1) - F(t_o) \qquad 3.73$$

where F is an arbitrary function, C is an arbitrary constant, then 3.72 or 3.69 can be written

$$\delta \int \left(\sum p_i \dot{q}_i - H \right) dt = \delta \int_{t_o}^{t_1} \left(\sum P_i \dot{Q}_i - K + \frac{dF}{dt} \right) dt = 0 \qquad 3.74$$

Thus

$$\sum_i p_i \dot{q}_i - H = \sum_i P_i \dot{Q}_i - K + \frac{dF}{dt} \qquad 3.75$$

The arbitrary function, F, is called the generating function of the transformation and could be a function of all of the variables

$$F = F(p_i, q_i, P_i, Q_i, t) \qquad 3.76$$

All of these variables, however, are not independent, since we must relate the P_i and Q_i to the p_i, q_i through a set of transformation equations. Thus, if there are n P_i and n Q_i, there are 2n equations of transformation relating the new coordinates to the old.

$$P_i = P_i(p_i, q_i, t)$$
$$Q_i = Q_i(p_i, q_i, t) \qquad 3.77$$

Using equations 3.77 we can eliminate any 2n of the variables from F and write F in one of the four forms:

$$F = F_1(q, Q, t)$$
$$F = F_2(q, P, t) \qquad 3.78$$
$$F = F_3(p, q, t)$$
$$F = F_4(P, Q, t)$$

Let us suppose that F is expressed as $F_1(q, Q, t)$. Then from 3.75

$$p_i \dot{q}_i - H = P_i \dot{Q}_i - K + \frac{dF_1}{dt}(q, Q, t) \qquad 3.79$$

if we expand $\frac{dF_1}{dt}$, we have

$$\frac{dF_1}{dt} = \sum_i \frac{\partial F_1}{\partial q_i} \dot{q}_i + \sum_i \frac{\partial F_1}{\partial Q_i} \dot{Q}_i + \frac{\partial F}{\partial t} \qquad 3.80$$

Thus, equating coefficients of the q_i and Q_i in 3.79

$$p_i = \frac{\partial F_1}{\partial q_i} \qquad 3.81$$

$$\therefore P_i = -\frac{\partial F_1}{\partial Q_i} \qquad 3.81$$

$$K = H + \frac{\partial F_1}{\partial t}$$

We can - in principle at least - solve the first of these equations for the Q_i in terms of the p_i and q_i,

$$Q_i = Q_i\,(p_i,\ q_i,\ t)$$

Then the second of equations 3.81, after substituting for the Q_i, would furnish the P_i in terms of the p_i and q_i and t. Thus, fixing the function F determines the transformation equations. For this reason F is called the generating function of the transformation.

For our immediate purposes, we wish to obtain a generating function of the form

$$F = F_2\,(q,\ P,\ t)$$

Equations 3.81 suggest that we can obtain a suitable function by means of the Legendre transformation, 3.30. Accordingly, we define

$$F_2\,(q,\ P,\ t) = F_1\,(q,\ Q,\ t) + \sum_i P_i\,Q_i \qquad 3.82$$

consequently,

$$F_1(q_1,\ Q,\ t) = F_2\,(q,\ P,\ t) - \sum_i P_i\,Q_i \qquad 3.83$$

Then equation 3.79 becomes

$$\sum p_i\,\dot{q}_i - H = P_i\,\dot{Q}_i - K + \frac{d}{dt}\left\{F_2 - \sum_i P_i\,Q_1\right\} \qquad 3.84$$

If we repeat the process of differentiating the collecting coefficients, we obtain

$$p_i = \frac{\partial F_2}{\partial q_i}$$

$$Q_i = \frac{\partial F_2}{\partial P_i} \qquad 3.85$$

$$K = H + \frac{\partial F_2}{\partial t}$$

In summary, given a function of the generalized coordinates and conjugate momenta and possibly of the time, such that Hamilton's Principle is satisfied and therefore canonical equations of motion can be obtained:

$$\frac{\partial H}{\partial p_i} = \dot{q}_i \qquad \frac{\partial H}{\partial q_i} = -\dot{p}_i \tag{3.51}$$

If we wish to change variables from p_i, q_i to P_i, Q_i we can select a function $F(q, P, t)$, for example, and let

$$p_i = \frac{\partial F_2}{\partial q_i}$$

$$Q_i = \frac{\partial F_2}{\partial P_i} \tag{3.86}$$

The new Hamiltonian is defined by

$$K = H + \frac{\partial F}{\partial t} \tag{3.87}$$

The new equations of motion are canonical

$$\frac{\partial K}{\partial P_i} = \dot{Q}_i$$

$$\frac{\partial K}{\partial Q_i} = -\dot{P}_i \tag{3.88}$$

and Hamilton's Principle is satisfied using K as the new Hamiltonian and P_i, Q_i as the new momenta and coordinates.

Examples of canonical transformations (Ref: Goldstein, pp. 244, 258)

1) Let $\quad F = F_2 = \sum_i q_i P_i$

then from 3.86

$$p_i = P_i$$

$$Q_i = q_i$$

This is an identity transformation, apparently useless - yet see Example 4.

2) Let $F = F_2 = \sum_i f_i(q_1, \ldots, q_n, t) P_i$

$$Q_i = \frac{\partial F_2}{\partial P_i} = f_i\,(q_1, \ldots, q_n, t)$$

The new coordinates do not involve the conjugate momenta, i.e., this is a transformation of coordinates, called a point transformation.

3) Let $F = F_1 = \sum_i q_i Q_i$

then

$$p_i = \frac{\partial F}{\partial q_i} = Q_i$$

$$P_i = \frac{\partial F}{\partial Q_i} = -q_i$$

This transformation simply interchanges the coordinates and momenta - what we formerly called coordinates are now called momenta and vice versa. We should, therefore, not attach any particular significance to the names "coordinates" and "conjugate momenta". These are simply two traditional designations for family I coordinates and family II coordinates which are connected by the relation

$$I = \frac{\partial L}{\partial \dot{II}}$$

4) Infinitesimal Transformation. The identity transformation was:

$$F = F_2 = \sum_i q_i\, P_i$$

Now let

$$F = F_2 = \sum_i q_i\, P_i + \varepsilon G\,(q_i, P_i)$$

where ε is an infinitesimal.

$$p_i = \frac{\partial F_2}{\partial q_i} = P_i + \varepsilon \frac{\partial G}{\partial q_i}$$

$$P_i - p_i = \delta p_i = -\varepsilon \frac{\partial G}{\partial q_i}$$

$$Q_i = \frac{\partial F_2}{\partial P_i} = q_i + \varepsilon \frac{\partial G}{\partial P_i}$$

$$Q_i - q_i = \delta q_i = \varepsilon \frac{\partial G}{\partial P_i}$$

Since δp_i is an infinitesimal, we can replace P_i by p_i in $\frac{\partial G}{\partial P_i}$ and the error in δq_i will be of second order in $\mathcal{E}$.

$$\delta q_i = \mathcal{E} \frac{\partial G}{\partial p_i} (q_i, p_i)$$

to first order in $\mathcal{E}$. Now let

$$G (q_i P_i) \cong G (q_i, p_i) = H(q_i, p_i)$$

and let $\mathcal{E}$ be a small time increment, $\mathcal{E} = dt$. Then

$$\delta p_i = - dt \frac{\partial H}{\partial q_i} = \dot{p}_i \, dt$$

$$\delta q_i = dt \frac{\partial H}{\partial p_i} = \dot{q}_i \, dt$$

Thus, the Hamiltonian itself may be thought of as the generating function of a transformation, which transforms the coordinates and momenta at time t into the new coordinates and momenta at time t + dt. The motion is the continual time evolution of the canonical transformation generated by the Hamiltonian. Conversely, there must exist a canonical transformation of the coordinates and momenta from any time to any other time. In particular, there must exist a transformation from any time to the initial time t_o. Thus a transformation

$$\begin{aligned} q_i &= q_i (q_{jo}, p_{jo}, t) \\ p_i &= p_i (q_{jo}, p_{jo}, t) \end{aligned} \qquad 3.89$$

exists. This transformation, which specifies coordinates and momenta in terms of initial conditions and time is the description of the motion we seek. This is the solution of the motion problem.

The new variables certainly will be constant in time if we require that the transformed Hamiltonian be identically zero, since then

$$\frac{\partial K}{\partial P_i} = \dot{Q}_i = 0 \qquad \frac{\partial K}{\partial Q_i} = - \dot{P}_i = 0 \qquad 3.90$$

Note that even though only derivatives of K determine $\dot{Q}_i$ and $\dot{P}_i$, the statement $K \equiv 0$ implies much more than K = constant. For example, suppose

$$G = xy$$

then

$$\frac{\partial G}{\partial x} = y \qquad\qquad \frac{\partial G}{\partial y} = x$$

The statement G = constant does not imply $\frac{\partial G}{\partial x} = 0$, but merely $\frac{\partial G}{\partial x}\,dx + \frac{\partial G}{\partial y}\,dy = 0$. Our statement K = 0 means that K is not a function of any of the P_i or Q_i, so that not only is dK = 0 but also $\frac{\partial K}{\partial P_i} = 0$, $\frac{\partial K}{\partial Q_i} = 0$.

Since $K = H + \frac{\partial F}{\partial t}$, K will be zero if F satisfies

$$H + \frac{\partial F}{\partial t} = 0 \qquad 3.91$$

for all values p, q, P, Q, t which occur. We will use $F = F_2(q, P, t)$ and, from 3.86

$$p_i = \frac{\partial F}{\partial q_i} \qquad 3.92$$

Equation 3.91 becomes:

$$H(p_i, q_i, t) + \frac{\partial F}{\partial t} = H\left(\frac{\partial F_2}{\partial q_i}, q_i, t\right) + \frac{\partial F_2}{\partial t} = 0 \qquad 3.93$$

Equation 3.93 is known as the Hamilton-Jacobi equation. It is a partial differential equation, of first order, in n + 1 variables. We should then obtain n + 1 arbitrary constants of integration from its solution. The function, F_2, which satisfies 3.93 is customarily called Hamilton's principal function and is denoted by the letter S. Note that if S is a solution of 3.93, then S + const. also is a solution. Thus, one of the n + 1 constants of integration has no effect on the transformation, which involves only derivatives of S. A complete solution of 3.93 for our purposes can be written in the form

$$S = S(q_1, \ldots, q_n, \alpha_1, \ldots, \alpha_n, t) \qquad 3.94$$

where the α's are independent constants of integration and none of them is simply an additive constant. Since the α's are arbitrary, independent constants, we are free to let them be the new constant momenta, P_i, so that S is indeed of the proper form

$$S = S(q_i, P_i, t) \qquad 3.95$$

After finding the principal function, S, we can find the transformation equations

$$p_i = \frac{\partial S}{\partial q_i}(q_i, \alpha_i, t) \qquad 3.96$$

$$Q_i = \text{const.} = \beta_i = \frac{\partial S}{\partial \alpha_i}(q_i, \alpha_i, t)$$

By setting $t = t_o$, $p_i = p_{io}$, $q_i = q_{io}$ in equations 3.96, the α_i and β_i can be evaluated in terms of the initial conditions. Then the second of equations 3.96 can be (in principle) solved for the q_i in terms of the α_i, β_i and t. These can then be put in the first of equations 3.96, the final result being:

$$p_i = p_i(\alpha_i, \beta_i, t)$$

$$q_i = q_i(\alpha_i, \beta_i, t) \qquad 3.97$$

The problem is now solved, in principle. We have insisted on hedging by constantly making statements which are true "in principle". As might be expected, much of the time practical difficulties arise in actually carrying out the operations which are possible "in principle". Nevertheless, the method is very powerful, and often can be used as a systematic means of solving variational problems other than the dynamics problems which we consider here.

As a first example of the application of the Hamilton-Jacobi method, consider the one-dimensional motion of a mass restrained by a spring.

$$T = \frac{m\dot{x}^2}{2} \qquad 3.98$$

$$F = -kx \quad \therefore \quad U = \frac{kx^2}{2} \qquad 3.99$$

$$L = T - U = \frac{m\dot{x}^2}{2} - \frac{kx^2}{2} \qquad 3.100$$

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x} \qquad q = x \qquad 3.101$$

$$H = \sum_i p_i \dot{q}_i - L = (p)\left(\frac{p}{m}\right) - \frac{m}{2}\left(\frac{p^2}{m^2}\right) + \frac{kq^2}{2} \qquad 3.102$$

$$H = \frac{p^2}{2m} + \frac{kq^2}{2} \qquad 3.103$$

$$H + \frac{\partial S}{\partial t} = 0 \qquad 3.104$$

$$p = \frac{\partial S}{\partial q} \qquad 3.105$$

$$\frac{1}{2m}\left(\frac{\partial S}{\partial q}\right)^2 + \frac{kq^2}{2} + \frac{\partial S}{\partial t} = 0 \qquad 3.106$$

Because the dependence on t appears only in the last term of 3.106, $\frac{\partial S}{\partial t}$ must be at most a constant,

$$\frac{\partial S}{\partial t} = -\alpha \qquad 3.107$$

Then

$$S = W(q, \alpha) - \alpha t \qquad 3.108$$

3.106 becomes:

$$\frac{1}{2m}\left(\frac{\partial W}{\partial q}\right)^2 + \frac{kq^2}{2} = \alpha \qquad 3.109$$

$$W = \sqrt{mk}\int\sqrt{\frac{2\alpha}{k} - q^2}\quad dq \qquad 3.110$$

$$S = \sqrt{mk}\int\sqrt{\frac{2\alpha}{k} - q^2}\quad dq - \alpha t \qquad 3.111$$

$$Q = \beta = \frac{\partial S}{\partial \alpha} = \sqrt{\frac{m}{k}}\int\frac{dq}{\sqrt{\frac{2\alpha}{k} - q^2}} - t \qquad 3.112$$

$$t + \beta = \sqrt{\frac{m}{k}}\ \text{arc cos}\quad q\sqrt{\frac{k}{2\alpha}} \qquad 3.113$$

Let

$$w = \sqrt{\frac{k}{m}} \qquad 3.114$$

$$q = \sqrt{\frac{2\alpha}{k}}\ \cos w\,(t + \beta)$$

CHAPTER IV - THE CANONIC CONSTANTS FOR THE ELLIPTIC ORBIT

Reference: Celestial Mechanics, Smart, Chapter 9.

The Hamilton-Jacobi theory developed in Chapter III will be applied to the problem of motion in a central force field.

The position of a heavenly body is customarily specified by spherical coordinates which are slightly modified from the spherical coordinates normally used in analytic geometry.

Let the origin of coordinates be at the center of force of the central force field. A reference plane is passed through the center of force. If the axes are heliocentric, the reference plane usually will be the plane of the ecliptic, the plane containing the earth's orbit about the sun. If the axes are geocentric, the reference plane usually will be the plane of the equator, i.e. the plane through the center of mass of the earth, normal to the earth's axis of rotation. The reference direction in either case is the vernal equinox, ♈. This line is the line of intersection of the equatorial and ecliptic planes. At one time this line was in the direction of a star in the constellation Aries, the Ram, called the First Point of Aries. Due to precession of the line of apsides of the earth's orbit, the vernal equinox is now shifted and lies in the constellation Pisces. However, it is a well defined direction and the symbol ♈ denotes the vernal equinox even though it is no longer in the Ram. We will specify the position of the planet, or satellite, by the radius from the center of force, r; the angular distance of the satellite from the reference plane measured in the plane normal to the reference plane containing the body and the center of force, called the latitude, L; and the angular distance of this normal plane from the vernal equinox, measured in the reference plane, and called the (ecliptic or equatorial) longitude. See Figure 4-1.

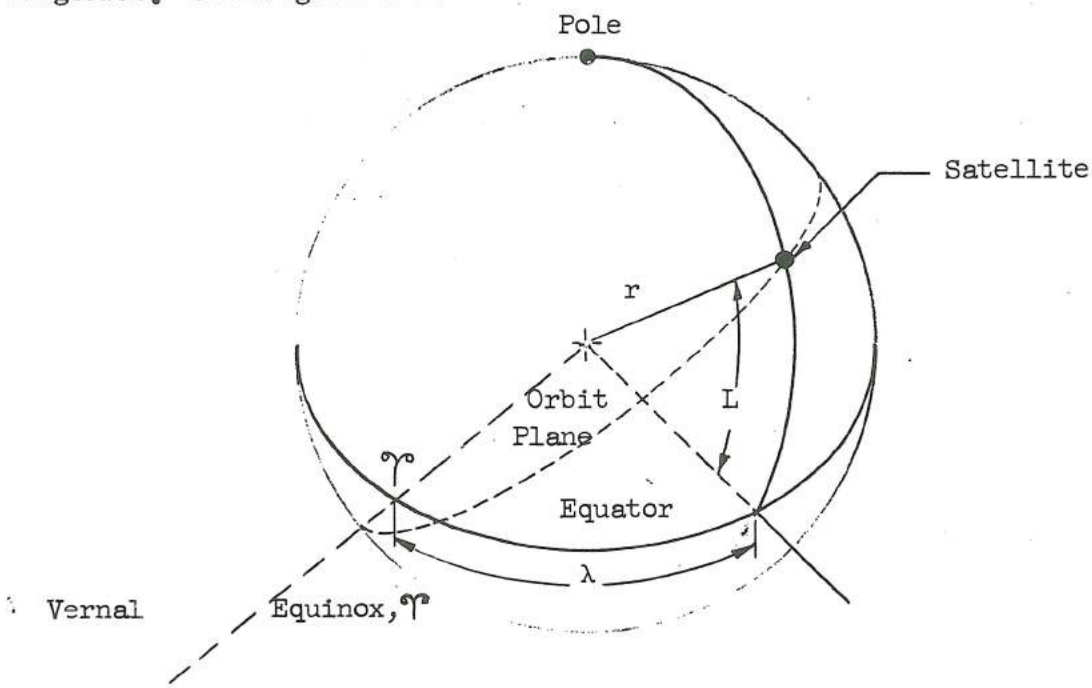


Figure 4-1

If we align the x axis of a Cartesian coordinate system with the vernal equinox and the z axis with the normal to the reference plane, then

$$x = r \cos \lambda \cos L$$

$$y = r \sin \lambda \cos L \qquad 4.1$$

$$z = r \sin L$$

The kinetic energy is

$$T = \tfrac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \qquad 4.2$$

or from 4.1

$$T = \tfrac{1}{2} m (\dot{r}^2 + r^2 \cos^2 L \; \dot{\lambda}^2 + r^2 \dot{L}^2) \qquad 4.3$$

the potential, U is (from 3.62)

$$U = -\frac{k}{r} \qquad 4.4$$

The Lagrangian $\mathcal{L}$, is

$$\mathcal{L} = T - U = \tfrac{1}{2} m(\dot{r}^2 + r^2 \cos^2 L \; \dot{\lambda}^2 + r^2 \dot{L}^2) + \frac{k}{r} \qquad 4.5$$

In order to avoid carrying the constant m throughout the development, let us divide 4.5 by m, or alternatively take m to be the unit of mass. Let $\frac{\mu}{r} = \frac{k}{mr}$ be the potential per unit mass.

$$\mathcal{L}' = T' - U' = \tfrac{1}{2} (\dot{r}^2 + r^2 \cos^2 L \; \dot{\lambda}^2 + r^2 \dot{L}^2) + \frac{\mu}{r} \qquad 4.5a$$

The conjugate momenta:

$$\frac{\partial \mathcal{L}'}{\partial \dot{r}} = p_r = \dot{r}$$

$$\frac{\partial \mathcal{L}'}{\partial \dot{\lambda}} = p_\lambda = r^2 \cos^2 L \, \dot{\lambda} \qquad 4.6$$

$$\frac{\partial \mathcal{L}'}{\partial \dot{L}} = p_L = r^2 \dot{L}$$

The Hamiltonian:

$$H = \frac{1}{2}\left(p_r^2 + \frac{p_\lambda^2}{r^2\cos^2 L} + \frac{p_L^2}{r^2}\right) - \frac{\mu}{r} \qquad 4.7$$

Remembering that $p_i = \dfrac{\partial S}{\partial q_i}$, the Hamilton-Jacobi equation is

$$\frac{\partial S}{\partial t} + \frac{1}{2}\left\{\left(\frac{\partial S}{\partial r}\right)^2 + \frac{1}{r^2\cos^2 L}\left(\frac{\partial S}{\partial \lambda}\right)^2 + \frac{1}{r^2}\left(\frac{\partial S}{\partial L}\right)^2\right\} - \frac{\mu}{r} = 0 \qquad 4.8$$

Since neither t nor λ appears explicitly in the Hamiltonian,

$$H = \text{const} = +\ \alpha_1 \qquad \frac{\partial S}{\partial t} = -\ \alpha_1 \qquad 4.9$$

$$\frac{\partial H}{\partial \lambda} = 0 = \dot{p}_\lambda \qquad p_\lambda = \text{const} = \frac{\partial S}{\partial \lambda} = \alpha_3 \qquad 4.10$$

We, therefore, will try a solution of the Hamilton-Jacobi equation of the form:

$$S = -\ \alpha_1 t + \alpha_3 \lambda + S'(r, L) \qquad 4.11$$

Substituting 4.11 in 4.8

$$-\ \alpha_1 + \frac{1}{2}\left\{\left(\frac{\partial S'}{\partial r}\right)^2 + \frac{1}{r^2\cos^2 L}(\alpha)^2 + \frac{1}{r^2}\left(\frac{\partial S'}{\partial L}\right)^2\right\} - \frac{\mu}{r} = 0 \qquad 4.12$$

$$2\,\alpha_1 r^2 - r^2\left(\frac{\partial S'}{\partial r}\right)^2 + 2\mu r = \left(\frac{\partial S'}{\partial L}\right)^2 + \frac{\alpha_3^2}{\cos^2 L} \qquad 4.13$$

The variables are here separated, so we try

$$S' = S_1(r) + S_2(L) \qquad 4.14$$

Each side of 4.13 must then be equal to a constant, which we will call α_2^2

$$2\,\alpha_1 r^2 - r^2\left(\frac{\partial S_1}{\partial r}\right)^2 + 2\mu r - \alpha_2^2 = 0 \qquad 4.15$$

$$\left(\frac{\partial S_2}{\partial L}\right)^2 + \frac{\alpha_3^2}{\cos^2 L} - \alpha_2^2 = 0 \qquad 4.16$$

$$S_1 = \int_{r_1}^{r} \sqrt{2\,\alpha_1 r^2 + 2\mu r - \alpha_2^2}\quad \frac{dr}{r} \qquad 4.17$$

$$S_2 = \int_{0}^{L} \sqrt{\alpha_2^2 - \alpha_3^2/\cos^2 L}\quad dL \qquad 4.18$$

It appears that we have introduced a new constant into our solution, namely r_1 in 4.17. Therefore, this constant must be specified in terms of the three independent constants α_1, α_2 and α_3. We let r_1 be the smaller of the two roots, r_1 and r_2 of

$$2\,\alpha_1 r^2 + 2\mu r - \alpha_2^2 = 0 \qquad 4.19$$

Then $r_1 + r_2 = -\frac{\mu}{\alpha_1}$

$$r_1 r_2 = \frac{-\alpha_2^2}{2\,\alpha_1} \qquad 4.20$$

Therefore $r_1 = r_1(\alpha_1\,\alpha_2)$ and no new constant has been introduced. The coordinates of the satellite, r, λ , L are determined from the transformation equations:

$$\frac{\partial S}{\partial \alpha_1} = \beta_1,\ \frac{\partial S}{\partial \alpha_2} = \beta_2,\ \frac{\partial S}{\partial \alpha_3} = \beta_3 \qquad 4.21$$

Since

$$S = -\alpha_1 t + \alpha_3 \lambda + S_1(r) + S_2(L) \qquad 4.22$$

and α_3 does not appear in S_1, α_1 does not appear in S_2:

$$-t + \frac{\partial S_1}{\partial \alpha_1} = \beta_1$$

$$\lambda + \frac{\partial S_2}{\partial \alpha_3} = \beta_3 \qquad 4.23$$

$$\frac{\partial S_1}{\partial \alpha_2} + \frac{\partial S_2}{\partial \alpha_2} = \beta_2$$

From equations 4.17 and 4.18, substituted into the first of 4.23:

$$t + \beta_1 = \int_{r_1}^{r} \frac{r^2}{\sqrt{2\alpha_1 r^2 + 2\mu r - \alpha_2^2}} \frac{dr}{r} \qquad 4.24$$

$$- \frac{\partial r}{\partial \alpha_1} \left[\frac{1}{r} \sqrt{2\alpha_1 r^2 + 2\mu r - \alpha_2^2} \right]_{r = r_1}$$

The last part of the expression on the right side of 4.23 vanishes by virtue of the definition of r_1

$$t + \beta_1 = \sqrt{\frac{1}{-2\alpha_1}} \int_{r_1}^{r} \frac{r\, dr}{\sqrt{(r-r_1)(r_2-r)}} \qquad 4.25$$

Where it will be recalled that r_1 is the smaller and r_2 the larger of the two roots of

$$r^2 + \frac{\mu}{\alpha_1} r - \frac{\alpha_2^2}{2\alpha_1} = 0 \qquad 4.26$$

Hence r lies between r_1 and r_2 and $(r - r_1)(r_2 - r)$ is a positive quantity. The integral 4.24 vanishes when $r = r_1$, therefore, $t + \beta_1 = 0$ when $r = r_1$ and $\beta_1 = -\tau$, where τ is the epoch, the time of passage of the point on the orbit where r is minimum, the periapsis (perigee or perihelion).

