NASA CONTRACTOR REPORT

NASA CR-1005

NASA CR-1005

GPO PRICE $ _____

CFSTI PRICE(S) $ 3.00

Hard copy (HC) _____

Microfiche (MF) .65

ff 653 July 65

FACILITY FORM 602

N68-15783 (ACCESSION NUMBER)

122 (PAGES)

(NASA CR OR TMX OR AD NUMBER)

1 (CODE)

30 (CATEGORY)

[redacted]

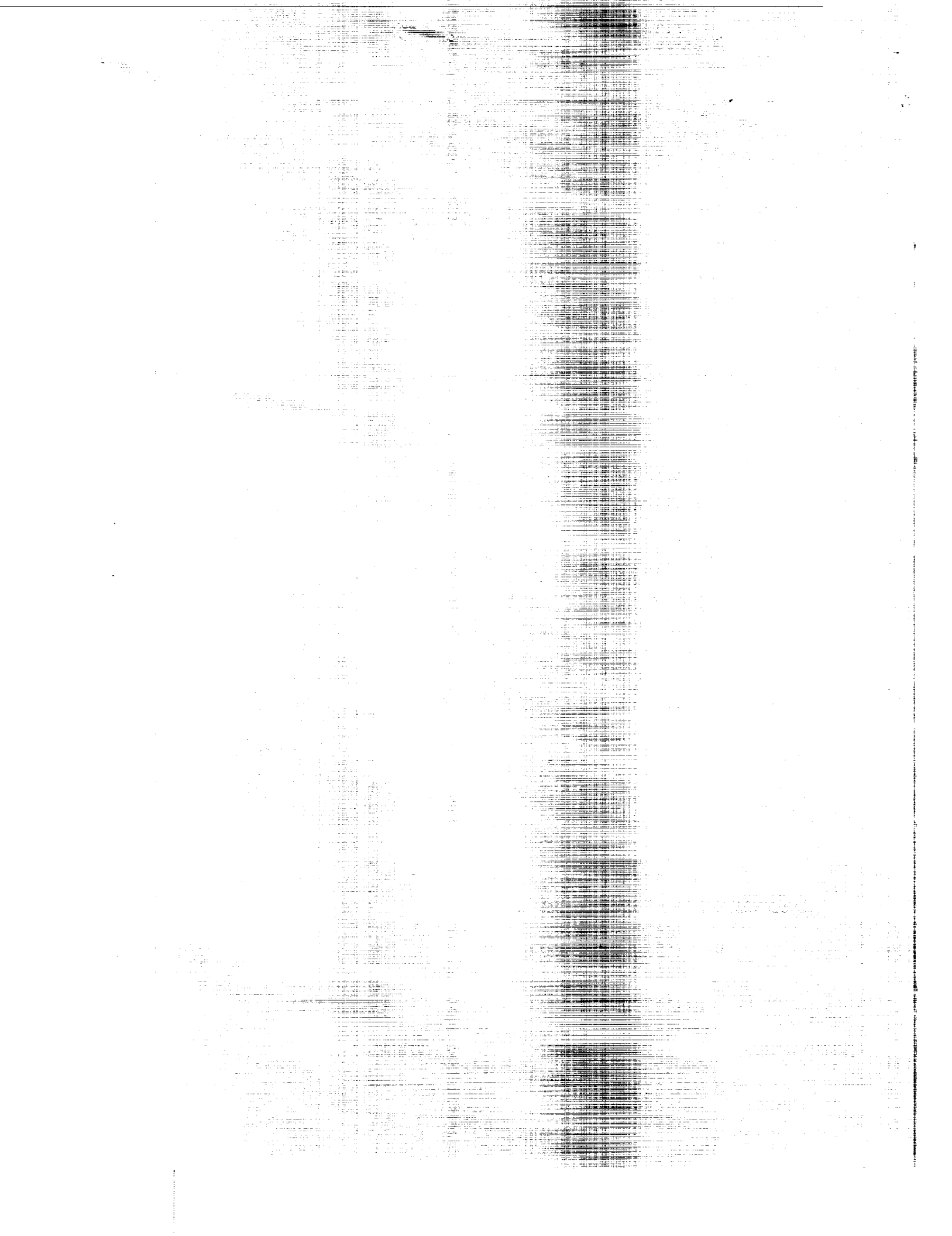
GU[redacted]N
AN[redacted]A

Volu[redacted]
Spec[redacted]

by N[redacted]de
and

Prepa[redacted]
NOR[redacted]
Down[redacted]
for G[redacted]