We define two new constants, a and e, by:

$$r_1 = a\,(1-e)$$

$$r_2 = a\,(1+e) \qquad 4.27$$

Thus from 4.20

$$\alpha_1 = -\frac{\mu}{2a}$$

$$\alpha_2 = \sqrt{\mu a\,(1-e^2)} \qquad 4.28$$

If we now introduce the new variable E defined by

$$r = a\,(1 - e\cos E) \qquad 4.29$$

such that $E = 0$ when $r = r_1$

$$r - r_1 = ae\,(1 - \cos E)$$

$$r_2 - r = ae\,(1 + \cos E) \qquad 4.30$$

$$dr = \sin E\, dE$$

Then 4.25 becomes:

$$\sqrt{-2\,\alpha_1}\,(t + \beta_1) = a \int_0^E (1 - e\cos E)\, dE \qquad 4.31$$

$$E - e\sin E = \frac{1}{a}\sqrt{-2\,\alpha_1}\,(t + \beta_1) \qquad 4.32$$

$$E - e\sin E = n\,(t - \tau) \qquad 4.33$$

which is Kepler's equation. See Equation 2.29. Thus, we have identified the following quantities:

$$\beta_1 = -\tau$$

$$\alpha_1 = -\frac{\mu}{2a} \qquad 4.34$$

$$\alpha_2 = \sqrt{\mu a\,(1 - e^2)}$$

Now consider the integral

$$\lambda + \frac{d\,S_2}{d\,\alpha_3} = \beta_3 \qquad 4.35$$

From 4.18

$$\beta_3 = \lambda - \alpha_3 \int_0^L \frac{\sec^2 L\, dL}{\sqrt{\alpha_2^2 - \alpha_3^2 \sec L}}$$

$$= \lambda - \int_0^L \frac{\sec^2 L\, dL}{\sqrt{\dfrac{\alpha_2^2 - \alpha_3^2}{\alpha_3^2} - \tan^2 L}} \qquad 4.36$$

In order for the denominator of 4.36 to be real, $\alpha_2^2 > \alpha_3^2$. If we define a new angle, ϕ, by:

$$\alpha_2^2 - \alpha_3^2 = \alpha_3^2 \tan^2 \phi \qquad 4.37$$

and require that ϕ be a positive angle, then 4.36 becomes

$$\beta_3 = \lambda - \int_0^L \frac{\sec^2 L \, dL}{\sqrt{\tan^2 \phi - \tan^2 L}} \qquad 4.38$$

which integrates to

$$\beta_3 = \lambda - \sin^{-1} \left(\frac{\tan L}{\tan \phi} \right) \qquad 4.39$$

(See Dwight, Tables of Integrals and Other Mathematical Date, formulas 427.3, 512.1)

Recalling that

$$x = r \cos \lambda \cos L$$

$$y = r \sin \lambda \cos L \qquad 4.1$$

$$z = r \sin L$$

We rewrite 4.39 as:

$$\sin (\lambda - \beta_3) = \frac{\tan L}{\tan \phi} \qquad 4.40$$

$$r \sin (\lambda - \beta_3) \cos L = r \frac{\sin L}{\tan \phi} = z \cot \phi$$

$$y \cos \beta_3 - x \sin \beta_3 = z \cot \phi \qquad 4.41$$

4.41 is the equation of a plane through the center of force

$$a x + by + cz = 0$$

Further, from 4.40 we see that when $L = 0$, i.e. when the satellite passes through the reference plane,

$$\lambda = \beta_3 \qquad 4.42$$

Therefore, β_3 is the longitude of the ascending node, ☊. Since the maximum value of sin (λ - β_3) is one, the maximum value of L is ϕ. But the maximum value of L for a particle which moves in a plane inclined to the reference plane by an angle i, is the inclination angle, i. Thus

$$\Phi = i \tag{4.43}$$

Accordingly, from 4.37

$$\alpha_3 = \alpha_2 \cos i \tag{4.44}$$

or from 4.34

$$\alpha_3 = \sqrt{\mu a \,(1-e^2)} \cos i \tag{4.45}$$

We have now identified

$$\begin{aligned} \beta_1 &= -\tau \\ \alpha_1 &= -\mu/2a \\ \alpha_2 &= \sqrt{\mu a\,(1-e^2)} \\ \beta_3 &= \Omega \\ \alpha_3 &= \sqrt{\mu a\,(1-e^2)} \cos i \end{aligned} \tag{4.46}$$

We now turn to the final integral:

$$\beta_2 = \frac{\partial S_1}{\partial \alpha_2} + \frac{\partial S_2}{\partial \alpha_1} \tag{4.23}$$

$$\begin{aligned} \beta_2 = \frac{\partial}{\partial \alpha_2} \int_{r_1(\alpha_1\, \alpha_2)}^{r} \sqrt{2\,\alpha_1 r^2 + 2\mu r - \alpha_2^2}\;\frac{dr}{r} \\ + \frac{\partial}{\partial \alpha_2} \int_{o}^{L} \sqrt{\alpha_2^2 - \frac{\alpha_3^2}{\cos^2 L}}\; dL \end{aligned} \tag{4.47}$$

$$\beta_2 = \alpha_2 \int_o^L \frac{dL}{\sqrt{\alpha_2^2 - \frac{\alpha_3^2}{\cos^2 L}}} - \alpha_2 \int_{r_1}^r \frac{dr}{r\sqrt{2\alpha_1 r^2 + 2\mu r - \alpha_2^2}} \qquad 4.48$$

$$- \left[\frac{1}{r} \sqrt{2\alpha_1 r^2 + 2\mu r - \alpha_2^2} \quad \frac{\partial r}{\partial \alpha_2} \right]_{r = r_1}$$

Again the final term of 4.48 is zero. Let $\beta_2 = I_1 - I_2$

$$I_1 = \alpha_2 \int_o^L \frac{\cos L \, dL}{\sqrt{\alpha_2^2 - \alpha_3^2 - \alpha_2^2 \sin^2 L}} \qquad 4.49$$

$$I_1 = \int_o^L \frac{\cos L \, dL}{\sqrt{\sin^2 i - \sin^2 L}}$$

from 4.37. 4.49 may now be integrated to obtain

$$I_1 = \sin^{-1} \frac{\sin L}{\sin i} \qquad 4.50$$

From the Law of Sines of spherical trigonometry, I is the angle from the node to the satellite which we will call u, the argument of the latitude.

$$I_1 = u \qquad 4.51$$

$$\sin U = \frac{\sin L}{\sin i}$$

Turning now to the second integral in 4.48,

$$I_2 = \alpha_2 \int_{r_1}^r \frac{dr}{r\sqrt{2\alpha_1 r^2 + 2\mu r - \alpha_2^2}} \qquad 4.52$$

$$I_2 = \frac{\alpha_2}{\sqrt{-2\alpha_1}} \int_{r_1}^{r} \frac{dr}{r\sqrt{(r-r_1)(r_2-r)}} \qquad 4.53$$

In terms of the eccentric anomaly,

$$I_2 = \frac{1}{a} \int_{o}^{E} \frac{dE}{\sqrt{1-e\cos E}} \qquad 4.54$$

$$I_2 = \frac{1}{a\,(1-e)} \int_{o}^{E} \frac{\sec^2\frac{E}{2}\,dE}{1+\frac{1+e}{1-e}\tan^2\frac{E}{2}} \qquad 4.55$$

If we let v be defined from

$$\tan\frac{v}{2} = \sqrt{\frac{1+e}{1-e}}\,\tan\frac{E}{2} \qquad 4.56$$

then

$$I_2 = v \qquad 4.57$$

Equations 4.56 and 4.57 will be more easily interpreted by recalling that for motion in an elliptic orbit we had

$$r = \frac{a\,(1-e^2)}{1+e\cos v} \qquad 4.58$$

and

$$r = a\,(1+e\cos E) \qquad 4.59$$

hence

$$(1+e\cos E)(1+e\cos v) = 1-e^2 \qquad 4.60$$

which can also be written as:

$$\tan\frac{v}{2} = \sqrt{\frac{1+e}{1-e}}\,\tan\frac{E}{2} \qquad 4.56$$

or

$$\sin E = \frac{(1-e^2)^{\frac{1}{2}}\sin v}{1+e\cos v} \qquad 4.61$$

Thus

$$\beta_2 = u - v = \omega \qquad 4.62$$

ω is called the argument of periapsis, and is evidently the angle from the node line to the periapsis.

Summary of the Canonic Constants

The canonic momenta and coordinates are:

$$\alpha_1 = \frac{-\mu}{2a} \qquad \beta_1 = -\tau$$

$$\alpha_2 = \sqrt{\mu a\,(1-e^2)} \qquad \beta_2 = \omega \qquad 4.63$$

$$\alpha_3 = \sqrt{\mu a\,(1-e^2)}\cos i \qquad \beta_3 = \Omega$$

The elements of the orbit in terms of the canonic constants are:

$$a = \frac{-\mu}{2\alpha_1} \qquad \tau = -\beta_1$$

$$e = \sqrt{1 + \frac{2\alpha_1 \alpha_2^2}{\mu^2}} \qquad \omega = \beta_2 \qquad 4.64$$

$$i = \cos^{-1}\left(\frac{\alpha_3}{\alpha_2}\right) \qquad \Omega = \beta_3$$

Discussion of the Canonical Constants

We have already found that when the Hamiltonian itself does not contain the time explicitly, then it is a constant of the motion. We have called that constant

$$\alpha_1 = \frac{E}{m} = -\frac{\mu}{2a} \qquad 4.65$$

Note that from the form of U,

$$\frac{U}{m} = -\frac{\mu}{r} \qquad 4.66$$

the zero level of potential energy is at $r \rightarrow \infty$.

The third canonical constant, α_3, was given by equation 4.10 as

$$p_\lambda = \alpha_3 \qquad 4.10$$

But from 4.6

$$\alpha_3 = p_\lambda = \frac{\partial L'}{\partial \dot{\lambda}} = r^2 \cos^2 L \, \dot{\lambda} \qquad 4.6$$

$r^2 \cos^2 L \, \dot{\lambda}$ is the component of angular momentum per unit mass along the fixed polar axis, normal to the reference plane, and from 4.6 is therefore conserved.

We write equation 4.16 in the form

$$p_L^2 + \frac{\alpha_3^2}{\cos^2 L} = \alpha_2^2 \qquad 4.67$$

Now in plane polar coordinates, the Hamiltonian per unit mass is:

$$H = \frac{1}{2}\left[p_r^2 + \frac{p^2}{r^2} \right] - \frac{\mu}{r} \qquad 4.68$$

Comparison of 4.68 with 4.7

$$H = \frac{1}{2}\left(p_r^2 + \frac{p_\lambda^2}{r^2 \cos^2 L} + \frac{p_L^2}{r^2} \right) - \frac{\mu}{r} \qquad 4.7$$

shows that p^2 must be identified with α_2. In 4.68, p^2 is the total angular momentum of the particle about the center of force. Hence α_2 is the magnitude of the total angular momentum, which is conserved. Note that the two statements, (magnitude of angular momentum vector and component of angular vector along the polar axis are both constant) are not redundant. In fact, in order to specify that the angular momentum vector is constant requires the third statement; the motion remains in a fixed plane ($\beta_3 = \Omega$).

FOREWORD TO APPENDIX B

Derivation of a complete set of equations of motion suitable for machine solution is not a completely trivial matter. The equations presented in the paper which is reproduced here as Appendix B have certain advantages with respect to computer scaling and accuracy. It is hoped that they will be useful to users of these notes.

CHAPTER V

PERTURBATIONS

Analysis of the motion of a body in a central force field is the basis of most of the knowledge which we have of planetary motion. The detailed analysis of the motion inevitably must recognize that the pure central force field is a mathematical idealization, and that the motion of real bodies in the solar system is influenced by deviations from the central force field. Fortunately, the deviation often is not large. If this were not the case, we would know only a few solutions of idealized equations of motion. Perturbation analysis is a method of treating the motion as a deviation from the conic orbit which characterizes the central force field with inverse-square variation.

Before making the perturbation analysis, let us examine a simpler problem briefly, and illustrate the method. (Reference: Celestial Mechanics, Moulton, 2nd Edition, Chapter X.)

Consider the solution of

$$\ddot{x} + kx = -\mu\dot{x}^3 + \upsilon \cos lt \tag{5.1}$$

Rewriting the equation so that it is in the form of the canonical equations of motion, let

$$\dot{x} = p$$

$$x = q \tag{5.2}$$

Then 5.1 becomes

$$\dot{q} - p = 0$$

$$\dot{p} + kq = -\mu p^3 + \upsilon \cos lt \tag{5.3}$$

Considering the left hand side of 5.1 to represent the undisturbed motion, and the right hand side to represent the perturbing force, the equations of the undisturbed motion are

$$\dot{q} - p = 0$$

$$\dot{p} + kq = 0 \tag{5.4}$$

(See pp. 64-66 where this problem was solved by the Hamilton-Jacobi method.)

The solution of 5.4 is the well-known harmonic motion:

$$q = \alpha \cos \sqrt{k}\, t + \beta \sin \sqrt{k}\, t$$
$$p = -\sqrt{k}\, \alpha \sin \sqrt{k}\, t + \sqrt{k}\beta \cos \sqrt{k}\, t \tag{5.5}$$

In our terminology, α and β are the elements of the orbit.

Let us forget, for the moment, that 5.5 are the solutions of 5.4 and consider 5.5 merely as a set of transformation equations, from the old variables, p and q to new variables α and β. We now wish to find $\alpha(t)$ and $\beta(t)$ to satisfy 5.3. Substituting 5.5 in 5.3 and remembering that α and β are now functions of time, we have:

$$(\cos \sqrt{k}\, t)\dot{\alpha} + (\sin \sqrt{k}\, t)\dot{\beta} = 0$$
$$(-\sin\sqrt{k}\, t)\dot{\alpha} + (\cos \sqrt{k}\, t)\dot{\beta} = \mu k[\alpha \sin\sqrt{k}\, t - \beta \cos\sqrt{k}\, t]^3 + \frac{\upsilon}{\sqrt{k}} \cos 1t \tag{5.6}$$

These equations can be solved for $\dot{\alpha}$ and $\dot{\beta}$:

$$\dot{\alpha} = -\mu k[\alpha \sin\sqrt{k}\, t - \beta \cos \sqrt{k}\, t]^3 \sin\sqrt{k}\, t - \frac{\upsilon}{\sqrt{k}} \cos 1t \sin\sqrt{k}t \tag{5.7}$$
$$\dot{\beta} = \mu k[\alpha \sin\sqrt{k}t] - \beta \cos\sqrt{k}t]^3 \cos\sqrt{k}t] + \frac{\upsilon}{\sqrt{k}} \cos 1t \cos\sqrt{k}t] \tag{5.8}$$

5.7 and 5.8 are just as valid as 5.3, since no approximation was made in deriving them--we have merely changed variables. Recall that if μ and υ are zero, the solution of 5.7 and 5.8 is $\dot{\alpha} = 0$, $\dot{\beta} = 0$ and the solution of 5.3 becomes just the transformation Equations (5.5) with α and β constant. It is by no means a coincidence that this reminds us strongly of the Hamilton-Jacobi theory.

Equations 5.7 and 5.8 are at least as difficult to solve analytically as 5.1 or 5.3. But suppose that μ and υ are small, so that the forces represented by the right hand side of 5.1 are small perturbing forces when compared with the individual terms on the left-hand side. Then for a short period of time we can consider α and β to be constants in the right-hand side of 5.7 and 5.8, which therefore can be integrated to give $\alpha(t)$ and $\beta(t)$. (The solution is given in

Moulton.) The result is of the form

$$\dot{\alpha} = f(\alpha, \beta, t)$$
$$\dot{\beta} = g(\alpha, \beta, t) \qquad (5.9)$$

Thus as α and β change slowly with time, they also change f and g which must be continually changed as the error builds up.

The example which has just been cited is an example of the general method known as the method of "variation of arbitrary constants"--somewhat unfortunately. The essence of the method is to seek a transformation of variables such that the new variables will change very slowly with time, and therefore numerical integration methods can be applied, with the approximation being reasonably valid across a wide time interval.

One obvious application of the method is to the central force problem with the addition of a set of forces which slightly perturb the central force field. For example, if we are concerned with the motion of a body in a nearly inverse-square central force field due to the closest or largest attracting body together with a slight perturbing force due to a distant or small body, we should expect that the motion will differ not too greatly from the pure conic. Such a solution could be written in terms of the elements of the conic, but now we expect the elements to change (slowly) with time.

Our first task will be to write the equations of motion in terms of the new variables, α_i and β_i. We have already accomplished this, (in principle) since this was just the object of the Hamilton-Jacobi theory.

It will be recalled that the transformation equations:

$$P_i = \frac{\partial S}{\partial q_i}$$
$$\beta_i = \frac{\partial S}{\partial \alpha_i} \qquad (5.10)$$

were a set of equations which related the old variables p_i, q_i (p_r, p_L, p_λ, r, λ, L) to new variables P_i, Q_i, (α_1, α_2, α_3, β_1, β_2, β_3). The new variables α_i, β_i are the canonic constants (which are related to the elements of the orbit). The elements of the orbit are constants for the Kepler problem. When there is an additional perturbation of the force field, we will expect these "constants" to change slowly with time.

The transformation Equation (5.10) can be applied even if the α_i and β_i are not constants. We said that these quantities must be constant if the transformed Hamiltonian, K, were zero, since

$$\begin{aligned} \dot{\alpha}_i &= -\frac{\partial K}{\partial \beta_i} \\ \dot{\beta}_i &= \frac{\partial K}{\partial \alpha_i} \end{aligned} \tag{5.11}$$

where $K = H + \frac{\partial S}{\partial t} = 0$.

Now suppose that K is not zero:

$$K = K_1 + K_2 \tag{5.12}$$

where

$$K_1 = H + \frac{\partial S}{\partial t} \equiv 0 \tag{5.13}$$

$$K_2 = -R(\alpha, \beta, t) \tag{5.14}$$

We select the function R to be the perturbation force function.

$$U = -\frac{K}{r} - R \tag{5.15}$$

The gradient of R is the perturbing force. Then

$$K = K_1 + K_2 = H + \frac{\partial S}{\partial t} - R \tag{5.16}$$

$$K = -R \tag{5.17}$$

The canonical equations of motion in terms of the new variables α_i, β_i are:

$$\begin{aligned} \dot{\alpha}_i &= -\frac{\partial K}{\partial \beta} = \frac{\partial R}{\partial \beta} \\ \dot{\beta}_i &= \frac{\partial K}{\partial \alpha} = -\frac{\partial R}{\partial \alpha} \end{aligned} \tag{5.18}$$

We have thus derived a set of Equations (5.18) for the time rate of change of the canonic "constants," and since R is a small quantity we expect the canonic constants to change slowly.

While the problem appears to be solved very neatly, there are a number of practical hookers which can generate serious problems when the method is applied. The first practical problem is that of determining the force function, R, and expressing it in terms of the canonic variables or the elements of the osculating orbit. (The osculating orbit is the orbit the body would describe if all perturbing forces were removed at any given time. The elements of this orbit therefore change with time and are equal at any time to the elements which would be determined by Equations 5.18.) Moulton devotes Chapters IX and X to determination of R and integration of Equations 5.18 or their equivalent. Smart also devotes several chapters to these problems. Because of the lengthy algebraic manipulations, we will present here only a portion of the results and refer the reader to the books by Moulton, Smart and to Brown's "Lunar Theory."

Equations of Motion in Terms of the Elements of the Osculating Ellipse

Because of custom, it is usual to use the equations of motion in terms of the elements of the orbit, rather than in terms of the canonic constants as indicated by 5.18.

If we denote any element of the orbit by $a_m (m = 1, 2, \ldots 6)$ then:

$$\dot{a}_m = \sum_{r=1}^{3} \left(\frac{\partial a_m}{\partial \alpha_r} \dot{\alpha}_r + \frac{\partial a_m}{\partial \beta_r} \dot{\beta}_r \right) \tag{5.19}$$

$$= \sum_{r=1}^{3} \left(- \frac{\partial a_m}{\partial \alpha_r} \frac{\partial K}{\partial \beta_r} + \frac{\partial a_m}{\partial \beta_r} \frac{\partial K}{\partial \alpha_r} \right) \tag{5.20}$$

But

$$\frac{\partial K}{\partial \alpha_r} = \sum_{S=1}^{6} \frac{\partial K}{\partial a_S} \frac{\partial a_S}{\partial \alpha_r}$$

$$\frac{\partial K}{\partial \beta_r} = \sum_{S=1}^{6} \frac{\partial K}{\partial a_S} \frac{\partial a_S}{\partial \beta_r} \tag{5.21}$$

The operations indicated by 5.20 and 5.21 are lengthy but not difficult and are carried out, for example, in Moulton.

It should be noted that the equations of motion which result from 5.18, 5.19, 5.20, and 5.21 are completely rigorous. No approximation has been made in deriving them. The equations are:

$$\frac{d\Omega}{dt} = \frac{1}{na^2\sqrt{1-e^2}\,\sin i}\,\frac{\partial R}{\partial i}$$

$$\frac{di}{dt} = \frac{\cot i}{na^2\sqrt{1-e^2}}\left[\frac{\partial R}{\partial \omega} \quad \frac{1}{\cos i}\,\frac{\partial R}{\partial \Omega}\right]$$

$$\frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{na^2 e}\,\frac{\partial R}{\partial e} \quad \frac{\cot i}{na^2\sqrt{1-e^2}}\,\frac{\partial R}{\partial i} \tag{5.22}$$

$$\frac{da}{dt} = \frac{2}{na}\,\frac{\partial R}{\partial \chi}$$

$$\frac{de}{dt} = \frac{1-e^2}{na^2 e}\,\frac{\partial R}{\partial \chi} - \frac{\sqrt{1-e^2}}{na^2 e}\,\frac{\partial R}{\partial \omega}$$

$$\frac{d\chi}{dt} = -\frac{1}{na^2}\left[\frac{1-e^2}{e}\,\frac{\partial R}{\partial e} + 2a\,\frac{\partial R}{\partial a}\right]$$

where:

$$\chi = -n\tau \tag{5.23}$$

$$n = \frac{K}{a^{3/2}} = \frac{\gamma(M+m)}{a^{3/2}} \tag{5.24}$$

n is the mean motion and τ is the epoch, the time of perigee passage.

In order to obtain a solution of the above equations, one must be able to express the perturbation potential, R, in terms of the elements of the orbit. This may be difficult or even impossible. For numerical treatment, using a step-by-step integration process, it is convenient to express the perturbation acceleration in terms of components.

We resolve the perturbing acceleration into components W, S, R where W is the component of acceleration perpendicular to the plane of the orbit, positive toward the north pole; S is the component in the plane of the orbit at right angles to the radius vector, positive in the direction of motion; R is the component along the radius vector, positive in the positive r direction. (Note that we use R to represent both the perturbation potential and the radial component of acceleration. This unfortunate choice of symbols is fairly standard.)

In terms of W, S. R the equations of motion are: (See Moulton, for example)

$$\frac{d\Omega}{dt} = \frac{r \sin u}{na^2\sqrt{1-e^2} \sin i} W$$

$$\frac{di}{dt} = \frac{r \cos u}{na^2\sqrt{1-e^2}} W$$

$$\frac{d\omega}{dt} = \frac{-\sqrt{1-e^2} \cos v}{nae} R + \frac{\sqrt{1-e^2}}{nae} \left(1 + \frac{r}{p}\right) \sin v\, S$$

$$- \frac{r \sin u \cot i}{n a^2 \sqrt{1-e^2}} W \qquad (5.25)$$

$$\frac{de}{dt} = \frac{\sqrt{1-e^2} \sin v}{n a} R + \frac{\sqrt{1-e^2}}{na^2 e} \left[\frac{a^2(1-e^2)}{r} - r\right] S$$

$$\frac{da}{dt} = \frac{2 e \sin v}{n\sqrt{1-e^2}} R + \frac{2a\sqrt{1-e^2}}{nr} S$$

$$\frac{dX}{dt} = - \frac{1}{na}\left[\frac{2r}{a} - \frac{1-e^2}{e} \cos v\right] R$$

$$- \frac{(1-e^2)}{nae}\left[1 + \frac{r}{p}\right] \sin v\, S$$

Because of the many different possible combinations of elliptic elements, there are approximately as many different forms of 5.25 as there are authors writing about them. A number of these are given in Moulton, Smart and "Space Flight" by Krafft Ehricke. Several of them are of sufficient interest and utility that we will reproduce them here.

It is sometimes convenient to have the perturbation potential expressed in terms of T, N and W, where T is the component in the direction of motion, N is normal to T, positive toward the interior of the ellipse. The equations are given completely by Ehricke, but we are particularly interested in the case where N = W = 0, T being either a force of resistance, such as drag, or a tangential propulsive force. For this case:

$$\frac{d\Omega}{dt} = 0$$

$$\frac{di}{dt} = 0$$

$$\frac{d\omega}{dt} = \frac{2\sqrt{1-e^2}}{nae} \; \frac{\sin v}{1 + e^2 + 2e \cos v} \; T$$

$$\frac{da}{dt} = 2\,\frac{\sqrt{1+e^2+2e \cos v}}{n\sqrt{1-e^2}} \; T \tag{5.26}$$

$$\frac{de}{dt} = \frac{2\sqrt{1-e^2}\,(\cos v + e)}{na\sqrt{1+e^2+2e \cos v}} \; T$$

$$\frac{d\chi}{dt} = -\frac{2(1-e^2)(1+e^2+2e \cos v) \sin v}{nae(1+e \cos v)\sqrt{1+e^2+2e \cos v}} \; T$$

When the right-hand side of Equations 5.25 or 5.26 does not depend explicitly on time, we can express the equations in terms of one of the angle variables, v or u. The equations are then particularly useful for numerical work. We will briefly review a Russian study of the lifetime of an artificial satellite, using u as the independent variable, however, let us first examine Equations 5.25 which furnish a considerable amount of qualitative information about trajectories and maneuvers.

First we note that the longitude of the node, Ω, and the inclination, i, can be changed only by application of a normal component, W. If the problem is to change from one orbit plane to another, we take the reference plane to be the plane of the target orbit. Then W will change the interceptor plane inclination most effectively when $r \cos u$ is maximum, near the node line for small e. We should select the ascending or descending node according to the value of r.

For small eccentricity, e, the component S is more effective than R in changing the semi-major axis, a, which determines the orbit energy level.

There are a great number of conclusions which can be drawn from inspection of the equations of motion in terms of the elements of the osculating ellipse (the ellipse which the body would follow if all perturbing forces ceased). It should be emphasized again that these equations are exact--no approximation was made in deriving them. A word of caution--when e = 0, the elements ω and τ, and therefore v are undefined.

EXAMPLE: THE LIFETIME OF AN ARTIFICIAL EARTH SATELLITE

Reference: "Determining the Lifetime of an Artificial Earth Satellite and an Investigation of Secular Perturbations of its Orbit," by D. Ye. Ohhotsimskey, T. M. Eneyev and G. P. Taratynova. Translation in Vol. VI, Annals of the IGY.

The following is presented as an application of the method of variation of elements. Because this work was completed several years ago, the data used are now obsolete and therefore the results are not numerically accurate. Those concerned with problems of this type should make a careful survey of the current literature on the subject. The starting point for this investigation was the equations of motion in the following form:

$$
\begin{aligned}
\frac{dp}{dt} &= \frac{2r\sqrt{p}}{\gamma M} S \\
\frac{de}{dt} &= \frac{\sqrt{p}}{\gamma M}(\sin v)R + \frac{\sqrt{p}}{\sqrt{\gamma M}}\left(1 + \frac{r}{p}\cos v + e\frac{r}{p}\right) S \\
\frac{d\omega}{dt} &= -\frac{\sqrt{p}}{e\sqrt{\gamma M}}(\cos v)R + \frac{\sqrt{p}}{e\sqrt{\gamma M}}\left(1 + \frac{r}{p}\right)\sin v\, S \\
&\qquad - \frac{\sqrt{p}}{\sqrt{\gamma M}}\cot i \sin u\, W \\
\frac{d\Omega}{dt} &= \frac{\sqrt{p}}{\sqrt{\gamma M}}\frac{r}{p}\frac{\sin u}{\sin i} W \\
\frac{di}{dt} &= \frac{\sqrt{p}}{\sqrt{\gamma M}}\frac{r}{p}\cos u\, W \\
\frac{d\tau}{dt} &= \frac{r^2}{e\gamma M}\left[-(\cos v - e(\sin v)N)R + \frac{p}{r}NS\right]
\end{aligned}
\tag{5.27}
$$

where the symbols have their usual meaning and

$$u = \omega + v \qquad r = \frac{p}{1 + e\cos v} \tag{5.28}$$

$$N = \frac{2p^2}{r^2}\int_0^v \frac{\cos v\, dv}{(1 + e\cos v)^3} \tag{5.29}$$

and v is related to t by

$$t - \tau = \frac{p^{3/2}}{\sqrt{\gamma M}} \int_0^v \frac{dv}{(1 + e\cos v)^2} \qquad (5.30)$$

γ is the gravitational constant, and M is the mass of the earth. Note that e in the expression for N and $t - \tau$ is a function of time. If the right-hand sides of 5.27 do not depend explicitly on time it is convenient to change independent variables, introducing the argument of the latitude, u, as independent variable.

It can be shown geometrically that

$$\frac{du}{dt} = \sqrt{\gamma M}\frac{\sqrt{p}}{r^2}\left(1 - \frac{r^3}{\alpha M p} \cot i(\sin u)W\right) \qquad (5.31)$$

Using 5.31, we can transform 5.27 to be:

$$\frac{dp}{du} = \frac{2\alpha}{\gamma M} r^3 S$$

$$\frac{de}{du} = \frac{r^2\alpha}{\gamma M}\left[\sin v\cdot R + \cos v(1 + \frac{r}{p})S + e\frac{r}{p}S\right]$$

$$\frac{d\omega}{dt} = \frac{r^2\alpha}{\gamma Me}\left[\cos v \cdot R + \sin v(1 + \frac{r}{p}) S - e\frac{r}{p} \cot i \sin u W\right] \qquad (5.32)$$

$$\frac{d\Omega}{dt} = -\frac{r^3\alpha}{\gamma Mp}\frac{\sin u}{\sin i} W$$

$$\frac{di}{du} = -\frac{r^3\alpha}{\gamma Mp} \cos u W$$

where

$$\alpha = \frac{1}{1 - \frac{r^3}{\gamma Mp} \cot i \sin u \cdot W} \qquad (5.33)$$

The equation for τ is not used and is not presented. It is pointed out by the authors that Equations 5.32 are also valid if we simply write v as the independent variable and define α differently:

$$\alpha = \frac{1}{1 + \frac{r^2}{\gamma Me} \cos v\cdot S - \frac{r^2}{\gamma Me} 1 + \frac{r}{p}\sin v\cdot T} \qquad (5.34)$$

Generally, the equations using u as the variable, rather than v, are more useful since for very low eccentricities v is hard to define.

If we neglect the rotation of the atmosphere with the earth, and assume that the earth's gravitational field is a central-force field, Equations 5.32 reduce to:

$$\frac{dp}{du} = \frac{2r^3}{\gamma M} S$$

$$\frac{de}{du} = \frac{r^2}{\gamma M}\left[\sin v \cdot R + \cos v \left(1 + \frac{r}{p}\right)S + e\frac{r}{p}S\right]$$

$$\frac{d\omega}{du} = \frac{r^2}{\gamma Me}\left[\cos v \cdot R + \sin v\left(1 + \frac{r}{p}\right)S\right] \tag{5.35}$$

$$\frac{d\Omega}{du} = 0$$

$$\frac{di}{du} = 0$$

The aerodynamic drag of the body is

$$D = \frac{1}{2} \rho V^2 A C_D$$

where ρ is the density, V the total velocity, A the area on which the drag coefficient, C_D, is based. Then the tangential acceleration, T is

$$T = -\frac{D}{m} \tag{5.36}$$

and the components R and S are

$$R = -\frac{C_D A}{m}\frac{\rho V}{2} V_r$$

$$S = \frac{C_D A}{m}\frac{\rho V}{2} V_n \tag{5.37}$$

where V_r and V_n are the velocity components along and normal to the radius vector.

Using the following formulae:

$$r = \frac{p}{1 + e \cos V}$$

$$V = \sqrt{\frac{\gamma M}{p}} \sqrt{1 + 2\,e \cos v + e^2}$$

$$V_r = \sqrt{\frac{\gamma M}{p}}\; e \sin v \qquad (5.38)$$

$$V_n = \sqrt{\frac{\gamma M}{p}}\,(1 + e \cos v)$$

(Problem: prove 5.38.)

We obtain

$$\frac{dp}{du} = \left(\frac{-C_D A}{m}\right)(\rho)\, \frac{p}{1 + e \cos v} \sqrt{1 + 2e \cos v + e^2}$$

$$\frac{de}{du} = \left(\frac{-C_D A}{m}\right)(\rho)\, \frac{p(e + \cos v)}{(1 + e \cos v)^2} \sqrt{1 + 2e \cos v + e^2} \qquad (5.39)$$

$$\frac{d\omega}{du} = \left(\frac{-C_D A}{m}\right)(\rho)\, \frac{p \sin v}{2e(1 + e \cos v)} \sqrt{1 + 2e \cos v + e^2}$$

Equations 5.39 could be integrated, but, since satellite lifetime may be measured in years, (or centuries) an enormous amount of machine time and a considerable accumulated error could result. Instead, the authors followed the scheme given below.

Integrating 5.39 from 0 to 2π

$$\Delta p = -\,\rho_1 \frac{C_D A}{m} \int_0^{2\pi} \phi\,(p, e, w, u)\, du$$

$$\Delta e = -\,\rho_1 \frac{C_D A}{m} \int_0^{2\pi} \psi\,(p, e, w, u)\, du \qquad (5.40)$$

$$\Delta\omega = -\,\rho_1 \frac{C_D A}{m} \int_0^{2\pi} \chi\,(p, e, w, u)\, du$$

where Δp, Δe, and $\Delta\omega$ are the changes in the semi-latus rectum, eccentricity and argument of perigee in one revolution. In calculating these integrals, p, e, and ω are considered constant on the right-hand side of 5.40.

$$\phi = \frac{\rho}{\rho_1} \frac{p^2 \sqrt{1 + 2e \cos v + e^2}}{(1 + e \cos v)^2}$$

$$\psi = \frac{\rho}{\rho_1} \frac{p \sqrt{1 + 2e \cos v + e^2}}{(1 + e \cos v)^2} (e + \cos v) \qquad (5.41)$$

$$\chi = \frac{\rho}{\rho_1} \frac{p \sin \sqrt{1 + 2e \cos v + e^2}}{2e(1 + e \cos v)^2}$$

and $v = u - \omega$, ρ/ρ_1 is a function of r and ρ_1 is the density at a reference altitude.

Under the assumption that p, e, and ω are constant, the last integral of 5.40 vanishes:

$$\Delta\omega = 0$$

For small eccentricities, this appears to be very questionable. However, only the change of p is significant in determining the lifetime, and from the first of Equations 5.35, it is seen that p depends on ω only through r and the effect on S. As long as $\frac{d\omega}{dt}$ is considerably less than the angular velocity of the body in orbit, the location of perigee should be unimportant for an integral over a full revolution. The change of ω will be considered to have negligible effect on the life of the satellite.

Assuming that the changes of p and e are small during a single revolution, we then can write:

$$\frac{dp}{dN} = - \frac{C_D A \rho_1}{m} \int_0^{2\pi} \phi \, (p,e,u) \, du$$

$$\frac{de}{dN} = - \frac{C_D A \rho_1}{m} \int_0^{2\pi} \psi \, (p,e,u) \, du \qquad (5.42)$$

Assuming $\omega = \omega_o$ = const.

If we let $\upsilon = N \frac{C_D A}{m}$, Equations (5.42) can be written

$$\frac{de}{dp} = \frac{\int_o^{2\pi} \psi \, du}{\int_o^{2\pi} \phi \, du} ; \quad \frac{du}{dp} = - \frac{1}{\rho_1 \int_o^{2\pi} \phi \, du} \tag{5.43}$$

Given the law of variation of density with altitude, 5.43 can be integrated by machine for various initial values of p and e, or of perigee altitude, h_π, and apogee altitude, h_α. The authors chose to assign initial values of h_π and h_α. Since each point on an integral curve could be considered as the initial point of some other motion, there was no necessity to vary both h_{α_o} and h_{π_o}. The equations were integrated for h_{α_o} = 1600 km and a range of h_{π_o} from 160 km to 500 km. It was assumed that when the satellite perigee reached 100 km the remaining life would be negligible.

One of the important features of this method is that Equations 5.43 are independent of the structural parameters of the satellite (the drag coefficient, mass, frontal area, etc.).

The authors present plots and tables from which one can immediately determine υ for any given initial height of perigee and apogee. From the value of υ, one then determines

$$N = \upsilon / \frac{C_D A}{m}$$

and then can determine the approximate lifetime by multiplying N by an average period--90 to 100 minutes, or by numerical integration,

$$t = \sum \text{(no. of revolutions)(time per revolution)}$$

The tables and charts presented by the authors appear to be in very useful form, but probably should be recomputed using the latest atmospheric data.

Since this phase of the investigation neglected the effects of oblateness of the earth and of the rotation of the atmosphere, the results are only approximate.

Effects of the Oblate Earth

If the earth were a homegeneous sphere, the potential U would be

$$U = - \frac{k}{r}$$

Because the earth is flattened at the poles, this expression must be modified. A commonly used expression for the potential of the oblate earth (see Ehricke, Space Flight, pages 156, 157):

$$U = -\frac{\gamma M}{r} + \frac{\epsilon}{3r^3}(3\sin^2 L-1) \qquad (5.44)$$

where

$$\epsilon = \gamma\, M\, r_e^2\, (\alpha - \frac{\beta}{2})$$

$$\beta = \omega_e^2\, r_e / g_e$$

L is the latitude, r_e is the equatorial radius of the earth, ω_e is the angular velocity of the earth's daily rotation, α is the ellipticity of the earth and g_e is the measured acceleration of gravity at the equator. (The basic expression was given by Clairuat in 1743 according to K. Ehricke, p. 156) The "ellipticity" α is given by

$$\alpha = \frac{a-b}{a} = \frac{1}{297}$$

$$\alpha - \frac{\beta}{2} = 1.6485 \times 10^{-3} \qquad (5.45)$$

The projections of the perturbing acceleration due to the earth's oblateness are obtained from 5.44.

$$R = \frac{\epsilon}{r^4}\,[3\sin^2 i \sin^2 u - 1]$$

$$S = -\frac{\epsilon}{r^4}[\sin^2 i \sin 2u] \qquad (5.46)$$

$$W = -\frac{\epsilon}{r^4}\,[\sin 2i \sin u]$$

When these expressions are substituted in 5.32 which then may be integrated for a single revolution, it is found that only the longitude of the node, Ω, and the argument of perigee, ω, are perturbed secularly. The authors give:

$$\frac{d\Omega}{dN} = -\frac{2\pi\epsilon}{r_{ave}^2\gamma M}\cos i$$

$$\frac{d\omega}{dN} = \frac{\pi\epsilon}{r_{ave}^2\gamma M}(5\cos^2 i-1) \qquad (5.47)$$

where $\frac{d}{dN}$ is the change per revolution.

r_{ave} is not defined, but probably it is close to the mean distance, or semi-major axis.

Effect of Rotation of the Atmosphere on the Elements of the Orbit

Since the satellite velocity is large compared with the linear velocity due to earth rotation, the drag forces can be expanded in powers of $\frac{\omega_e r}{V}$ to obtain:

$$
\begin{aligned}
R &= \frac{C_D A}{m} \frac{1}{2} \frac{V_n}{V} V_r \, \rho \, \omega_e \, r \cos i \\
S &= \frac{C_D A}{m} \frac{1}{2} \frac{V^2 + V_n^2}{V} \, \rho \, \omega_e \, r \cos i \\
W &= - \frac{C_D A}{m} \frac{1}{2} \, V \rho r \, \omega_e \sin i \cos u
\end{aligned}
\tag{5.48}
$$

Substituting these expressions for R, S, and W, assuming a variation of ρ with altitude and following the procedure described earlier results in expressions for the rates of change of each of the orbital elements. (The expressions are not given here because they involve an integral which is not tabulated.)

The results indicate that the lifetime of a satellite moving toward the east will be greater than that of one moving toward the west, with the difference amounting to as much as 10-12% for an equatorial orbit. The effect on the semi-latus rectum is proportional to cos i.

Note: The effects of the rotating atmosphere are studied in some detail, using the equations in terms of the orbital elements, by T. E. Sterne, "Effect of the Rotation of a Planetary Atmosphere upon the Orbit of a Close Satellite," ARS JOURNAL, October, 1959.

APPENDIX A

Mathematical Background: Calculus of Variations

References:

(a) Differential and Integral Calculus, by R. Courant, Vol. II, Nordeman Publishing Company, New York, 1945.

(b) An Introduction to the Calculus of Variations, by Charles Fox, Oxford University Press, 1954.

(c) Methods of Mathematical Physics, Courant-Hilbert, Vol. I, Interscience.

The usual condition that a function of n variables, $f(x_1 \ldots x_n)$ have a stationary value is stated as:

$$df = 0$$

or

$$\text{grad } f = 0$$

or

$$\partial f/\partial x_i = 0 \qquad \text{A.1}$$

The calculus of variations is concerned with investigating extreme—or, more generally, stationary—properties of a function. However, we are in this case presented with a "function of a function," and it is necessary to develop new methods for investigating the stationary properties. Our procedure will be analogous to the procedure by which Equations 1.1 are obtained.

Attention to problems of the variational type was called by Bernoulli in 1696 who stated the "BRACHISTOCHRONE" problem (shortest time problem) as follows:

In a vertical xy plane, a point A is to be joined to a point B which is lower (but not directly below) than A by a smooth curve $y = u(x)$ in such a way that a frictionless particle will slide from A to B along y in the shortest possible time.

Since there is no friction,

$$\frac{1}{2} mv^2 + mgh = \text{const} \qquad A.2$$

If we take v = 0 at h = 0

$$v = \sqrt{-2gh} \qquad A.3$$

Let us take $y = -h$

$$v = \sqrt{2gy} \qquad A.4$$

Since

$$v = \frac{ds}{dt} = \frac{ds}{dx}\frac{dx}{dt} \qquad A.5$$

$$T = \int_0^T dt = \int_{x_A}^{x_B} \frac{1}{v}\frac{ds}{dx}\,dx$$

also

$$ds^2 = dx^2 + dy^2 \qquad A.6$$

and

$$\frac{ds}{dx} = \sqrt{1 + \left(\frac{dy}{dx}\right)^2} \qquad A.7$$

Finally

$$T = \frac{1}{\sqrt{2g}} \int_{x_A}^{x_B} \sqrt{\frac{1 + (y')^2}{y}}\,dx \qquad A.8$$

where we have written y' for dy/dx. Now let us state the brachistochrone problem as:

Among all curves $y = \phi(x)$ which are continuously differentiable and which pass through the points $(x_A, 0)$ and (x_B, y_B) find that particular y(x) for which the integral

$$I = \frac{1}{\sqrt{2g}} \int_{x_A}^{x_B} \sqrt{\frac{1 + y'^2}{y}}\,dx \qquad A.8$$

will have the least possible value.

This is the problem that started it all. It is an example of the simplest kind of problem of the calculus of variations. Note the essential difference between this problem and the ordinary maximum-minimum problem. Here the value of the integral, I, depends on the behavior of y throughout the entire interval of x. The value of I cannot be specified by specifying a finite number of variables, but rather must be specified by prescribing a continuous function. I is called a functional, or a function of a function.

Problems of this type are of great importance in trajectory analysis. For example, consider an obviously important problem:

How should the attitude of a given booster be programmed in order to boost a satellite to a specified oribital altitude with maximum horizontal component of velocity?

Apparently, the horizontal velocity attained will depend on the entire attitude history, or program, during powered flight.

Some of these problems are treated in Chapter VII wherein some of the methods of the calculus of variations are applied.

Consider the following general problem:

Among all twice differentiable curves, $y = \phi(x)$ find the particular $\phi(x)$ which passes through the points (x_0, y_0) and (x_1, y_1) and which causes the integral

$$I\{\phi\} = \int_{x_0}^{x_1} F(x, y, \frac{dy}{dx})\, dx \qquad \text{A.9}$$

to assume a stationary value.

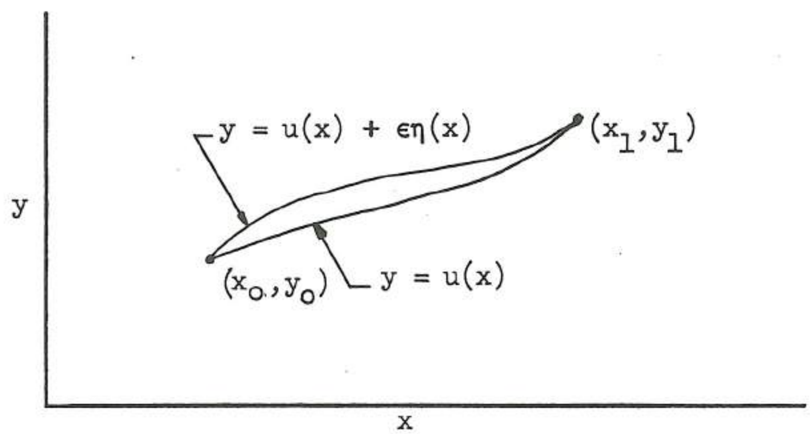
(The requirement for continuous derivatives will be examined later.)

Our procedure in finding that $\phi(x)$ which satisfies the requirements is analogous to the procedure for finding the extreme value of a function. We assume that $\phi = u(x)$ is the particular curve desired. Then any other curve which is admissible, for example

$$\phi = u(x) + \epsilon\eta(x) \qquad \text{A.10}$$

must cause I to move from the extreme value but if I is expanded in powers of ϵ, the first degree term must vanish. For example, if u is the admissible function which minimizes I, then $u + \epsilon\eta$ will cause I to increase, and the charge of I will be at most proportional to ϵ^2. Here ϵ is a constant and $\eta(x)$ is a function with continuous first and second derivations which vanishes at x_0 and x_1.

$$\eta(x_0) = \eta(x_1) = 0 \qquad \text{A.11}$$



We have

$$y = \phi(x) = u + \epsilon\eta(x)$$

$$\frac{dy}{dx} = y' = \phi' = u'(x) + \epsilon\eta'(x) \qquad \text{A.12}$$

and

$$I = \int_{x_0}^{x_1} F(x,y,y')dx = \int_{x_0}^{x_1} F(x,\ u + \epsilon\eta,\ u' + \epsilon\eta')dx$$

This integral is a function of ϵ

$$\psi(\epsilon) = I = \int_{x_0}^{x_1} F(x,\ u + \epsilon\eta,\ u' + \epsilon\eta')dx \qquad \text{A.13}$$

Now we can state in more familiar terms the condition that u(x) is the function which causes I to be stationary. Clearly if this be true then

$$\frac{d\psi}{d\epsilon} = 0 \quad \text{at} \quad \epsilon = 0 \qquad \text{A.14}$$

Remembering that $y = u+\epsilon\eta$, we compute $\frac{d\psi}{d\epsilon}$ as

$$\frac{d\psi}{d\epsilon} = \frac{d}{d\epsilon} \int_{x_0}^{x_1} F(x,y,y')\, dx$$

$$= \int_{x_0}^{x_1} \frac{\partial F}{\partial y}\frac{dy}{d\epsilon} + \frac{\partial F}{\partial y'}\frac{dy'}{d\epsilon}\, dx$$

$$\frac{d\psi}{d\epsilon} = \int_{x_0}^{x_1} (\eta F_y + \eta' F_{y'})\, dx \qquad \text{A.15}$$

where we have written

$$\frac{\partial F(x,y,y')}{\partial y} = F_y \quad \text{etc.}$$

and the conditions on the functions u and η assure that it is permissible to interchange the order of integration and differentiation.

$$\frac{d\psi}{d\epsilon} = \int_{x_0}^{x_1} (\eta F_y + \eta' F_{y'})\, dx$$

$$\left.\frac{d\psi}{d\epsilon}\right|_{\epsilon=0} = \int_{x_0}^{x_1} (\eta F_u + \eta' F_{u'})\, dx = 0 \qquad \text{A.16}$$

Equation A.16 expresses the condition that $y = u(x)$ causes I to be stationary. Equation A.16 can be put in more convenient form if we integrate the last half of the expression by parts.

$$\int_{x_0}^{x_1} \eta' F_{u'} dx = \eta F_{u'} \Big|_{x_0}^{x_1} - \int_{x_0}^{x_1} \eta \frac{d}{dx} F_{u'}\, dx \qquad \text{A.17}$$

Since we required (A.11) that

$$\eta(x_o) = \eta(x_1) = 0 \qquad \text{A.11}$$

$$\int_{x_o}^{x_1} \eta' F_{u'}\, dx = -\int_{x_o}^{x_1} \eta \frac{d}{dx} F_{u'}\, dx \qquad \text{A.18}$$

$$\left.\frac{d\psi}{d\epsilon}\right|_{\epsilon=0} = \int_{x_o}^{x_1} \eta(F_u - \frac{d}{dx} F_{u'})\, dx = 0 \qquad \text{A.19}$$

Now the function $\eta(x)$ is arbitrary, and therefore

$$F_y - \frac{d}{dx} F_{y'} = 0 \quad \text{when } y = u(x) \qquad \text{A.20}$$

Equation A.20 follows from the fact that $\eta(x)$ is arbitrary, in the following way: Suppose that $F_y - \frac{d}{dx} F_{y'} \neq 0$; for example, suppose it is positive in the interval

$$x_3 < x < x_4.$$

Then, since $\eta(x)$ is arbitrary, we can choose $\eta(x) = 0$ $x_o \leqslant x \leqslant x_3$ and also in $x_4 \leqslant x \leqslant x_1$, but let $\eta(x)$ be positive in $x_3 < x < x_4$. Then the integrand in A.19 is zero everywhere in the interval of integration except in $x_3 < x < x_4$ and it is positive throughout this interval. Therefore the integral of A.19 is positive, which violates A.19. and so A.20 must be true.

Equation A.20 is known as Euler's equation, discovered by him in 1744. Written out in full:

$$F_y - \frac{d}{dx} F_{y'} = F_y - F_{y'y'}y'' - F_{yy'}y' - F_{xy'} = 0 \qquad \text{A.20}$$

Euler's equation is an ordinary differential equation of second order. The solutions are called <u>extremals</u> of the variational problem. The particular extremal (if any) which satisfies the boundary conditions is the desired solution of the variational problem. Note that, since A.20 is a second order equation, two arbitrary constants of integration are available for satisfying boundary conditions.

Example:

$$I = \frac{1}{\sqrt{2g}} \int_{x_o}^{x_1} \sqrt{\frac{1 + y'^2}{y}} \, dx \qquad A.8$$

Euler's equation:

$$\left(\sqrt{1 + y'^2}\right) \left(-\frac{1}{2}\right) \left(y^{-3/2}\right) - \frac{d}{dx}\left\{\left(\frac{1}{\sqrt{y}}\right)\left(\frac{1}{2}\right) (1 + y'^2)^{-1/2} \, (2y')\right\} = 0$$

The solutions of this equation are the brachistochrones. In this case, however, the Euler equation can be simplified directly. Note that the explicit dependence on x is missing. When this is true,

$$\frac{d}{dx}\,(y'F_{y'} - F) = 0 \qquad A.21$$

and

$$y'F_{y'} - F = \text{const} \qquad A.22$$

(Since:

$$\frac{d}{dx}\,(y'F_{y'} - F) = y''F_{y'} + y' \frac{d}{dx} F_{y'} - F_{y'}y'' - F_y y'$$

$$= y' \frac{d}{dx} F_{y'} - y' \, Fy$$

$$= y'(\frac{d}{dx} F_{y'} - F_y) = 0$$

$$\therefore \frac{d}{dx} F_{y'} - F_y = 0$$

implies $y'F_{y'} - F = \text{const}$)

The Euler equation for the brachistochrone problem can be written:

$$\sqrt{\frac{1 + y'^2}{y}} - \frac{y'}{\sqrt{y}} \left(y'/\sqrt{1 + y'^2}\right) = \text{const}$$

$$\sqrt{\frac{1 + y'^2}{y}} - \frac{y'^2}{\sqrt{y(1 + y'^2)}} = \text{const}$$

$$\frac{\left(\sqrt{1 + y'^2}\right)^2 - y'^2}{\sqrt{y(1 + y'^2)}} = \text{const}$$

$$\frac{1}{\sqrt{y(1 + y'^2)}} = \text{const} = -\frac{1}{C} \qquad \text{A.23}$$

Let $y = 1/2\ C^2\ (1 - \cos t)$ A.24

from A.23, $y' = \sqrt{\frac{C^2 - y}{y}} = \cot\ (t/2)$

$$x = \int \frac{dy}{y'} = \int \tan \frac{t}{2} \frac{dy}{dt}\, dt \qquad \text{A.25}$$

$$x = \int \tan \frac{t}{2}\ (C^2/2 \sin t)\ dt$$

$$x = \int C^2 \sin^2 \frac{t}{2}\, dt$$

$$\begin{cases} x = \dfrac{C^2}{2}\ (t - \sin t) + \text{const} & \text{A.26} \\ y = \dfrac{C^2}{2}\ (1 - \cos t) & \text{A.24} \end{cases}$$

x and y are given as functions of the parameter, t. The brachistochrones are cycloids described by points on the circumference of a circle of radius $C^2/2$ which rolls on the x axis.

Generalizations of the Simple Variational Problem.

1. Integrals with more than one argument function.

Our first problem concerned the case where

$$I\{\phi\} = \int_{x_0}^{x_1} F(x,y,y')\,dx \qquad \text{A.9}$$

$$y = \phi(x)$$

and we were to find the single function of x, $y = u(x)$ which caused I to assume a stationary value.

Now suppose that there are a number of unknown functions, $y_i = \phi_i(x)$, to be determined. That is, let $F(x,y_1,y_2,\ldots y_n,\ y_1'\ldots y_n')$ be a function of the 2n + 1 arguments, x, $\phi_1\ldots\phi_n$, $\phi_1'\ldots\phi_n'$ which is continuous and has continuous first and second derivatives at least. Since the $y_i = \phi_i(x)$ are all functions of the single variable x, F is a function of the single independent variable, x, and

$$I\{\phi_1\ldots\phi_n\} = \int_{x_0}^{x_1} F\,dx \qquad \text{A.25}$$

is a definite number; the number depending on the choice of the functions ϕ_i. Our problem will be to find the particular set of n functions, $y_i = u_i(x)$ which causes the integral I to assume a stationary value. To show that I has reached a stationary value, we must compare the number, I, obtained by setting $y_i = u_i(x)$ with the number obtained by setting $y_i = \phi_i(x)$ where $u_i(x)$ is the set of functions which satisfies our problem and the $\phi_i(x)$ are all other functions which have the appropriate continuity properties and which satisfy the boundary conditions:

$$\phi_i(x_0) = A_i$$

$$\phi_i(x_1) = B_i \qquad \text{A.26}$$

Our procedure in this case is very similar to the procedure used in obtaining Euler's equation for the simplest problem.

We take $y_i = u_i(x)$ to be the particular set of functions which satisfies our problem. We then imbed this set of functions in a one-parameter family of functions depending on a single parameter, ϵ, as follows:

Let $\eta_1(x) \ldots \eta_n(x)$ be n arbitrary functions of x with continuous first and second derivatives, which vanish at $x = x_o$ and $x = x_1$

$$\eta_i(x_o) = \eta_i(x_1) = 0 \qquad \text{A.27}$$

Then consider

$$\psi(\epsilon) = \int_{x_o}^{x_1} F(x, u_1 + \epsilon\eta_1, \ldots u_n + \epsilon\eta_n, u_1' + \epsilon\eta_1' \ldots u_n' + \epsilon\eta_n')\, dx \qquad \text{A.28}$$

A necessary condition that I be stationary when $y_i = u_i(x)$, (that is, when $\epsilon = 0$), is that

$$\left.\frac{d\psi}{d\epsilon}\right|_{\epsilon=0} = 0 \qquad \text{A.29}$$

This must be true <u>no matter how the functions</u> $\eta_i(x)$ <u>are chosen</u>. We are at liberty, for example, to choose

$$\eta_2 = \eta_3 \ldots = \eta_n = 0$$

Then

$$\frac{d\psi}{d\epsilon} = \int_{x_o}^{x_1} \left(\eta_1 \frac{\partial F}{\partial y_1} + \eta_1' \frac{\partial F}{\partial y_1'}\right) dx \qquad \text{A.30}$$

and this implies, with A.29 above,

$$\frac{\partial F}{\partial y_1} - \frac{d}{dx} \frac{\partial F}{\partial y_1'} = 0 \qquad \text{A.31}$$

A similar result, of course, holds for each of the n functions and the following thecrem may be stated: (Reference a. p. 508).

"A necessary and sufficient condition that the integral $I\{u_1 \ldots u_n\}$ may be stationary is that the n functions $u_i(x)$ shall satisfy the system of Euler's equation

$$F_{u_i} - \frac{d}{dx} F_{u_i'} = 0 \quad (i = 1, 2 \ldots n)."$$

This is a system of n second order differential equations for the n functions $u_i(x)$. All solutions of the system of equations are called "extremals" of the variation problem.

If F does not contain the independent variable, x, explicitly, a first integral of the system of Euler equations is given by

$$F - \sum u_i' F_{u_i'} = \text{const.} \qquad \text{A.32}$$

The proof is very similar to the proof for the simplest case. Differentiation of the above expression with respect to x results in an expression which must be equal to zero if the Euler equations are valid. Details of the proof are left to the reader.

Example: The brachistochrone problem in three dimensions.

We again take gravity to act in the positive y direction. The problem now is to minimize

$$I = \int_{x_0}^{x_1} \sqrt{\frac{1 + y'^2 + z'^2}{y}}\, dx \qquad \text{A.33}$$

$$I = \int_{x_0}^{x_1} F(y, y', z')\, dx \qquad \text{A.34}$$

Euler's equations are:

$$(1) \quad F_z - \frac{d}{dx} F_{z'} = \frac{d}{dx} \frac{z'}{\sqrt{y}} \frac{1}{\sqrt{1 + y'^2 + z'^2}} = 0$$

$$\frac{z'}{\sqrt{y}} \frac{1}{\sqrt{1 + y'^2 + z'^2}} = a \qquad \text{A.35}$$

(2) $F - y'F_{y'} - z'F_{z'} = b$

$$\frac{1}{\sqrt{y}} \quad \frac{1}{\sqrt{1 + y'^2 + z'^2}} = b$$

Dividing (1) by (2), we see that z' = const = C. Therefore the curve for which the integral is constant must lie in the plane

$$z = Cx + d \qquad \text{A.36}$$

Substitution of this expression into either (1) or (2) results in an equation for y which is formally identical to that obtained in the two-dimensional case. The answer is the equation of a cycloid with undetermined constants sufficient to allow the boundary conditions to be satisfied.

Variational Problem with Constraints.

Very often a variational problem is presented where, in addition to the problem of determining a function which causes an integral to assume a stationary value, the resulting function must satisfy subsidiary conditions. That is, the field of functions to be investigated is restricted to functions which satisfy a subsidiary condition. An example is the so-called "isoperimetric" problem:

Find the curve, y(x), which encloses the greatest area in the xy plane, and which has a given length.

Here the given length of the curve imposes a constraint. The general problem of this type is to find the particular function, $y = u(x)$ among all admissible functions $y = \phi(x)$, which causes

$$I\{\phi\} = \int_{x_0}^{x_1} F(x,y,y')\, dx \qquad \text{A.9}$$

to be stationary, the functions ϕ being subject to the further subsidiary condition

$$H\{\phi\} = \int_{x_0}^{x_1} G(x,y,y')\, dx = \text{const.} \qquad \text{A.37}$$

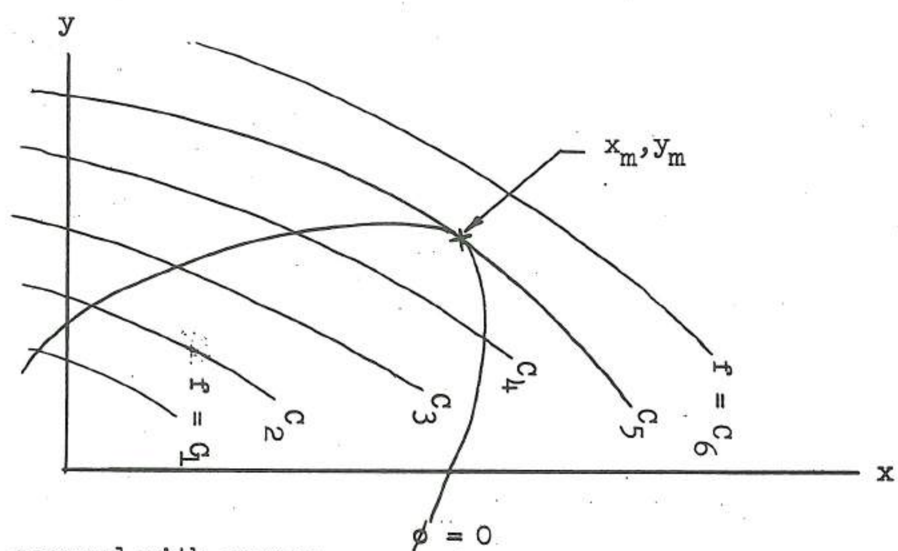
Before developing the means for solving A.9 subject to A.37, we shall briefly review the subject of maxima and minima of functions of several variables. (Reference a, pages 183 - 204.).

Maxima and Minima with Subsidiary Conditions:

We consider the problem of finding stationary values of a function, f(x,y) when the two "independent" variables are not actually independent, but are related by the subsidiary condition

$$\phi(x,y) = 0 \qquad \text{A.38}$$

This is not a fundamentally new problem, since we can (in theory, at least) solve A.38 for one variable in terms of the other, substitute this expression in f(x,y) which then is a function of one variable only, and proceed along well-known lines. It is more convenient and (according to Courant, ref. a) "also more elegant" to preserve the symmetry of the problem and express the conditions for a stationary value in a way which gives no preference to either variable. A very practical reason for preserving the symmetry is that often the subsidiary expression is such that it cannot readily be solved for one variable in terms of the other, or, if solvable, the resulting expression is very clumsy to handle. The problem may be visualized with the aid of the figure.



The plane is covered with curves

$$f(x,y) = C$$

which intersect the curve

$$\phi(x,y) = 0$$

Our problem is to find the coordinates of the point (x_m, y_m) where $f(x,y)$ has reached an extreme value, in this case C_5 and simultaneously $\phi = 0$. At the point (x_m, y_m) the two curves, $f = C$ and $\phi = 0$ will have the same tangent. It will be recalled that the slope of a curve

$$F(x,y) = C$$

is computed as

$$\frac{dy}{dx} = -\frac{F_x}{F_y} \qquad \text{A.39}$$

Thus in the case being considered here, assuming that ϕ_x and ϕ_y do not vanish simultaneously we can write

$$\frac{f_x}{f_y} = \frac{\phi_x}{\phi_y} \qquad \text{A.40}$$

at the point x_m, y_m or:

$$\frac{f_x}{\phi_x} = \frac{f_y}{\phi_y} = -\lambda \qquad \text{A.41}$$

Solving A.41, we have

$$\begin{aligned} f_x + \lambda\phi_x &= 0 \\ f_y + \lambda\phi_y &= 0 \end{aligned} \qquad \text{A.42}$$

Equations A.42, together with the constraint equation

$$\phi(x,y) = 0 \qquad \text{A.38}$$

are three equations in the three unknowns, x_m, y_m and λ.

The above discussion is only intended to make the following rule seem plausible. The rule is proven rigorously in many calculus books, for example Reference a. The factor λ is known as Lagranges' Multiplier and the following rule is known as Lagranges' Method of Undetermined Multipliers. (Ref. a, p. 191.)

"To find the extreme value of the function $f(x,y)$ subject to the subsidiary condition $\phi(x,y) = 0$, we add to $f(x,y)$ the product of $\phi(x,y)$ and an unknown factor λ, independent of x and y, and write down the known necessary conditions

$$f_x + \lambda\phi_x = 0$$

$$f_y + \lambda\phi_y = 0 \qquad \text{A.42}$$

for an extreme value of

$$F = f + \lambda\phi \qquad \text{A.43}$$

Equations A.42 together with $\phi = 0$ are sufficient to determine the coordinates of the extreme value and the constant of proportionality, λ."

Example:

Find the extreme values of

$$u = xy \qquad \text{A.44}$$

on the unit circle with center at the origin, that is, subject to the constraint

$$\phi(x,y) = x^2 + y^2 - 1 = 0 \qquad \text{A.45}$$

Let

$$F = xy + \lambda(x^2 + y^2 - 1) \qquad \text{A.46}$$

$$\left.\begin{aligned} F_x &= y + 2\lambda x = 0 \\ F_y &= x + 2\lambda y = 0 \\ \phi &= x^2 + y^2 - 1 = 0 \end{aligned}\right\} \qquad \text{A.47}$$

Equations A.47 may be solved for x,y,λ. We obtain the four points

$$x = \pm\sqrt{2}/2$$

$$y = \pm\sqrt{2}/2$$

We also find

$$\lambda = \pm \frac{1}{2}$$

and

$$u = xy = \pm \frac{1}{2}$$

I.e., there are four points on the circle where u is stationary--two maxima and two minima.

We will find that Lagrange's method of undetermined multipliers is very useful in handling variational problems with constraints.

Returning now to the variational problem with constraints, we have to find the stationary value of

$$I = \int_{x_0}^{x_1} F(x,y,y')\, dx \qquad \text{A.9}$$

subject to the constraint

$$H = \int_{x_0}^{x_1} G(x,y,y')\, dx = \text{const} \qquad \text{A.37}$$

As usual, we assume that $y = u(x)$ is the curve which satisfies our requirements. Again we form a function which is varied from $u(x)$, but in order to allow freedom to satisfy the constraint equation we introduce two parameters, ϵ_1, ϵ_2 and let $y(x)$ be a member of the two parameter family

$$y = u(x) + \epsilon_1\eta(x) + \epsilon_2\zeta(x) \qquad \text{A.48}$$

where η and ζ are twice differentiable and

$$\eta(x_1) = \eta(x_0) = \zeta(x_0) = \zeta(x_1) = 0 \qquad \text{A.49}$$

Now the two integrals I and H are functions of the two parameters, ϵ_1 and ϵ_2 and

$$\psi(\epsilon_1,\epsilon_2) = \int_{x_0}^{x_1} F(x,u + \epsilon_1\eta + \epsilon_2\zeta,\ u' + \epsilon_1\eta' + \epsilon_2\zeta')\, dx \qquad \text{A.50}$$

must be stationary at $\epsilon_1 = \epsilon_2 = 0$ with respect to small values of ϵ_1 and ϵ_2. ϵ_1 and ϵ_2 are connected by the relation

$$\phi(\epsilon_1,\epsilon_2) = \int_{x_0}^{x_1} G(x,u + \epsilon_1\eta + \epsilon_2\zeta,\ u' + \epsilon_1\eta' + \epsilon_2\zeta')\,dx = \text{const} \qquad \text{A.51}$$

But now Lagrange's method of undetermined multipliers tells us that to find the stationary value of $\psi(\epsilon_1,\epsilon_2)$ subject to the subsidiary condition $\phi(\epsilon_1,\epsilon_2)$ = const we should form the function

$$\phi^* = \psi + \lambda\phi$$

and that

$$\left.\begin{aligned} \psi_{\epsilon_1} + \lambda\phi_{\epsilon_1} &= 0 \\ \psi_{\epsilon_2} + \lambda\phi_{\epsilon_2} &= 0 \end{aligned}\right\} \text{ at } \epsilon_1 = \epsilon_2 = 0 \qquad \text{A.52}$$

are necessary conditions for the existence of a stationary value of ψ. The choice $y = u(x)$ assures that the stationary value occurs at $\epsilon_1 = \epsilon_2 = 0$.

Performing the operations indicated by A.52, and after the usual integration by parts, we obtain

$$\int_{x_0}^{x_1} \left\{\left(F_y - \frac{d}{dx} F_{y'}\right) + \lambda\left(G_y - \frac{d}{dx} G_{y'}\right)\right\} \eta\,dx = 0 \qquad \text{A.53}$$

$$\int_{x_0}^{x_1} \left\{\left(F_y - \frac{d}{dx} F_{y'}\right) + \lambda\left(G_y - \frac{d}{dx} G_{y'}\right)\right\} \zeta\,dx = 0$$

The integrand of the first equation must be zero for any η unless λ depends on η, and if λ depends on η, the second equation cannot be true. That is, the first of equations A.53 only requires that the quantity in curly brackets be orthogonal to the function η. But if it is orthogonal to an arbitrary function η, then it cannot also be orthogonal to a second arbitrary function ζ. We therefore conclude that the integrand is zero and

$$F^*_{y} - \frac{d}{dx} F^*_{y'} = 0 \qquad \text{A.54}$$

where

$$F^* = F + \lambda G$$

The integral of Equation A.54 contains the parameter λ in addition to the two constants of integration. The values of these three constants are determined from the boundary conditions and the constraint Equation A.37.

Example: The isoperimetric problem

$$I = \frac{1}{2} \int_0^{\pi} r^2 \, d\theta \qquad \text{A.55}$$

$$\ell = \int_A^B dS \qquad \text{A.56}$$

Here we have the problem of joining two fixed points, A and B by a plane curve of given length, ℓ, so that the area enclosed by the curve and the chord through A and B is maximum.

$$\ell = \int_A^B dS = \int_0^{\pi} \left\{ r^2 + \left(\frac{dr}{d\theta}\right)^2 \right\}^{1/2} d\theta \qquad \text{A.57}$$

Here

$$F(r, r', \theta) = \frac{1}{2} r^2 \qquad \text{A.58}$$

$$G(r, r', \theta) = (r^2 + r'^2)^{1/2} \qquad \text{A.60}$$

$$F^* = F + \lambda G \qquad \text{A.61}$$

$$F^* = \frac{1}{2} r^2 + \lambda (r^2 + r'^2)^{1/2} \qquad \text{A.62}$$

The Euler equation is

$$r + \frac{\lambda r}{(r^2 + r'^2)^{1/2}} - \frac{d}{d\theta} \left\{ \frac{\lambda r'}{(r^2 + r'^2)^{1/2}} \right\} = 0 \qquad \text{A.63}$$

This leads to

$$\frac{rr_{\theta\theta} - 2r_\theta^2 - r^2}{(r^2 + r_\theta^2)^{3/2}} = \frac{1}{\lambda} \qquad \text{A.64}$$

The left side of A.64 is the expression for the curvature in polar coordinates. A.64 therefore states that the curvature, $1/p$, is to be constant, and the required curve is a circular arc of radius λ passing through the two end points.

Problem:

Prove that the sphere is the solid of revolution which, for a given surface area, has maximum volume. Take

$$\text{Area, } A = 2\pi \int y \, ds = 2\pi \int_0^a y(1 + y_x)^2 \, dx$$

$$\text{Volume, } V = \pi \int_0^a y^2 \, dx$$

VARIATIONAL PROBLEMS WITH CONSTRAINTS, CTD.

In the last section the constraint was expressed in integral form. Numerous other types of constraints arise, a few of which are considered below.

Ordinary Constraints

We wish to find a curve, $x(t)$, $y(t)$, $z(t)$ ($t_o \leqslant t \leqslant T$) in three-dimensional space, subject to the constraint that the curve shall lie on a given surface, $G(x, y, z) = 0$ and shall pass through two given points A and B on that surface. We wish to optimize

$$I = \int_{t_o}^{T} F(x,y,z,\dot{x},\dot{y},\dot{z})dt$$

Subject to the constraint

$$G(x,y,z) = 0 \qquad \text{A.65}$$

We might immediately reduce this problem to that considered on pages A-3 to A-5. If we solve the relation $G(x,y,z) = 0$ for one of the variables, for example z, in terms of the other two,

$$z = g(x,y) \qquad \text{A.66}$$

Then A.65 can be written:

$$I = \int_{to}^{T} F[x,y,g(x,y),\dot{x},\dot{y},\dot{x}g_x + \dot{y}g_y]dt$$

$$I = \int_{to}^{T} H(x,y,\dot{x},\dot{y})dt \qquad \text{A.67}$$

Our two Euler expressions are:

$$\frac{d}{dt} H_{\dot{x}} - H_x = \frac{d}{dt} F_{\dot{x}} - F_x + \frac{d}{dt}(F_{\dot{z}}g_x) - F_z g_x - F_{\dot{z}}\frac{\partial \dot{z}}{\partial x} = 0$$

$$\frac{d}{dt} H_{\dot{y}} - H_y = \frac{d}{dt} F_{\dot{y}} - F_y + \frac{d}{dt}(F_{\dot{z}}g_y) - F_z g_y - F_{\dot{z}}\frac{\partial \dot{z}}{\partial y} = 0 \qquad \text{A.68}$$

In the last term of each of A.68 we can interchange the order of differentiation

$$\frac{d}{dt} g_x = \frac{\partial \dot{z}}{\partial x}, \qquad \frac{d}{dt} g_y = \frac{\partial \dot{z}}{\partial y} \qquad \text{A.69}$$

Therefore:

$$\frac{d}{dt} F_{\dot{x}} - F_x + g_x \left(\frac{d}{dt} F_{\dot{z}} - F_z\right) = 0$$

$$\frac{d}{dt} F_{\dot{y}} - F_y + g_y \left(\frac{d}{dt} F_{\dot{z}} - F_z\right) = 0 \qquad \text{A.70}$$

Now G_z and $d/dt\, F_{\dot{z}} - F_z$ are both functions of time. <u>If</u> we can find a function of time $\lambda(t)$ such that

$$\frac{d}{dt} F_{\dot{z}} - F_z = \lambda G_z \qquad \text{A.71}$$

then, using the fact that

$$g_x = -\frac{G_x}{G_z}, \qquad g_y = -\frac{G_y}{G_z} \qquad \text{A.72}$$

we have from A.70, A.71, and A.72

$$\left.\begin{aligned} \frac{d}{dt} F_{\dot{x}} - F_x &= \lambda G_x \\ \frac{d}{dt} F_{\dot{y}} - F_y &= \lambda G_y \\ \frac{d}{dt} F_{\dot{z}} - F_z &= \lambda G_z \end{aligned}\right] \qquad \text{A.73}$$

Thus, a necessary condition for an extreme value of I subject to the constraint $G = 0$ is the existence of a function $\lambda(t)$ such that equations A.73 are satisfied in addition to the constraint equation. We have four equations in the four unknown functions of time, x, y, z, λ.

Non-holonomic Constraints

The two types of constraint considered so far (integral and ordinary) are special cases of a more general type. If the constraint is a non-integrable differential expression of the form

$$G(x,y,z,y',z') = 0 \qquad \text{A.74}$$

which cannot be obtained by differentiating an expression

$$H(x,y,z) = 0 \qquad \text{A.75}$$

it is called a non-holonomic constraint.

Most differential equations are non-holonomic constraints. For example:

$$y' - x = 0 \qquad \text{A.76}$$

is a holonomic constraint, since it can be obtained by differentiating

$$y - \frac{x^2}{2} = \text{const.} \qquad \text{A.77}$$

However,

$$y' - z = 0 \qquad \text{A.78}$$

is a non-holonomic constraint, since we cannot find an expression like A.77 which can be differentiated to obtain A.78.

We state without proof the following general rule, which applies to holonomic and non-holonomic constraints: (Ref. Courant-Hilbert, pp.221, 222).

"If the solution does not satisfy the Euler equations corresponding to the expression G, then there exists a multiplier, $\lambda(x)$ such that the Euler equations corresponding to the expression

$$F^* = F + \lambda G \qquad \text{A.79}$$

are satisfied."

Let us apply the above rule to the variational problem with an integral constraint, considered on pages A-14 to A-15. Let

$$z' - G = 0 \qquad \text{A.80}$$

be the constraint equation, equivalent to A.37. Written this way, it is seen that the problem on page A-14 is a special case of optimizing the integral

$$I = \int_{x_0}^{x_1} F(x,y,z,y',z')\, dx, \qquad \text{A.81}$$

Subject to the constraint

$$G(x,y,z,y',z') = 0 \quad . \qquad \text{A.82}$$

In this special case, z and z' are missing from F, and z is missing from G. Applying the general rule, we form the function

$$F^* = F(x,y,y') + \lambda(x)(z' - G(x,y,y')) \qquad A.83$$

and write the Euler equations

$$\frac{d}{dx}\frac{\partial F^*}{\partial y'} - \frac{\partial F^*}{\partial y} = 0$$

$$\frac{d}{dx}\frac{\partial F^*}{\partial z'} - \frac{\partial F^*}{\partial z} = 0 \qquad A.84$$

The second equation is:

$$\frac{d}{dx}\lambda = 0$$

$$\lambda = \text{const.} \qquad A.85$$

and the first equation then becomes the same as A.54.

Extension of the Simple Variational Problem
(Reference C, pp 208-210)

For the simplest variational problem--discussed on pages A-3 to A-5--the Euler expression was derived by imbedding the extremal, $y = u(x)$, in a one parameter family of functions, $y = u + \epsilon\eta$. The fact that this parameter, ϵ, occurs linearly was merely convenient for the development, not necessary. More generally, we may wish to define a family of functions, $y(x,\epsilon)$, with $y(x,0) = u(x)$ and

$$\eta(x) = \frac{\partial}{\partial\epsilon} y(x,\epsilon)\bigg|_{\epsilon=0} \qquad A.86$$

If we introduce the terminology

$$\delta y = \epsilon\eta(x) \qquad A.87$$

and

$$\delta I = \epsilon \left. \frac{d\psi}{d\epsilon} \right|_{\epsilon=0} \qquad \text{A.88}$$

where $\psi(\epsilon)$ is defined similarly to A.13, then

$$\delta I = \epsilon \left. \frac{d\psi}{d\epsilon} \right|_{\epsilon=0} = \epsilon \int_{x_0}^{x_1} (\eta F_y + \eta' F_{y'}) dx \qquad \text{A.89}$$

After the usual integration by parts,

$$\delta I = \epsilon \int_{x_0}^{x_1} (F_y - \frac{d}{dx} F_{y'}) \eta dx + \epsilon F_{y'} \eta \Big|_{x=x_0}^{x=x_1} \qquad \text{A.90}$$

$$\delta I = \int_{x_0}^{x_1} (F_y - \frac{d}{dx} F_{y'}) \delta y dx + F_{y'} \delta y \Big|_{x=x_0}^{x=x_1} \qquad \text{A.91}$$

The expression A.91 is called the first variation of the integral I even if $\eta(x)$, hence δy, does not vanish at the boundary. Suppose that we do not impose conditions on y at the boundaries, $x = x_0$ and $x = x_1$. A.91 still is a necessary condition for I to be stationary. But if I is stationary in this case with arbitrary δy at the boundaries, it is certainly stationary in the more restricted case where δy vanishes at the boundary as in A.13. This then implies that the Euler expression vanishes:

$$F_y - \frac{d}{dx} F_{y'} = 0 \qquad \text{A.92}$$

and consequently because of the arbitrariness of δy everywhere, including the boundaries:

$$F_{y'} \Big|_{\substack{x=x_0 \\ x=x_1}} = 0 \qquad \text{A.93}$$

Now, suppose that the boundary values of y occur explicitly in the expression to be optimized:

$$I = \int_{x_0}^{x_1} F(x,y,y')dx - \phi(y_0) + \psi(y_1) \tag{A.94}$$

where

$$y_0 = y(x_0)$$
$$y_1 = y(x_1) \tag{A.95}$$

and y_0 and y_1 are not prescribed. This case occurs very frequently in trajectory problems, where we often are interested only in the terminal value of some quantity, for example, the perigee altitude or the horizontal component of velocity, etc.

For I to be stationary, we require that the first variation of A.94 vanish:

$$\delta I = \int_{x_0}^{x_1} (F_y - \frac{d}{dx} F_{y'})\delta y dx + \left[\psi' + F_{y'}\Big|_{x=x_1}\right] \delta y_1$$
$$- \left[\phi' + F_{y'}\Big|_{x=x_0}\right] \delta y_0 = 0 \tag{A.96}$$

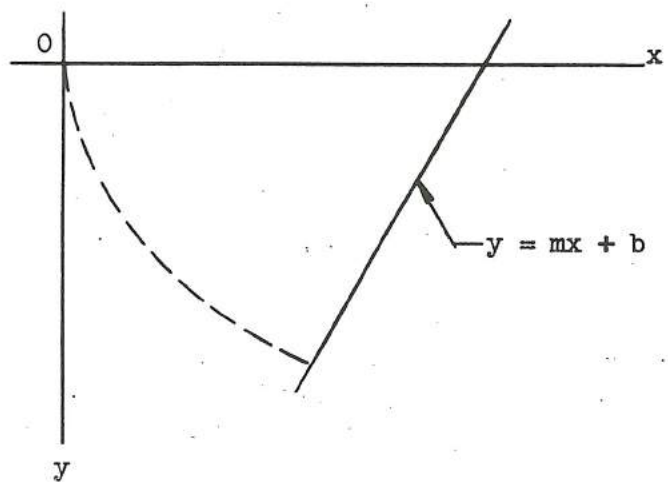
By the same reasoning used in obtaining A.92 and A.93 we obtain:

$$\left.\begin{aligned} F_y - \frac{d}{dx} F_{y'} &= 0 \\ (\psi' + F_{y'})\Big|_{x=x_1} &= 0 \\ (\phi' + F_{y'})\Big|_{x=x_0} &= 0 \end{aligned}\right] \tag{A.97}$$

The latter two expressions are called natural boundary conditions.

The variational problem with variable limits of integration.

Up to now we have considered only variational problems in which the limits of integration were fixed. Now consider the brachistochrone problem where the curve, $y(x)$ is to join a given fixed point, the origin, and a given line $y = mx + b$.



This problem is representative of a class of problems in which the interval of integration is variable. We will confine ourselves to the case of one independent and one dependent variable only.

Let

$$I = \int_a^b F(x,y,y')dx \qquad \text{A.98}$$

where A and B, the end points of the arc of integration have abscissas a and b, respectively, a and b are not necessarily constant. In finding the conditions for a stationary value of I we will vary not only the arc $y = u(x)$ which joins A and B, but we shall also allow B to move along a curve Γ_2 whose equation is

$$y = g_2(x) \qquad \text{A.99}$$

We will keep the point A fixed but the results can be generalized to allow for the case where A can be displaced along the curve Γ_1, whose equation is

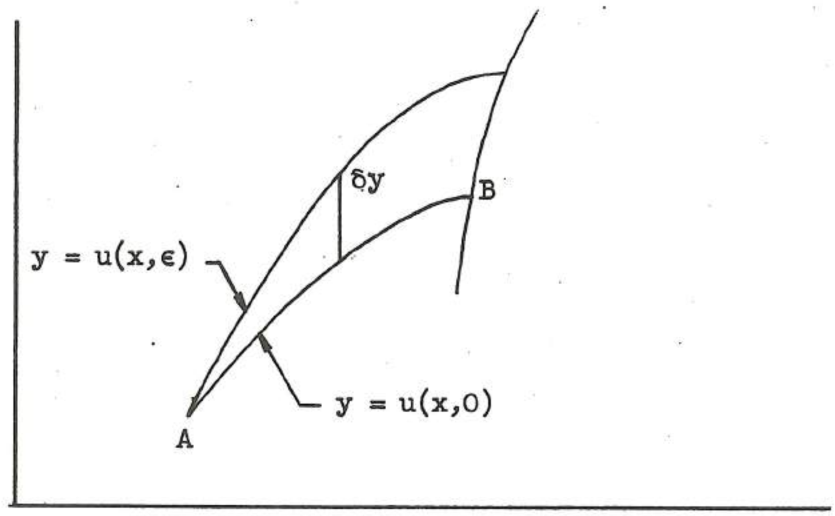
$$y = g_1(x) \qquad \text{A.100}$$

When the end points of the interval were fixed, we varied the function $y(x)$ from the extremal $u(x)$ by constructing the one parameter family

$$y = u(x) + \epsilon\eta(x).$$

Here we will vary $y(x)$ in the more general manner indicated by A.86, A.87.

For the extremal we take $y = u(x,0)$ and for the varied curve we take $y = u(x,\epsilon)$.



The criterion for a stationary value of I is $\delta I = 0$. When the end point B is allowed to vary, we have

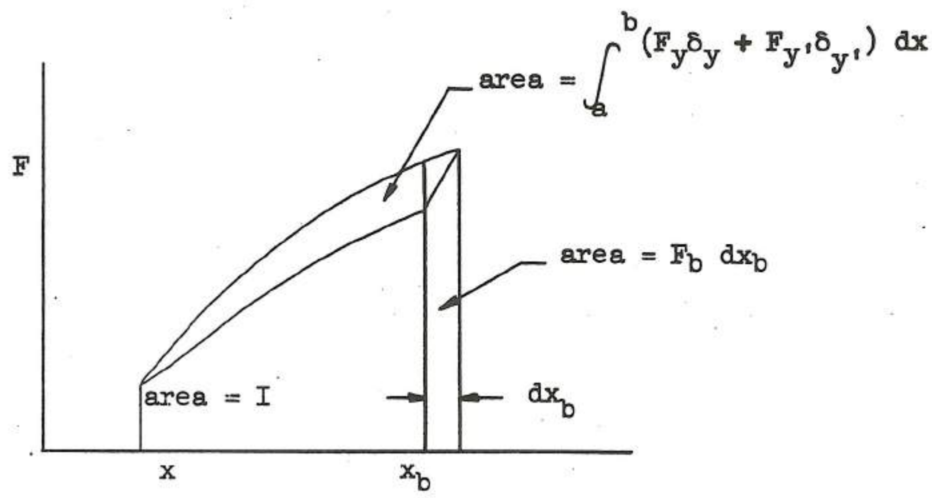
$$\delta I = \int_a^{b+dx_b} F(x, y + \delta y, y' + \delta y')\, dx \qquad \text{A.101}$$

$$- \int_a^b F(x,y,y')\, dx$$

$$\delta I = F_b dx_b + \int_a^b (F_y \delta_y + F_{y'} \delta_{y'})\, dx \qquad \text{A.102}$$

(A.102 can be obtained by differentiating $\int_a^b F\, dx$ with respect to ϵ, allowing for the fact that b is a function of ϵ.)

A.102 can be seen geometrically from the following figure.



A.102 can be integrated by parts

$$\delta I = F_b dx_b + \left.\frac{\partial F}{\partial y'}\right)_b \delta y_b$$

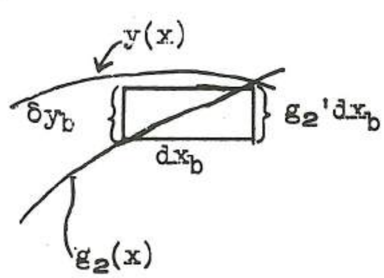
$$+ \int_a^b \delta_y \left\{ F_y - \frac{d}{dx} F_{y'} \right\} dx \qquad \text{A.103}$$

Here δy_b is $u(b,\epsilon) - u(b,0)$ and the subscript b indicates values corresponding to $x = b$. Since I must be stationary even if B is fixed, i.e., if the first two terms of A.103 are zero, we can repeat our earlier argument used in obtaining A.93 and conclude that

$$F_y - \frac{d}{dx} F_{y'} = 0 \qquad \text{A.104}$$

This condition, however, is not sufficient. If B varies, we must in addition require that

$$F_b dx_b + \left.\frac{\partial F}{\partial y'}\right)_b \delta y_b = 0 \qquad \text{A.105}$$



From the figure, it is apparent that for small dx_b

$$\delta y_b = g_2' \, dx_b - y' \, dx_b \qquad \text{A.106}$$

so that A.75 becomes

$$\left[F + (g_2' - y') \frac{\partial F}{\partial y'} \right]_{x=b} dx_b = 0 \qquad \text{A.107}$$

Thus in addition to A.105 we require, since dx_b is arbitrary,

$$\left\{ F + (g_2' - y') \frac{\partial F}{\partial y'} \right\}_{x=b} = 0 \qquad \text{A.109}$$

This equation is known as a transversality condition and the curve Γ_2 is said to be transversal to the extremal at B. The same development, of course, can be used when the lower limit of integration is variable. If the end point A can be displaced along the curve Γ_1, where the equation of Γ_1 is $y = g_1(x)$, in addition to B being variable along Γ_2, the following theorem applies: (Reference b, page 213).

"If the end points A and B of the range of integration of the integral $I = \int_a^b F(x,y,y')dx$ can be displaced along prescribed curves, then I is stationary when the following necessary conditions are satisfied:

(1) y, the ordinate of the extremal, satisfies the Eulerian equation

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} = 0$$

(2) at $x = a$, $F + (g_1' - y') \frac{\partial F}{\partial y'} = 0$

where a is the abscissa of the end point A, which can be displaced along the curve $y = g_1(x)$

(3) at $x = b$, $F + (g_2' - y') \frac{\partial F}{\partial y'} = 0$

where b is the abscissa of the end point B, which can be displaced along the curve $y = g_2(x)$. In these equations, y' is the slope of the extremal and g_1', g_2' are respectively the slopes of the displacement curves of A and B at $x = a$ and $x = b$."

Problem:

$$I = \int_A^B G(x,y)(1 + y'^2)^{1/2} \, dx$$

A and B are both variable. Prove that the extremal and Γ_1 intersect orthogonally at A, and that the extremal and Γ_2 intersect orthogonally at B. Orthogonality is shown by

$$g'y' + 1 = 0$$

THE EULER EXPRESSION AS A GRADIENT IN FUNCTION SPACE

If we have a function of n variables, $f(x_1, x_2 \ldots x_n)$, the condition for f to be stationary is expressed as

$$\text{grad } f = 0 \qquad \text{A.110}$$

where grad f is the n component vector

$$f_{x1} \vec{1}_{x1} + f_{x2} \vec{1}_{x2} + \ldots f_{xn} \vec{1}_{xn} \qquad \text{A.111}$$

If the n variables, x_i are all functions of a parameter, t, we compute the rate of change of f as

$$\dot{f}(t) = \sum_{i=1}^{n} \dot{x}_i f_{x_i} = \vec{V} \cdot \text{grad } f \qquad \text{A.112}$$

where $\vec{V}$ is the velocity vector with components $\dot{x}_i$. Similarly, the Euler expression which vanishes if the functional is stationary, may be regarded as the gradient of the functional in function space (also called Hilbert space).

For example, in the functional

$$I\ \{\Phi\} = \int_{x_0}^{x_1} F(x,y,y')dx \qquad \text{A.113}$$

(where $y = \Phi(x)$ are the admissible functions) assume that the $\Phi(x)$ are also functions of a parameter t (as they are functions of a parameter ϵ in A.12 and A.86).

$$y = \Phi(x,t) \qquad \text{A.114}$$

Then $I = I(t)$ as in A.13. Then, as in A.19,

$$\frac{dI}{dt} = \int_{x_2}^{x_1} \dot{\Phi}(F_y - \frac{d}{dx} F_{y'})dx \qquad \text{A.115}$$

This expression, A.115, is completely analogous to A.112 for the function $f(x_1,x_2 \ldots x_n)$. In general, (Courant-Hilbert, p. 223) one may define the gradient of a functional, $I\{\Phi\}$ as an expression $G\{\Phi\}$ such that if a parameter, t, is introduced to make the functions Φ a one-parameter family of functions, the relation

$$\frac{d}{dt} I\{\Phi\} = \int_{x_2}^{x_1} \dot{\Phi}\ G\{\Phi\}\ dx \qquad \text{A.116}$$

holds.

Variational Problems in the Canonical Form

(Ref.: (a) Whittaker, pp 265-267
(b) Courant-Hilbert, pp 238-242)

In the treatment of mechanics problems, based on Hamilton's Principle, Chapter III, we found that it was advantageous to transform the integrand--the Lagrangean--by means of a Legendre transformation, the resulting Euler equations of the variational problem then were in canonical form. With the equations of motion in canonical, or Hamiltonian, form we could apply the Hamilton-Jacobi theory and make use of all the results of the transformation theory of mechanics. This procedure can be applied to any variational problem, with one independent variable; the general method is due to Ostrogradsky (see Whittaker, pp 265-267).

Let $L(t, y, \dot{y}, \ddot{y} \ldots \overset{(m)}{y}, z, \dot{z}, \ddot{z} \ldots \overset{(n)}{z})$ be a function of the independent variable, t, and the two dependent variables y, z, and their derivatives up to order m and n, respectively. The following development applies to any number of dependent variables, if extended in the obvious manner.

By straight-forward extension of the proofs on pages A-3 and A-9 (see, for example, Reference A, pp 512-513), it is easy to show that the Euler expressions which furnish the conditions under which the integral

$$I = \int_{t_0}^{t_1} L \, dt \qquad \text{A.117}$$

may be stationary are:

$$\left.\begin{aligned} \frac{\partial L}{\partial y} - \frac{d}{dt}\frac{\partial L}{\partial \dot{y}} + \ldots + (-1)^m \frac{d^m}{dt^m}\frac{\partial L}{\partial \overset{m}{y}} &= 0 \\ \frac{\partial L}{\partial z} - \frac{d}{dt}\frac{\partial L}{\partial \dot{z}} + \ldots + (-1)^n \frac{d^n}{dt^n}\frac{\partial L}{d \overset{n}{z}} &= 0 \end{aligned}\right] \qquad \text{A.118}$$

We let

$$\left.\begin{aligned} p_1 &= \frac{\partial L}{\partial \dot{y}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{y}} + \ldots + (-1)^{m-1} \frac{d^{m-1}}{dt^{m-1}}\left(\frac{\partial L}{\partial \overset{m}{y}}\right) \\ p_2 &= \frac{\partial L}{\partial \ddot{y}} - \frac{d}{dt}\frac{\partial L}{\partial \dddot{y}} + \ldots + (-1)^{m-2} \frac{d^{m-2}}{dt^{m-2}}\left(\frac{\partial L}{\partial \overset{m}{y}}\right) \\ &\vdots \\ p_m &= \frac{\partial L}{\partial \overset{m}{y}} \end{aligned}\right] \qquad \text{A.119}$$

$$\left.\begin{aligned} p_{m+1} &= \frac{\partial L}{\partial \dot{z}} - \frac{d}{dt}\frac{\partial L}{\partial \ddot{z}} + \ldots + (-1)^{n-1} \frac{d^{n-1}}{dt^{n-1}}\left(\frac{\partial L}{\partial \overset{n}{z}}\right) \\ p_{m+2} &= \frac{\partial L}{\partial \ddot{z}} - \ldots\ldots\ldots\ldots + (-1)^{n-2} \frac{d^{n-2}}{dt^{n-2}}\left(\frac{\partial L}{\partial \overset{n}{z}}\right) \\ p_{m+n} &= \frac{\partial L}{\partial \overset{n}{z}} \end{aligned}\right] \qquad \text{A.119}$$

Further, let

$$\left.\begin{aligned} q_1 &= y, \quad q_2 = \dot{y}, \;\ldots.\; q_m = \overset{(m-1)}{y} \\ q_{m+1} &= z, \quad q_{m+2} = \dot{z}, \;\ldots.\; q_{m+n} = \overset{(n-1)}{z} \end{aligned}\right] \qquad \text{A.120}$$

We now define the Hamiltonian function

$$\left.\begin{aligned} H = &-L + p_1 q_2 + p_2 q_3 + \ldots. + p_{m-1} q_m \\ &+ p_m \overset{(m)}{y} + p_{m+1} q_{m+2} + \ldots. + p_{m+n-1} q_{m+n} \\ &+ p_{m+n} \overset{(n)}{z} \end{aligned}\right] \qquad \text{A.121}$$

where H is expressed as a function of t, q, ... q_{m+n}, p_1, ..., p_{m+n}. The quantities $\overset{(m)}{y}$ and $\overset{(n)}{z}$ are eliminated by use of the expressions $p_m = \partial L/\partial \overset{(m)}{y}$, $p_{m+n} = \partial L/\partial \overset{(n)}{z}$. Note that this definition of the Hamiltonian is more general (because of the more general p_i) than our earlier definition

$$H = \sum_i p_i \dot{q}_i - L$$

However, if only velocities, or first derivatives of the y_i and z_i appear in the Lagrangean, then the definitions are the same. Now if we calculate the variation of H,

$$\delta H = - \sum_{r=0}^{m-1} \frac{\partial L}{\partial \overset{r}{y}} \delta q_{r+1} - \frac{\partial L}{\partial \overset{m}{y}} \delta \overset{m}{y} - \sum_{r=0}^{n-1} \frac{\partial L}{\partial \overset{r}{z}} \delta q_{m+r+1}$$

$$- \frac{\partial L}{\partial \overset{n}{z}} \delta \overset{n}{z} + \sum_{r=1}^{m-1} p_r \, \delta q_{r+1} + p_m \, \delta \overset{m}{y}$$

$$+ \sum_{r=1}^{m-1} q_{r+1} \, \delta p_r + \overset{m}{y} \, \delta p_m \qquad \text{A.122}$$

$$+ \sum_{r=m+1}^{m+n-1} p_r \, \delta q_{r+1} + p_{m+n} \, \delta \overset{n}{z} + \sum_{r=m+1}^{m+n-1} q_{r+1} \, \delta p_r$$

$$+ \overset{n}{z} \, \delta p_{m+n}$$

Since

$$\frac{\partial L}{\partial y} = \dot{p}_1 \qquad \frac{\partial L}{\partial \dot{y}} = \dot{p}_2 + p_1$$

$$\frac{\partial L}{\partial \ddot{y}} = \dot{p}_3 + p_2 \; \cdots \; \frac{\partial L}{\partial \overset{m}{y}} = p_m$$

A.122 reduces to:

$$\delta H = - \sum_{r=1}^{m+n} \dot{p}_r \, \delta q_r + \sum_{r=1}^{m+n} \dot{q} \, \delta p_r \qquad \text{A.123}$$

and therefore, since

$$\delta H = \sum_{r=1}^{m+n} \frac{\partial H}{\partial q_r} \delta q_r + \sum_{r=1}^{m+n} \frac{\partial H}{\partial p_r} \delta p_r \qquad \text{A.124}$$

We have:

$$\frac{dq_r}{dt} = \frac{\partial H}{\partial p_r} \qquad \frac{dp_r}{dt} = - \frac{\partial H}{\partial q_r} \tag{A.125}$$

Thus the differential equations--the Euler relations--are expressed in canonical or Hamiltonian form.

As an example, consider the problem of maximizing some integral I;

$$I = \int_{t_o}^{t_1} F(x_1, x_2, \dot{x}_1, \dot{x}_2, t) \tag{A.126}$$

Subject to the equations of motion as constraints. Since the equations of motion are of second order, the constraint equations will be of the form

$$G(x_1, x_2, \dot{x}_1, \dot{x}_2, \ddot{x}_1, \ddot{x}_2) = 0 \tag{A.127}$$

According to the Lagrange Multiplier rule, we should form the function

$$F^* = F + \lambda(t)G \tag{A.128}$$

Thus

$$F^* = F^*(x_1, x_2, \dot{x}_1, \dot{x}_2, \ddot{x}_1, \ddot{x}_2, t) \tag{A.129}$$

Then according to Ostrogradsky's rule we should take:

$$\begin{aligned} q_1 &= x_1 & p_1 &= \frac{\partial F^*}{\partial \dot{x}_1} - \frac{d}{dt}\frac{\partial F^*}{\partial \ddot{x}_1} \\ q_2 &= \dot{x}_1 & p_2 &= \frac{\partial F^*}{\partial \ddot{x}_1} \\ q_3 &= x_2 & p_3 &= \frac{\partial F^*}{\partial \dot{x}_2} - \frac{d}{dt}\frac{\partial F^*}{\partial \ddot{x}_2} \\ q_4 &= \dot{x}_2 & p_4 &= \frac{\partial F^*}{\partial \ddot{x}_2} \end{aligned} \tag{A.130}$$

We form the Hamiltonian (here m = n = 2)

$$H = - F^* + p_1 q_2 + p_3 q_4 + p_2 \dot{x}_1 + p_4 \dot{x}_2 \qquad \text{A.131}$$

Since

$$p_2 = \frac{\partial F^*}{\partial \dot{x}_1}, \qquad p_4 = \frac{\partial F^*}{\partial \dot{x}_2}$$

we can compute these derivatives, solve the resulting expressions for $\dot{x}_1$, and $\dot{x}_2$ and eliminate $\dot{x}_1$ and $\dot{x}_2$ from H. Then the canonical equations of motion are:

$$\frac{dq_1}{dt} = \frac{\partial H}{\partial p_1} \qquad \frac{dp_1}{dt} = - \frac{\partial H}{\partial q_1}$$

$$\frac{dq_2}{dt} = \frac{\partial H}{\partial p_2} \qquad \frac{dp_2}{dt} = - \frac{\partial H}{\partial q_2} \qquad \text{A.132}$$

Note that these canonical equations of motion should contain only a single time derivative on the left-hand side, in each equation, i.e., they should be of the form:

$$\dot{q}_1 = f_1(q_1, q_2, p_1, p_2, t)$$

$$\dot{q}_2 = f_2(q_1, q_2, p_1, p_2, t)$$

$$\dot{p}_1 = f_3(q_1, q_2, p_1, p_2, t) \qquad \text{A.133}$$

$$\dot{p}_2 = f_4(q_1, q_2, p_1, p_2, t)$$

All the time derivatives on the right side should be eliminated by means of the expressions $\dot{q}_1 = q_2$, etc.

Numerical Methods. Reference (b) Chapter VII.
Reference (c) Chapter IV, pp 174 et. seq.

There is a relatively small number of analytical solutions to variational problems known. Consequently, numerical methods have been used extensively.

When faced with the necessity of solving a variational problem numerically one usually has the choice of attacking the problem directly or of reducing it to a differential equation (Euler's) and solving the differential equation numerically. (This is called the indirect method of the calculus of variations.) Since the class of solutions of Euler's equation is very restricted compared with the class of all functions which must be tried in the integral equation, and since machine methods of solution of differential equations are well established, it often is preferable, if possible, to deduce the Euler equation and solve it, rather than the original stationary value problem.

There are, however, several methods of attacking the problem directly which are well developed and which are well suited to machine methods. Two of these methods will be discussed very briefly here. For a more complete discussion, references b and c may be consulted and there is an extensive body of literature on direct solution of variational problems.

The Rayleigh-Ritz Method.

The object of this method is to replace the variational problem by that of finding extreme values of functions of several variables.

First it is assumed that y can be expressed in terms of known functions of x. For example, y might be assumed to be expandible in a power series, or in a Fourier series. On substituting the assumed expression for y in the integral, the integral can be evaluated, the coefficients in the expression for y remaining to be evaluated. By the usual methods of the calculus, the coefficients can be adjusted to maximize (or minimize) the integral.

Example:

$$I = \int_{-1}^{1} (1 - x^2)(y'^2)\, dx \qquad \text{A.134}$$

subject

$$\int_{-1}^{1} y^2\, dx = 1 \qquad \text{A.135}$$

Let us assume that y can be expanded in a power series and that the first three terms will give reasonable accuracy for our purposes

$$y = a + bx + cx^2 \qquad \text{A.136}$$

Substituting our assumed expression for y in A.134 and A.135

$$I = \frac{4}{3}\left(b^2 + \frac{4}{5}c^2\right) \qquad \text{A.137}$$

$$1 = 2\left(a^2 + \frac{b^2}{3} + \frac{2ac}{3} + \frac{c^2}{5}\right) \qquad \text{A.138}$$

A.137 is a function of three variables, a, b, c (a is missing) which is to be made stationary subject to the constraint equation A.138. But this is precisely the problem which we treated on pages A-11 to A-14. Our condition for a stationary value of I is, according to Equation A.51

$$\frac{\partial I/\partial a}{\partial \phi/\partial a} = \frac{\partial I/\partial b}{\partial \phi/\partial b} = \frac{\partial I/\partial c}{\partial \phi/\partial c} = -\lambda \qquad \text{A.139}$$

Evaluation of A.139 results in:

$$\frac{0}{4(a + c/3)} = \frac{8b/3}{4b/3} = \frac{32c/15}{4(a/3 + c/5)} = -\lambda \qquad \text{A.140}$$

The possible solutions are:

(1) $\lambda = -2, \quad a = c = 0$

(2) $\lambda = 0, \quad b = c = 0$

(3) $\lambda = -6, \quad b = 0 \quad c = -3a$

From A.138 we have for (1)

$b = \sqrt{3/2}$, for (2) $a = 1/\sqrt{2}$ and for (3) $a = \sqrt{5/8}$ and finally

$$\left.\begin{aligned} -\lambda &= 0 \quad y = 1/\sqrt{2} \\ -\lambda &= 2 \quad y = \sqrt{3/2x} \\ -\lambda &= 6 \quad y = \sqrt{5/8}\,(1 - 3x^2) \end{aligned}\right\} \qquad \text{A.141}$$

These functions are the first three Legendre functions, except for a constant multiplier

$$y = \left(\frac{2n + 1}{2}\right)^{1/2} P_n(x) \qquad n = 0,1,2 \qquad \text{A.142}$$

Problem:

Solve the above problem analytically, not by Rayleigh-Ritz.

Hint: Take $\lambda = n(n+1)$. Lengendre's equation is $\frac{d}{dx}\left\{\left(1+x^2\right)\frac{dy}{dx}\right\} + n(n+1)y = 0$.

In this particular case, the numerical method results in an exact solution of the variational problem. This, of course seldom happens except in text-books and for carefully doctored problems. In the real world, the result usually will be an approximation, the closeness of the result depending on the selection of the approximating functions. For this reason, before attempting a solution by the Rayleigh-Ritz method (or any other numerical method, for that matter) it is highly advisable carefully to investigate the problem analytically so that at least the gross characteristics of the solution are known.

The problem of investigating the degree of approximation achieved is highly important, it also can be very difficult. One indication is the closeness to which the left side of the Eulerian equation approaches zero. If it vanishes identically throughout the interval then the solution is exact.

Galerkin's Method.

The Rayleigh-Ritz method attacks the problem by converting it to an ordinary stationary value problem. The Galerkin method uses the condition for a stationary value, but does not convert the problem.

On page A-5, equation A.19, we showed that a necessary condition for a stationary value of I is that

$$\int_{x_0}^{x_1} \eta\left(Fu - \frac{d}{dx} Fu'\right) dx = 0 \qquad \text{A.19}$$

If u is an exact solution of the problem, then A.19 is true for any arbitrary η which satisfies the conditions of the problem. If u is not an exact solution of Euler's equation, then the quantity in parentheses (Euler's expression) does not vanish identically throughout the interval and A.19 is not satisfied by arbitrary η. Let us choose

$$y_n = \sum_{m=1}^{n} a_m f_m(x) \qquad \text{A.143}$$

as an approximate solution of Euler's equation A.20. Since A.19 is true for arbitrary η, we choose $\eta = f_m(x)$, $m = 1 \ldots n$. We substitute our approximate solution y_n in A.19 and obtain

$$\int_{x_0}^{x_1} f_m(x) \left\{\frac{\partial F}{\partial y_n} - \frac{d}{dx}\frac{\partial F}{\partial y'_n}\right\} dx = 0 \qquad A.144$$

Because y_n is not an exact solution of Euler's equation, the integrand does not, in general, vanish. In fact, it contains n arbitrary constants, a, and after the integration is carried out, the result is an equation in the n constants, a.

$$g(a_1, a_2 \ldots a_n) = 0 \qquad A.145$$

Since there are n different functions, f_m, we can determine n equations in the n unknown constants and satisfy equation A.144. It is not obvious--to me, at least--that this process requires y_n to converge to u, a solution of Euler's equation. However, we will take refuge in a well-known device and state that such an investigation exceeds the scope of the present work.

Problem: Use Galerkin's method to verify the above example of the Rayleigh-Ritz method.

The Method of Steepest Descent.

This method is developed in some detail in the text, Chapter IX.

CONCLUSION

In the time (and space) which we can devote to the calculus of variations, we have been able to give only the briefest consideration to a few important problems. Many questions have been ignored or gracefully sidestepped. For example, most of our development has assumed existence of a solution, we have often required the existence of higher derivatives (a restrictive condition that may not be required by the physics of the problem), we have given no consideration to determining the type of stationary solution which satisfies the Eulerian equation, i.e., whether we have found a max, min, or inflection point, etc. All of these considerations are of great importance in particular problems.

As an example of the type of troubles which can be encountered, consider the problem of joining two points A and B by an extremal of

$$J = \int_{x_0}^{x_1} (y'^2)(y' + 1)^2 \, dx. \qquad A.146$$

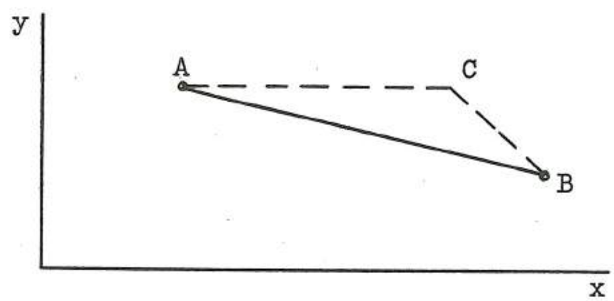
The Euler equation is

$$y'(y' + 1)^2 + y'^2(y' + 1) = \text{const} \qquad \text{A.147}$$

and the solution is

$$y = mx + b \qquad \text{A.148}$$

Suppose that A and B are located so that the slope of y is between -1 and 0. Since each term in the integrand is squared, certainly J has a positive value.



But now consider the path ACB where $y' = 0$ from A to C and $y' = -1$ from C to B. Then the integrand of J vanishes throughout the interval, $J = 0$, and the solution of the Eulerian equation did not furnish the minimum. The solution ACB was ruled out by the requirement for continuity imposed in obtaining Euler's equation.

This rather melancholy situation can be remedied, but the considerations are appreciably more sophisticated than we can handle here. In many cases such solutions can be ruled out on physical grounds--however, there are numerous physical problems where such discontinuous solutions are precisely the ones sought. (For example, the torque curves of an optimum bang-bang servo and the discontinuous thrust in some trajectory problems.) In such cases, it will be necessary to apply more advanced methods some of which are given, for example, in the book by Bliss.

APPENDIX B

ANALOG COMPUTER SOLUTION OF THE ORBITAL FLIGHT EQUATIONS*

by

L. E. Fogarty and R. M. Howe

A modified flight-path axis system is introduced to simplify computer mechanization of the orbital flight equations. Complete six degree-of-freedom equations are presented, including aerodynamic, powerplant, and control forces. Electronic differential analyzer mechanization of these equations in both real and fast time is described, including a novel technique for division which preserves favorable multiplier scaling. Specific machine results are presented which demonstrate accurate solution of close-satellite trajectories, including re-entry from satellite altitudes to sea level. With no change in circuit or scaling the same computer mechanization yields zero-drag orbits which close within several hundred feet of altitude.

---

* This work was supported by the United States Air Force, Wright Aeronautical Development Division, under USAF Contract No. AF 33(616)-5664

ANALOG COMPUTER SOLUTION OF THE ORBITAL FLIGHT EQUATIONS

by

L. E. Fogarty and R. M. Howe

1. Introduction

One of the most challenging of simulation problems is the computation of the trajectory and orientation of an orbital flight vehicle. The problem is particularly difficult if an analog computer is utilized for the solution of the orbital flight equations. This is because of the tremendous range of many of the variables, such as flight velocity, and because of the fact that in orbital flight the net radial acceleration is the small difference of two large forces, gravity and centrifugal. If the simulation must be conducted in real time, the extremely low rates of change of the variables present additional problems for an analog computer. In this paper an axis system is introduced which greatly reduces these difficulties, and examples of accurate analog solutions of orbital flight and re-entry to sea level for lifting vehicles are presented, both in speeded-up time and in real time. The axis system described here is also believed to offer advantages for digital mechanization of the problem.

2. Axis Systems

The equations which must be integrated to determine the motion have been well known since Euler's day. However, even though the form of the equations is rigidly prescribed, there is freedom to select axis systems and variables. The efficient use of a simulator computer requires that care be exercised in selecting computer variables, since the accuracy of the result is greatly affected by scaling.

The equations of motion describe the motion of the vehicle with respect to inertial space. This description of the motion is relatively uninteresting, since close satellite missions will be concerned with the surface of the earth. Thus we will have use for an earth reference frame as well as the inertial reference frame. Further, because the moments and products of inertia are most easily referred to vehicle body axes, and because aerodynamic and powerplant forces and moments are computed in a reference frame which moves with the vehicle, we will have use for a reference frame which is rigidly attached to the vehicle. Other sets of axes will be introduced for convenience.

The reference frames which will be used are illustrated in Figure 1 and include the following:

INERTIAL, I-FRAME

Motion of the vehicle with respect to "inertial" space is referred to the I-frame. Origin of the I-frame is the center of the earth, the reference plane is the equatorial plane, and the reference direction in the plane is the vernal equinox. Although the origin of this frame describes a near circular motion about the sun which produces an inertial force, it can be shown that this force is almost exactly cancelled by the sun's gravitational force for near satellites, and therefore it is valid to use the center of the earth as the origin for an inertial frame for this type of problem.

NAVIGATIONAL, N-FRAME

Motion of the vehicle with respect to the surface of the earth is referred to the N-frame, which is identical to the I-frame except that it rotates with respect to it with the earth's angular velocity r_n.

The reference direction of the N-frame is a line from the center of the earth through the intersection of the prime (Greenwich) meridian and the equator. In Figure 1 the unit vectors for the N-frame are $\vec{i}_n$, $\vec{j}_n$, $\vec{k}_n$. The vehicle at point P is located with respect to the N-frame by the longitude λ, latitude L, and radial distance r.

BODY AXES, B-FRAME

Conventional body axes x_b, y_b, and z_b are used with the origin at the center of gravity of the vehicle. In level flight the x_b axis points forward, the z_b axis downward, the y_b axis sideward to form a right-hand set. For a winged vehicle the x_b, z_b axes lie in the plane of symmetry of the vehicle.

EULER ANGLE, E-FRAME

Orientation of the vehicle body axes is specified by three conventional Euler angles ψ (heading), θ (pitch), and ϕ (bank); angles which give the orientation of the B-frame with respect to the E-frame. The E-frame origin is at the center of gravity of the vehicle, the reference plane is perpendicular to the radius vector to the center of the earth, and the reference direction in the reference plane is north. Thus for an ideal spherical earth the unit vectors $\vec{i}_e$, $\vec{j}_e$, and $\vec{k}_e$ for the E-frame point north, east, and downward, respectively.

MODIFIED FLIGHT PATH, H-FRAME

The vehicle translational equations of motion are referred to the H-frame. Origin of the H-frame is at the center of gravity of the vehicle, the

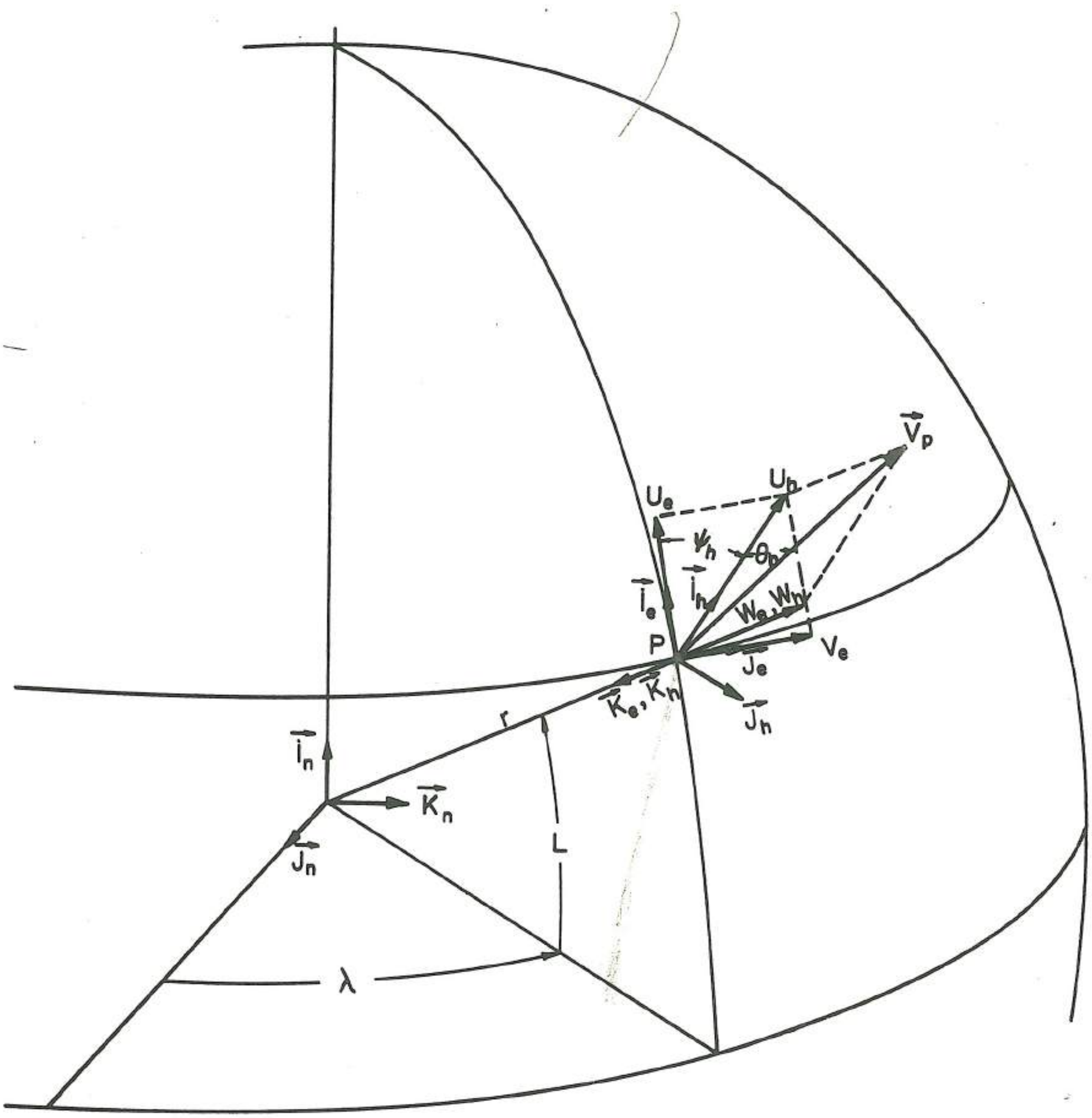


Figure B-1. Axis Systems Used for the Trajectory Equations.

reference plane is perpendicular to the radius vector to the center of the earth, and the reference direction is the projection of the vehicle inertial velocity vector $\vec{V}_p$ onto the reference plane. Thus for an ideal spherical earth the unit vectors $\vec{i}_h$, $\vec{j}_h$, and $\vec{k}_h$ for the H-frame point, respectively, in the direction of horizontal component of velocity, horizontally at right angles to this velocity, and downward. The H-frame is oriented with respect to the E-frame by the heading angle ψ_h, as shown in Figure 1. The vehicle velocity $\vec{V}_p$ is oriented with respect to the H-frame by the attitude angle θ_h.

3. Notation

Wherever possible, standard symbols have been used. Thus U, V, and W denote the x, y, and z components of a translational velocity vector; P, Q, and R are the x, y, and z components of an angular velocity vector. The vehicle velocity $\vec{V}_p$ (i.e., the translational velocity of the B-frame) with respect to the inertial, I-frame can be resolved into components along any of the axis systems, the component subscript denoting the particular system. Thus U_e, V_e, and W_e are the components of $\vec{V}_p$ along the three E-frame axes, respectively, whereas U_h, V_h, and W_h are the components of $\vec{V}_p$ along the three H-frame axes, respectively.

Motion of the vehicle (i.e., the B-frame) with respect to reference frames other than the inertial, I-frame and motion of reference frames other than the B-frame generally will be identified with a double subscript or a different symbol.

4. Translational Equations of Motion

Having defined the various axis systems which we will need to consider, let us turn to the derivation of the translational equations of motion using modified flight-path axes (the H-frame). First we note that Euler's equations of motion representing equilibrium of forces along three orthogonal moving axes x, y, and z are:

$$-m(\dot{U} - VR + WQ) + X = 0 \quad \text{B.1}$$

$$-m(\dot{V} - WP + UR) + Y = 0 \quad \text{B.2}$$

$$-m(\dot{W} - UQ + VP) + Z = 0 \quad \text{B.3}$$

Here m is the vehicle mass, U, V, W and P, Q, R are the translational and rotational inertial velocity components along the x, y, and z axes, respectively, and X, Y, and Z are external forces along these axes, respectively.

We will use the H-frame as the reference axes for equations (B.1-B.3). Since the origin of the H-frame coincides with the origin of the B-frame (the vehicle center of gravity), the translational velocity components of the H-frame are U_h, V_h, and W_h, the same as the vehicle velocity components along x_h, y_h, and z_h. Let us denote the H-frame components of the angular velocity of the H-frame with respect to the I-frame by p_h, q_h, and r_h. We note that

$$V_h = 0 \text{ (by definition of the H-frame)} \tag{B.4}$$

$$p_h = \frac{V_h}{r} = 0 \tag{B.5}$$

$$q_h = -\frac{U_h}{r} \tag{B.6}$$

Thus Eqs. B.1-B.3, referred to the H-frame, become

$$-m(\dot{U}_h - \frac{U_h W_h}{r}) + X_h = 0 \tag{B.7}$$

$$-m r_h U_h + Y_h = 0 \tag{B.8}$$

$$-m(\dot{W}_h + \frac{U_h^2}{r}) + Z_h + \frac{m g_o r_o^2}{r^2} = 0 \tag{B.9}$$

where Z_h is the z_h component of the external force vector not including the central force field term $m g_o r_o^2/r^2$, which has been included separately. Here g_o is the central force-field gravity acceleration at a constant reference radial distance r_o. Note that X_h, Y_h, and Z_h include the effects of oblateness on the gravity force.

Next we observe that $W_h = -\dot{r}$. Thus $\frac{1}{r}\frac{d}{dt}(rU_h) = \dot{U}_h + \frac{\dot{r}}{r} U_h = \dot{U}_h - \frac{W_h U_h}{r}$ and equation B.9 can be rewritten as

$$\frac{d}{dt}(rU_h) = \frac{rX_h}{m} \tag{B.10}$$

Integrating, we obtain

$$U_h = \frac{1}{r}\int_o^t \frac{rX_h}{m}\, d\tau + \frac{(rU_h)_o}{r} \tag{B.11}$$

This is essentially the angular momentum integral (actually mrU_h = angular momentum). In the absence of any external horizontal forces it allows direct calculation of horizontal velocity U_h as the radial distance r changes. For finite X_h the angular momentum is changed in accordance with the time integral of the moment rX_h.

Let us now calculate $\dot{\psi}_h$, the rate of change of H-frame heading angle. This will be the difference between H-frame angular velocity along $\vec{k}_h$ (downward) and E-frame angular velocity along $\vec{k}_h$(and $\vec{k}_e$). Thus

$$\dot{\psi}_h = r_h - \omega_{ez} \qquad \text{B.12}$$

where ω_{ez} is the z_e component of E-frame angular velocity $\vec{\omega}_e$. To determine ω_{ez} refer to Figure 1 and note that

$$\vec{\omega}_e = \frac{V_e}{r \cos L} \vec{i}_n - \frac{U_e}{r} \vec{j}_e \qquad \text{B.13}$$

and

$$\vec{i}_n = \cos L \, \vec{i}_e - \sin L \, \vec{k}_e \qquad \text{B.14}$$

Thus

$$\vec{\omega}_e = \frac{V_e}{r} \vec{i}_e - \frac{U_e}{r} \vec{j}_e - \frac{V_e}{r} \tan L \, \vec{k}_e \qquad \text{B.15}$$

and $\omega_{ez} = -\frac{V_e}{r} \tan L$. From this and Eq. B.8 we have from Eq. B.12

$$\dot{\psi}_h = \frac{Y_h}{mU_h} + \frac{V_e}{r} \tan L \qquad \text{B.16}$$

Eq. B.9 can be written as

$$\dot{W}_h = \frac{g_o r_o^2}{r^2} - \frac{U_h^2}{r} + \frac{Z_h}{m} \qquad \text{B.17}$$

Eqs. B.11, B.16, and B.17 are the three translational equations using the H-frame.

5. Latitude, Longitude, and Altitude

From Figure 1 it is evident that

$$\dot{L} = \frac{U_e}{r} = \frac{U_h \cos \psi_h}{r} \qquad \text{B.18}$$

and

$$\dot{\lambda} = \frac{V_e}{r \cos L} - r_n = \frac{U_h \sin \psi_h}{r \cos L} - r_n \qquad \text{B.19}$$

where r_n is the angular rate of the earth (N-frame) with respect to the I-frame, and is directed along the earth's spin axis, $\vec{k}_n$.

If we compute the radial distance from the center of the earth by the formula

$$r = r_o + \delta r \qquad \text{B.20}$$

where r_o is a convenient fixed radius, e.g., the mean radial distance of the nominal orbit, then

$$\dot{\delta r} = -W_h$$

and

$$h = h_o + \delta r + f_h(L) \qquad \text{B.21}$$

Here h is the altitude above sea level, h_o is the altitude at the reference radius r_o, and $f_h(L)$ is a function of latitude which introduces the geometrical effect of latitude on oblateness.

6. Angle of Attack, Angle of Sideslip, Dynamic Pressure, Mach Number

The aerodynamic forces and moments are functions of angle of attack α, angle of sideslip β, dynamic pressure q, and Mach number M.

Defining V_a, the total aircraft velocity with respect to the air mass, with body axis components U_{ab}, V_{ab}, W_{ab}, we have:

$$\alpha = \tan^{-1} \frac{W_{ab}}{U_{ab}} \qquad \text{B.22}$$

$$\beta = \tan^{-1} \frac{V_{ab} \cos \alpha}{U_{ab}} \qquad \text{B.23}$$

$$q = \frac{1}{2} \rho V_a^2 \qquad \text{B.24}$$

$$M = V_a/a \qquad \text{B.25}$$

$$V_a = \frac{U_{ab}}{\cos \alpha \cos \beta} \qquad \text{B.26}$$

where ρ is the atmospheric density and a is the speed of sound; both are functions of altitude.

Equations B.22, B.23, and B.26 are unusable when cos α or cos β equal zero, which might happen in orbital flight. There are other formulas for α, β and V_a which can be used; the best formula to use depends on the vehicle being simulated.

Determination of U_{ab}, V_{ab}, W_{ab} requires knowledge of the vehicle orientation with respect to the E-frame, treated in the next section. Because of the lengthy trigonometric formulas, they will be expressed in terms of conventional direction cosines relating the B-frame to the E-frame. Expressions for the direction cosines are given in Figure 2.

Let $\ell_{1,2,3}$, $m_{1,2,3}$, $n_{1,2,3}$ be the direction cosines relating the B-frame to the E-frame. Let w_x and w_y be the north and east components of wind. The velocity of a windless air mass with respect to the I-frame is $-r_n r \cos L$ directed along the y_e axis.

Then the B-frame components of the vehicle velocity with respect to the moving air mass are:

$$U_{ab} = \ell_1(U_e + w_e) + \ell_2(V_e + w_y - r_n r \cos L) + \ell_3 W_e \qquad \text{B.27}$$

$$V_{ab} = m_1(U_e + w_x) + m_2(V_e + w_y - r_n r \cos L) + m_3 W_e \qquad \text{B.28}$$

$$W_{ab} = n_1(U_e + w_x) + n_2(V_e + w_y - r_n r \cos L) + n_3 W_e \qquad \text{B.29}$$

7. Rotational Equations, Euler Angles, Direction Cosines

The equations derived in the previous sections, along with the rotational, Euler angle, direction cosine, and force-resolution equations are summarized

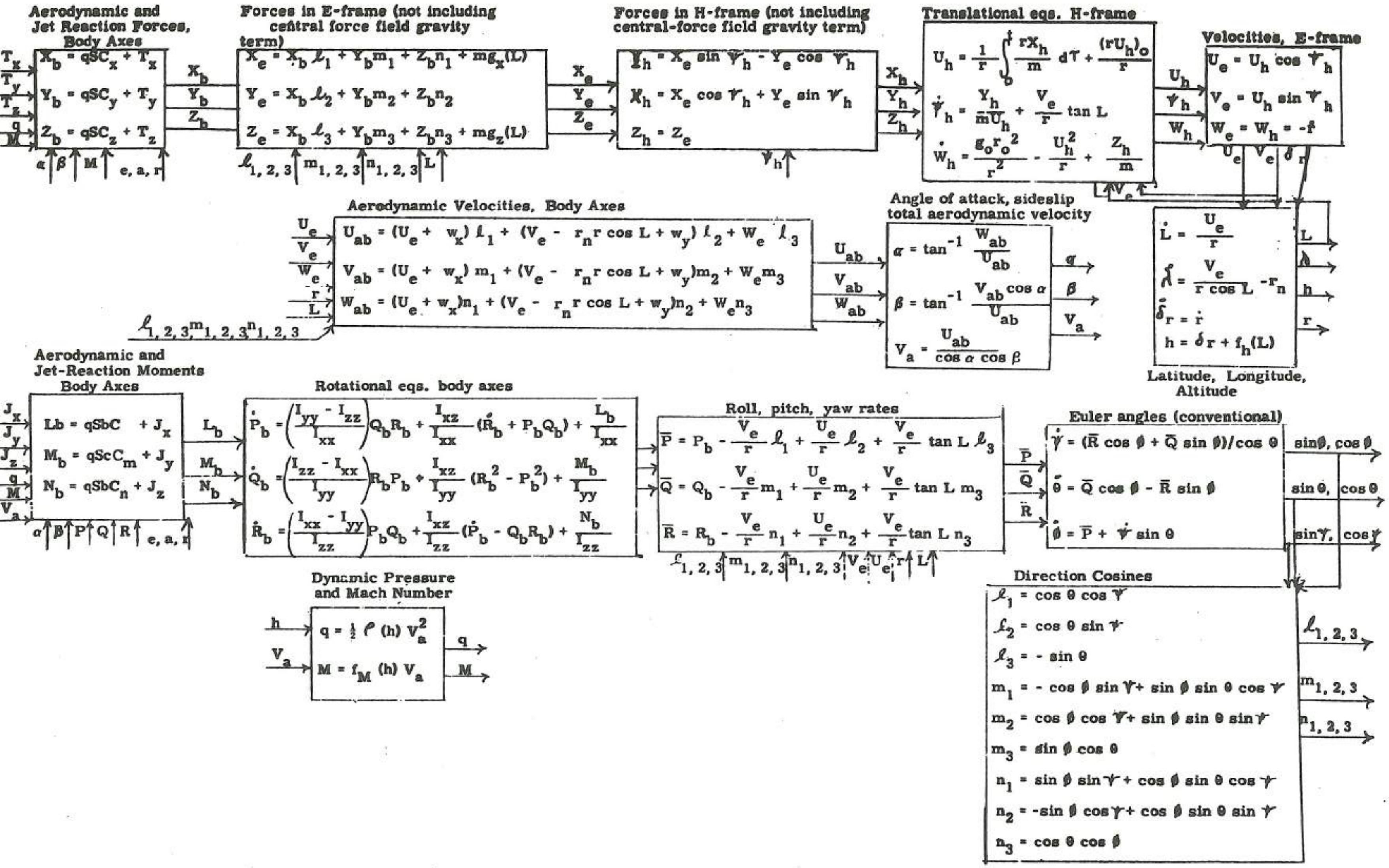


Figure 2. Block diagram of the Equations of motion using the H-frame for translational equations and the B-frame (Body Axes) for rotational equations.

in block-diagram form in Figure 2. The rotational equations result from equilibrium of moments about the x_b, y_b, z_b body axes and involve the components P_b, Q_b, and R_b of B-frame angular velocity along these axes, respectively. The equations are well known and are presented in the box labeled "Rotational eqs., body axes" in the lower center of the figure. In many cases most of the nonlinear terms in these equations are negligible.

To obtain the Euler angles ψ, Θ, and ϕ which orient the B-frame with respect to the E-frame it is necessary to calculate the B-frame angular velocity with respect to the E-frame. This is equal to the B-frame angular velocity with respect to the I-frame (P_b, Q_b, and R_b along x_b, y_b, z_b) minus the E-frame angular velocity with respect to the I-frame. The resulting roll, pitch, and yaw rates $\overline{P}$, $\overline{Q}$, and $\overline{R}$ are given by the formulas in Figure 2 shown in the box labeled "Roll, pitch, yaw rates" at the lower center part of the figure. To the right of this is a box which summarizes the familiar formulas for the Euler angles as computed from $\overline{P}$, $\overline{Q}$, and $\overline{R}$. Below this is a box containing the formulas for the 9 direction cosines.

It should be observed that there are other methods for representing the vehicle orientation, and there are alternative ways of computing either the Euler angles or direction cosines. The actual schemes used will depend on the required outputs from the simulation and the type of computer used to solve the equations.

8. Resolution of Translational Forces

Aerodynamic, control and power-plant forces and moments will usually be specified in terms of body-axis or stability-axis coefficients. Assuming that body axis coefficients are used, the moments may be used directly in the rotational equations of Figure B-2. Since the translational equations are written in the H-frame, the forces must be resolved to the H-frame. Because gravity forces due to oblateness are given in the E-frame, it appears that the force resolution is most easily accomplished by resolving the aerodynamic and power-plant forces into the E-frame where they may be combined with the gravity perturbation forces and the combined forces then resolved to the H-frame.

Let the aerodynamic and power-plant forces be X_b, Y_b, Z_b along the x_b, y_b, z_b axes and X_e, Y_e, Z_e along the x_e, y_e, z_e axes. Let $g_x(L)$ and $g_z(L)$ be the gravity perturbation forces along the x_e and z_e axes.

Then

$$X_e = X_b \ell_1 + Y_b m_1 + Z_b n_1 + g_x(L) \qquad \text{B.30}$$

$$Y_e = Y_b \ell_2 + Y_b m_2 + Z_b n_2 \qquad \text{B.31}$$

$$Z_e = X_b \ell_3 + Y_b m_3 + Z_b n_3 + g_z(L) \qquad B.32$$

and

$$Y_h = X_e \sin \psi_h - Y_e \cos \psi_h \qquad B.33$$

$$X_h = X_e \cos \psi_h + Y_e \sin \psi_h \qquad B.34$$

$$Z_h = Z_e \qquad B.35$$

These equations are also summarized in Figure B-2.

9. Mechanization of the Translational Equations for Analog Solution

Equation B.17 for $\dot{W}_h$, the vertical acceleration of the vehicle, is difficult to solve accurately on an analog computer because in orbital flight the U_h^2/r centrifugal term almost exactly cancels the gravity term, $g_o r_o^2/r^2$. Any small errors in computing either term can lead to errors in $\dot{W}_h$ and hence in the trajectory over an extended period of time. The required divisions by r can be accomplished with high accuracy, as we shall see later, by letting $r = r_o + \delta r$ and by using a multiplier driven by δr in the feedback loop of an amplifier. But the U_h^2 term must be calculated in some indirect fashion to avoid the inherent errors in any analog squaring device (for example, errors of 0.05% or less might be prohibitively large in computing U_h^2 in orbit).

To avoid this difficulty we let the horizontal velocity component U_h be given by

$$U_h = U_{ho} + \delta U_h \qquad B.36$$

where U_{ho} is a constant velocity defined by

$$U_{ho} = \sqrt{g_o r_o} \qquad B.37$$

Thus U_{ho} is the horizontal velocity for a circular orbit at radial distance r_o from the center of a spherical earth and δU_h is the deviation of the actual tangential velocity U_h from U_{ho}. The reference radius r_o is most conveniently set equal to the mean radius of the highest-apogee trajectory which must be computed, so that δr will range between equal positive and negative limits ($\delta r = r - r_o$).

Rewriting equation B.17 in terms of δU_h and δr, we obtain

$$-\dot{W}_h = \delta\ddot{r} = \frac{U_{ho}^2\ \delta r}{(r_o + \delta r)^2} + \frac{2U_{ho}\delta U_h + (\delta U_h)^2}{r_o + \delta r} - \frac{Z_h}{m} \qquad \text{B.38}$$

Here the centrifugal acceleration term U_{ho}^2/r_o has canceled the gravity term g_o and the only terms left in Eq. B.38 are all small for a near circular orbit of radius close to r_o. This is because $\delta r \ll r_o$ and $\delta U_h \ll U_{ho}$ in this case. The term $(\delta U_h)^2$ becomes small and the terms on the right side are approximately linear in δr and δU_h. Thus for small external forces Z_h (e.g., small aerodynamic forces), Eq. B.38 becomes an approximately linear equation, but it should be emphasized that no approximations have been made in Eq. B.38.

During the ascent or re-entry trajectory δU_h will become negative and may grow as large as U_{ho}, the circular reference-orbit velocity. In this case the $(2U_{ho}\delta U_h + \delta U_h^2)/(r_o + \delta r)$ term in Eq. B.38 essentially becomes the gravity acceleration. Any inaccuracies in computing $(\delta U_h)^2$ (e.g., errors of 0.1%) are not of great importance under these conditions, because the pilot will be in the computer loop and also because the length of time during which errors can build up will be short.

To simplify computer scaling it is desirable to introduce dimensionless distance, velocity, and time variables. Thus let

$$\delta\rho = \frac{\delta r}{r_o}, \quad \delta u_h = \frac{\delta U_h}{U_{ho}}, \quad w_h = \frac{W_h}{U_{ho}}, \quad \tau = \sqrt{g_o/r_o}\ t \qquad \text{B.39}$$

In terms of these variables the translational Equations B.11 and B.38 become

$$\delta u_h = \frac{1}{1 + \delta\rho}\left[\int (1 + \delta\rho)\frac{X_h}{mg_o}\,d\tau - \delta\rho\right] \qquad \text{B.40}$$

$$\frac{dw_h}{d\tau} = -\frac{d^2\rho}{d\tau^2} = -\frac{\delta\rho}{(1 + \delta\rho)^2} - \frac{2\delta u_h + (\delta u_h)^2}{1 + \delta\rho} + \frac{Z_h}{mg_o} \qquad \text{B.41}$$

Note that the external forces X_h and Z_h actually appear as equivalent accelerations in units of g_o, the gravity acceleration at the reference radius r_o. The initial condition on the integral in Eq. B-40 determines the initial horizontal velocity variable δu_h.

Equations B.40 and B.41 constitute the two-dimensional orbital flight equations. The output variables are dimensionless radial perturbation $\delta\rho$, dimensionless vertical velocity - w_h, and dimensionless horizontal velocity perturbation δu_h. Although $\delta\rho$ and δu_h are perturbations from corresponding values for a circular orbit at reference altitude, note that Eqs. B.40 and B.41 are exact no matter how large these variables become. An electronic differential analyzer circuit for solving these equations is shown in Figure 3. To solve the equations in real time the time constant of each integrator is nominally s seconds, where $s = \sqrt{r_o/g_o} = 829$ seconds for r_o corresponding to a mean altitude 80 statute miles above the surface of the spherical earth.

Several features of the circuit in Figure 3 should be mentioned. The circuit is scaled with +100 volts equal to unity in the problem variables. By computing $50\ \delta\rho$, we allow $\delta\rho$ to range over $\pm 1/50$, corresponding to a range in δr of ± 80 statute miles. The division of voltages by $1 + \delta\rho$, as required in Eqs. B.40 and B.41, is accomplished by using unity feedback directly across the amplifier and, in addition, by multiplying the amplifier output by $50\delta\rho$ and feeding it back to the amplifier input attenuated by a factor of 50. This circuit exactly produces the required division by $1+ \delta\rho$ and yet uses the multiplier over the full range of $\pm 50\ \delta\rho$. As a result any multiplier errors are reduced by a factor of 50 in their effect on the division. This is a very important feature of the mechanization and essentially eliminates multiplier errors from the division circuitry.

Note also in Figure 3 that δu_h is scaled to ± 1. This allows actual horizontal velocities ranging from zero to twice orbital velocity, so that as it stands the circuit can be used for simulation of ascent trajectories and re-entry to landing. The dimensionless velocity w_h is scaled to $\pm 1/5$, allowing vertical velocities of ± 5000 feet per second. For near orbital flight the scaling on δu_h and w_h is very unfavorable, but we wanted to demonstrate the accurate performance of the circuit in Figure 3 for possible simulation of the entire flight of an orbital vehicle, from takeoff through one or more orbits to landing.

To compute the angular distance θ_s which the vehicle travels in the two-dimensional orbit the following equation is used:

$$\theta_s = \int \frac{(1 + \delta u_h)}{1 + \delta\rho}\, d\tau \qquad \text{B.42}$$

10. Analog Solution for the Zero-Drag, Zero-Lift Case

For the case where there are no external forces other than the center-force field gravity terms, X_h and Z_h in Eqs. B.40 and B.41 are both zero. For small $\delta\rho$ and δu_h the resulting equations represent an undamped second order system with a natural frequency of one radian per dimensionless time unit τ. For any non-equilibrium set of initial conditions the resulting motion is approximately an undamped sinusoid which represents, in the case of $\delta\rho$, just the periodic deviation of the elliptical orbit from the reference circular

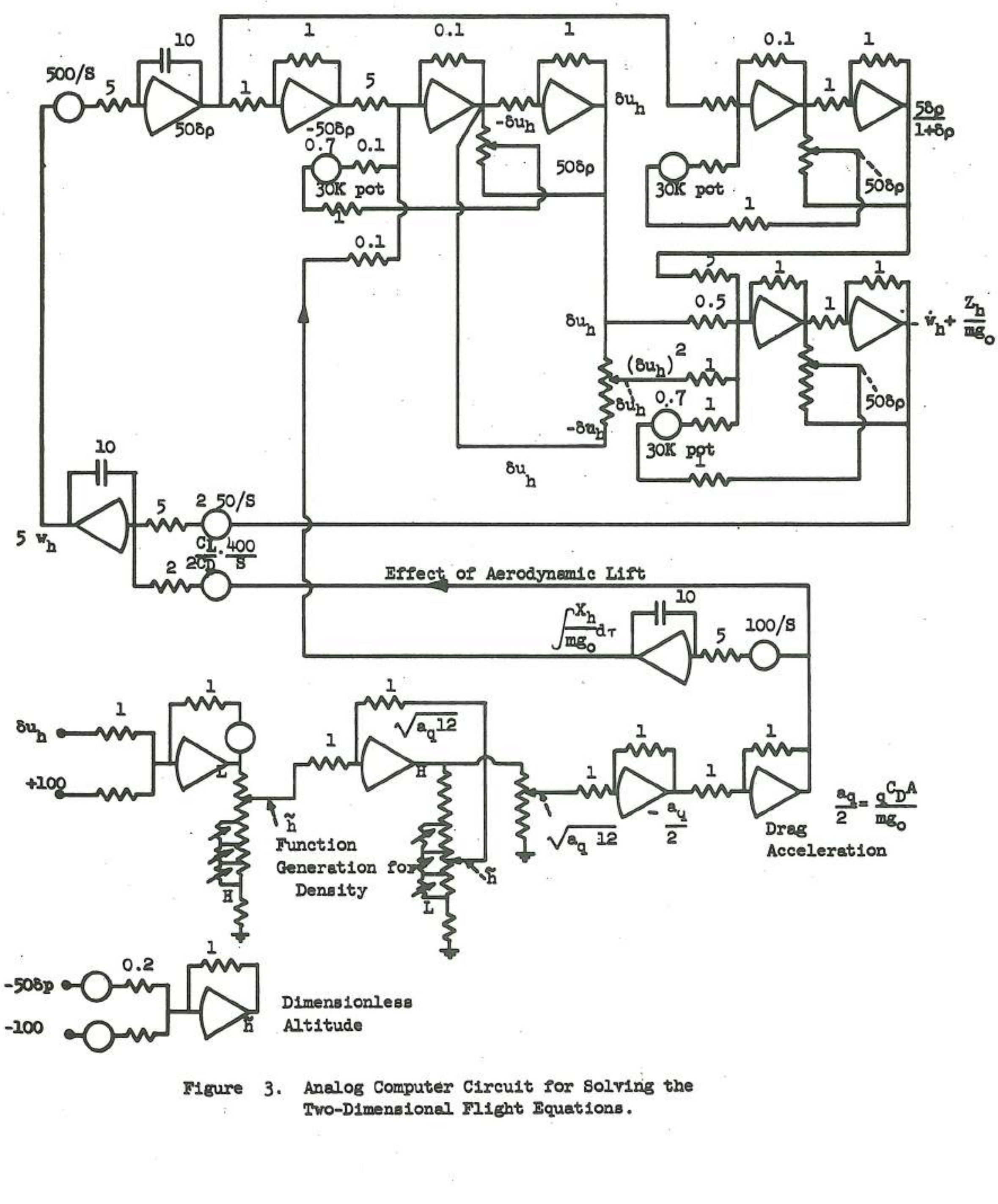


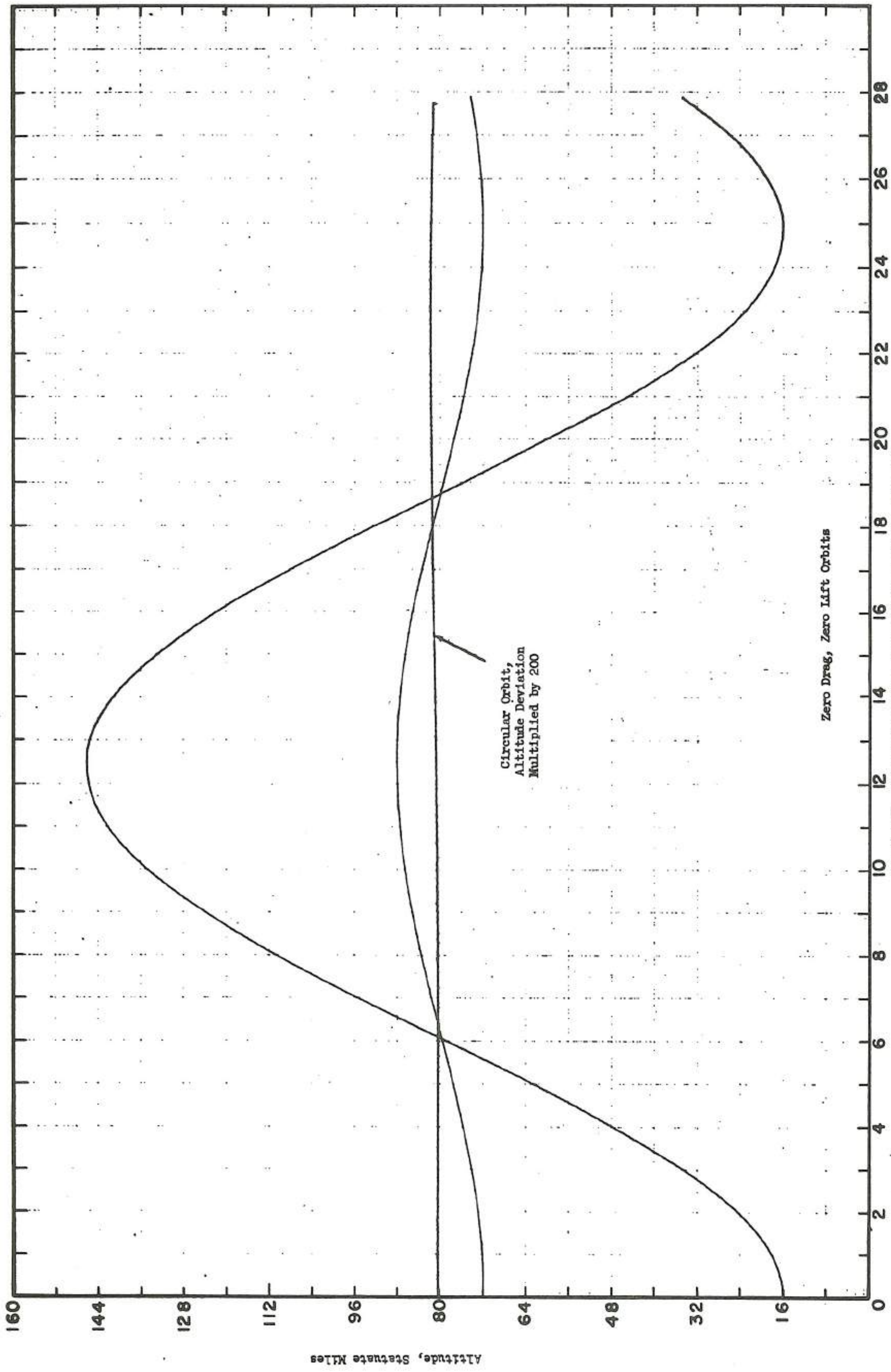
Figure 3. Analog Computer Circuit for Solving the Two-Dimensional Flight Equations.

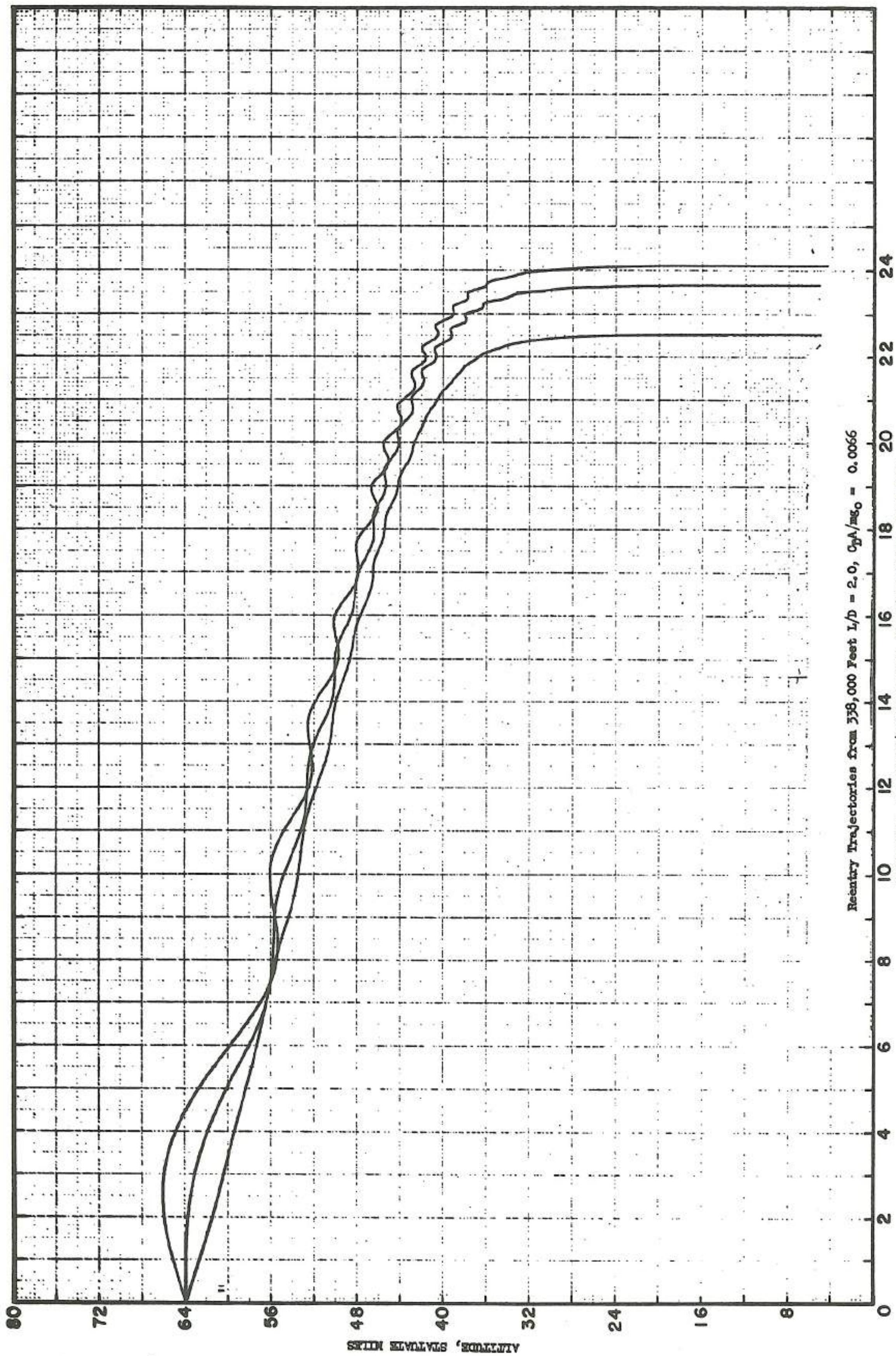
orbit. Under these conditions the principle sources of error in the analog solution are amplifier offsets which cause a slight bias in the average center value of the sinusoid in $\delta\rho$, and capacitor dielectric absorption and leakage resistance, which cause a fixed fractional loss of amplitude in peak $\delta\rho$ from one orbit to the next. Several typical analog solutions starting with the vehicle at perigee are shown in Figure 4. For the case where the initial perigee altitude is at 16 miles ($\delta\rho_o = -0.016$) the first apogee is within 200 feet of the correct value of 146.5 miles and the second perigee is within 200 feet of the initial value of 16 miles indicating orbit closure to that accuracy.

Also shown is a case where $\delta\rho_o = 0$, which represents injection into a circular orbit. The computer solution, blown up by a factor of 200, is shown in the figure and indicates that the altitude holds within 40 feet of the initial value over one orbital distance. The results were obtained at 100 times real time by reducing integrator resistors and capacitors by a factor of ten over those shown in Figure B-3. Comparable results were obtained in real time and at 10 times real time. It is well to remember that these results are possible with an analog computer because of the prior integration of the equation for horizontal velocity (except for the external torque term) before putting the problem on the computer. This eliminates an open-ended integration of $du_h/d\tau$ with the resulting long-term drift errors. Also, the deviation of horizontal velocity from circular orbital velocity is computed instead of horizontal velocity itself. This eliminates multiplier errors in the centrifugal term, providing, of course, that a servo multiplier is used to compute $(\delta u_h)^2$. When a servo is used, the error in computation of $(\delta u_h)^2$ diminishes as δu_h instead of remaining more or less fixed, as in the case with electronic multipliers.

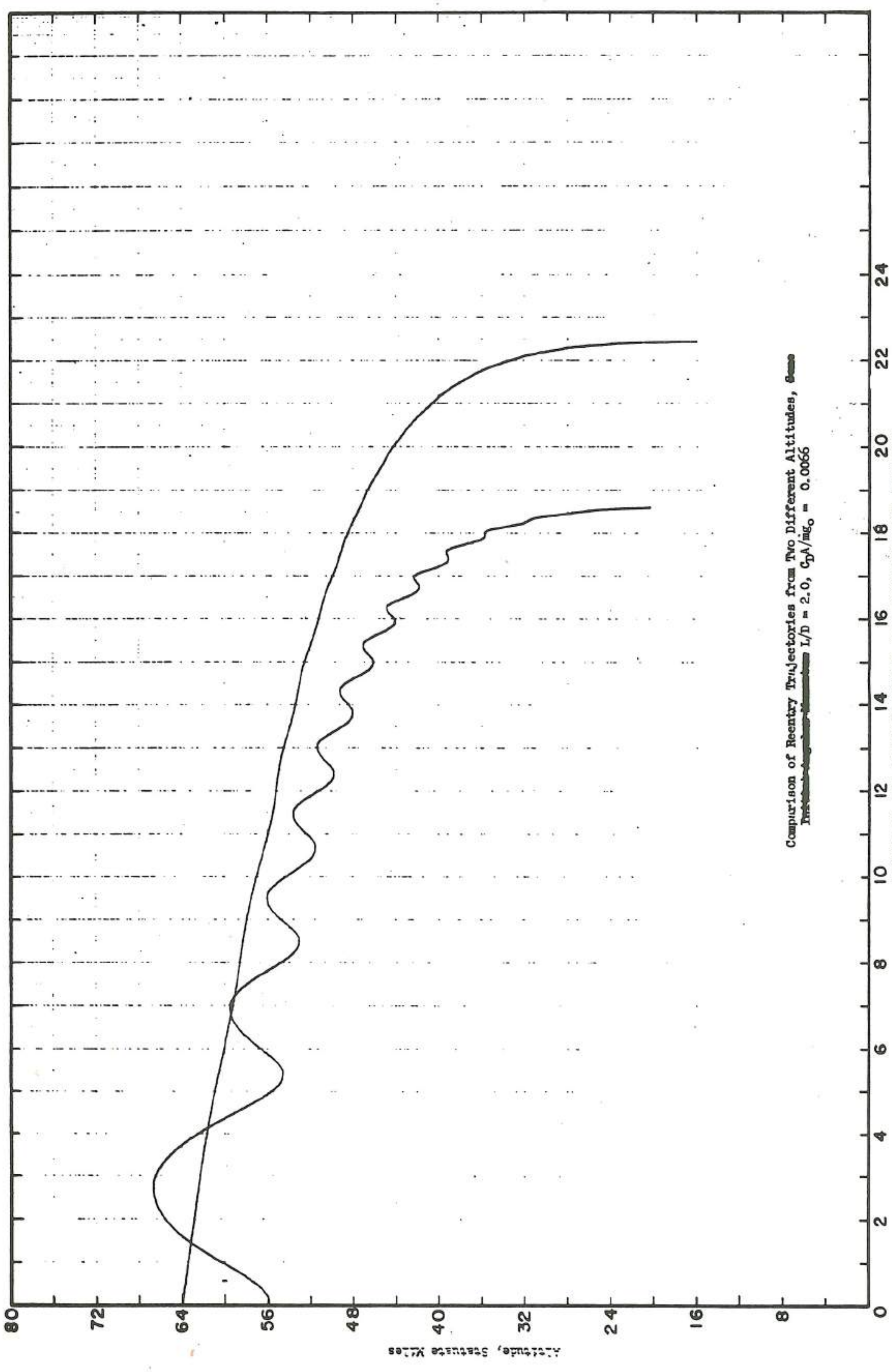
11. Analog Solution for Re-entry from Orbital Flight

The effectiveness of the circuit in Figure B-3 for the entire range of the variables is best demonstrated by considering a re-entry from orbital altitude and velocity to sea level and zero velocity. Dynamic pressure q was calculated by computing $\sqrt{q} = \sqrt{\rho/2}\ V_a$, as shown in Figure B-3, and then squaring the result. The variation of density ρ with altitude was simulated by approximating $\sqrt{\rho}$ in 16 segments with two servo-driven tapped pots. For a density variation over a range of 6 orders of magnitude $\sqrt{\rho}$ needs to be varied over 3 orders of magnitude. For simplicity it was assumed that the aerodynamic velocity $V_a \cong U_h$, and that drag acts horizontally, lift vertically. Since the flight path angle never exceeded a degree or two until the very end of re-entry, this assumption was reasonable. For a fixed lift coefficient $C_L = 0.4$, drag coefficient $C_D = 0.2$, and $C_D A/mg_o = 0.0066$ the computer solutions shown in Figure B-5 were obtained. The atmospheric density was scaled from sea level to 320,000 feet for these runs, which were made at 100 times real time. Repeatability of runs was within the width of the recorder-pen

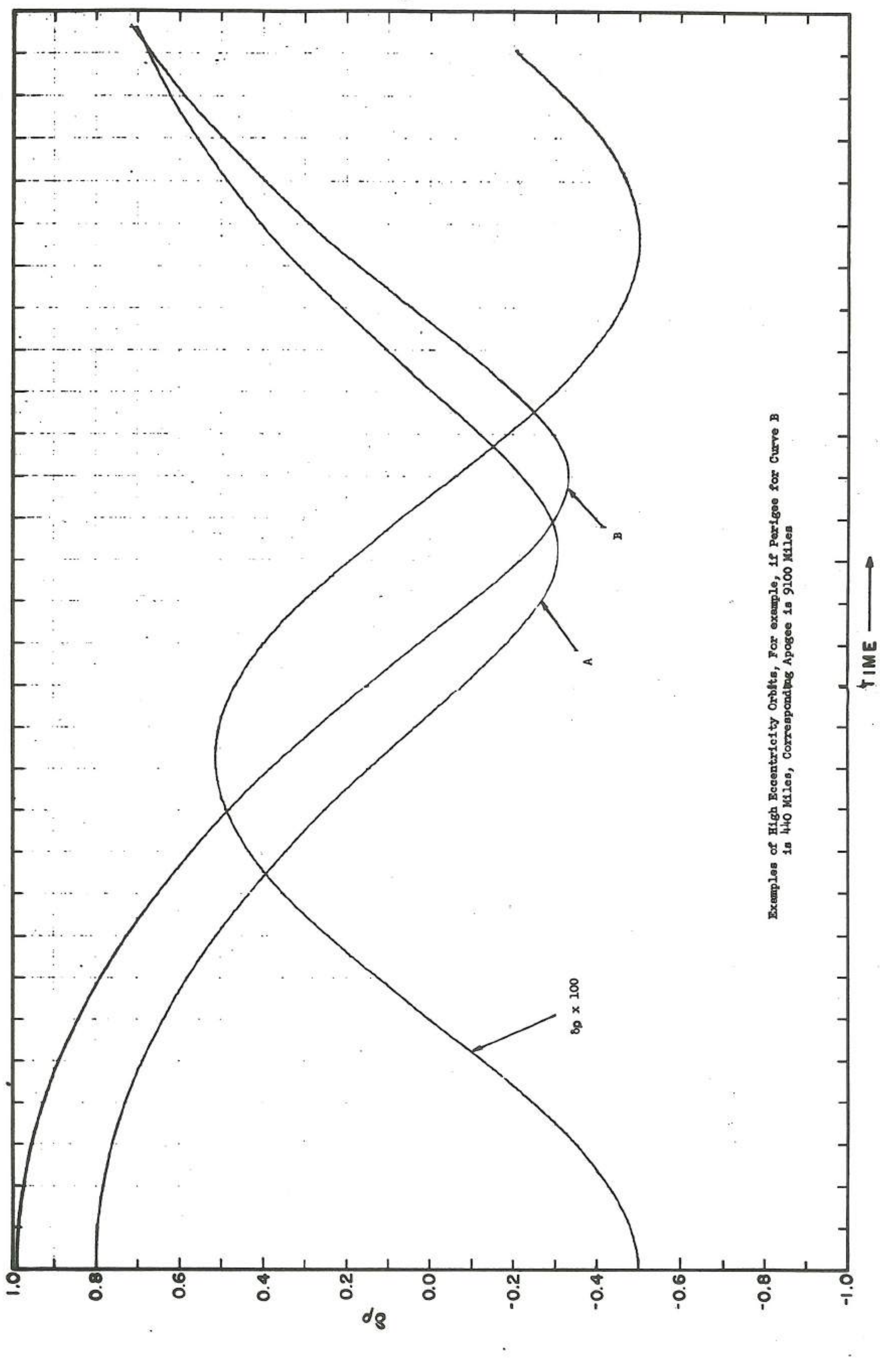

Circular Orbit,
Altitude Deviation
Multiplied by 200
Zero Drag, Zero Lift Orbits
RANGE, THOUSANDS OF STATUATE MILES
Altitude, Statuate Miles
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28
16 32 48 64 80 96 112 128 144 160



Reentry Trajectories from 338,000 Feet L/D = 2.0, C_DA/mg_o = 0.0066
RANGE, THOUSANDS OF STATUTE MILES
0
2
4
6
8
10
12
14
16
18
20
22
24
80
72
64
56
48
40
32
24
16
8
0



Comparison of Reentry Trajectories from Two Different Altitudes, Same [redacted] L/D = 2.0, C_DA/mg_o = 0.0066



Examples of High Eccentricity Orbits, For example, if Perigee for Curve B is 440 Miles, Corresponding Apogee is 9100 Miles line. In Figure B-6 are shown some runs with the atmospheric density scaled from 100,000 feet to 420,000 feet. Again the computer speed was 100 times real time. A single run at real time showed practical identical agreement with the equivalent fast run.

12. Conclusions

It is felt that the results in this paper demonstrate an efficient axis system for computer solution of the orbital flight equations and show the feasibility of employing an analog computer for this purpose, even in real time. The mechanization allows continuous simulation from orbital velocities and altitudes through re-entry to sea level without rescaling. In principle the simulation can be used for the ascent trajectory as well. Although only the two-dimensional translational equations were solved, this is the difficult part of the problem and the addition of the third dimension and the rotational equations of Figure B-2 should not cause difficulty. By use of the angular momentum integral the problem of open-ended integration to obtain horizontal velocity is eliminated. This is essentially equivalent to imposing an angular momentum constraint. An energy constraint could also have been imposed, but it does not seem worthwhile when appreciable aerodynamic forces are present. By computing velocity difference from circular reference orbit velocity and radial variation from circular reference orbit radius, one can eliminate the importance of multiplier errors for the near orbital simulation and yet be solving exact equations even at takeoff and landing.