NATIO[redacted]H[redacted]



NASA CR-1005

GUIDANCE, FLIGHT MECHANICS AND TRAJECTORY OPTIMIZATION

Volume VI - The N-Body Problem and Special Perturbation Techniques

By M. S. Allione, A. L. Blackford, J. C. Mendez,

and M. M. Whittouck

Issued by Originator as Report No. SID 65-1200-6

Prepared under Contract No. NAS 8-11495 by
NORTH AMERICAN AVIATION, INC.
Downey, Calif.

for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

FOREWORD

This report was prepared under contract NAS 8-11495 and is one of a series intended to illustrate analytical methods used in the fields of Guidance, Flight Mechanics, and Trajectory Optimization. Derivations, mechanizations and recommended procedures are given. Below is a complete list of the reports in the series.

Volume I	Coordinate Systems and Time Measure
Volume II	Observation Theory and Sensors
Volume III	The Two Body Problem
Volume IV	The Calculus of Variations and Modern Applications
Volume V	State Determination and/or Estimation
Volume VI	The N-Body Problem and Special Perturbation Techniques
Volume VII	The Pontryagin Maximum Principle
Volume VIII	Boost Guidance Equations
Volume IX	General Perturbations Theory
Volume X	Dynamic Programming
Volume XI	Guidance Equations for Orbital Operations
Volume XII	Relative Motion, Guidance Equations for Terminal Rendezvous
Volume XIII	Numerical Optimization Methods
Volume XIV	Entry Guidance Equations
Volume XV	Application of Optimization Techniques
Volume XVI	Mission Constraints and Trajectory Interfaces
Volume XVII	Guidance System Performance Analysis

The work was conducted under the direction of C. D. Baker, J. W. Winch, and D. P. Chandler, Aero-Astro Dynamics Laboratory, George C. Marshall Space Flight Center. The North American program was conducted under the direction of H. A. McCarty and G. E. Townsend.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

a	semi-major axis of a conic
e	eccentricity of a conic
$\bar{h}$	angular momentum
i	orbital inclination
m	mass
M	mean anomaly
n	mean motion
P	semilatus rectum for a conic
$\bar{r}$	position vector
T_o	time of periapse passage
$\bar{V}$	velocity vector
δ	central difference operator
Δ	forward difference operator
∇	backward difference operator
ϵ	energy per unit mass
μ	gravitional constant for central force field
τ	orbital period
Ω	longitude of ascending node
ω	argument of periapse

SUPERSCRIPT NOTATION

$-$	vector
$\cdot$	$\frac{d}{dt}$
$\cdot\cdot$	$\frac{d^2}{dt^2}$

SUBSCRIPT NOTATION

$--$	vector

1.0 STATEMENT OF THE PROBLEM

This monograph considers the n-body problem, its application to the problem of determining the motion of a spacecraft, and the technique available for its solution.

The formulation begins with the classic problem of n-bodies. Equations of motion are written for n point masses (with spherically symmetric force fields) interacting only through their mutual gravitational attractions. In an inertial reference frame, let m_i be the point masses and $\vec{r}_i$ designate their position vectors. Then

$$m_i \ddot{\vec{r}}_i = -k^2 m_i \sum_{j \neq i} m_j \frac{(\vec{r}_i - \vec{r}_j)}{|\vec{r}_i - \vec{r}_j|^3} \qquad , i = 1, \ldots, n \qquad (0.1)$$

where k^2 = the universal gravitational constant.

It can be easily shown (Section 2.0) that ten constants (or integrals) exist for these equations. Six of these constants show that the center of mass of the system (barycenter) moves rectilinearly with constant speed. Three more state that the angular momentum of the system is constant; and the tenth states that the sum of the kinetic and potential energy is constant. Considerable investigation has been directed toward the discovery of other constants of motion. To date, however, no other constants have been found. In fact, one reference, Baker, (Reference 1.1) states that Brun and Poincare have proved that no other algebraic integrals exist.

When considering the motion of a spacecraft in the light of the n-body problem, therefore, simplifying assumptions are made:

1. The mass of the spacecraft is negligible compared to the mass of the other bodies; and

2. the motions of the other n-1 bodies are known to a degree greater than that required of the motions for the n^{th}.

Due to numerous perturbations (Reference 1.2), the motion of the bodies in the solar system do not follow a path that can be represented precisely by an analytic formula. However, accurate predictions of future positions and velocities can be estimated from observations of their motion and from a knowledge of the nature of the perturbations. The reduction of observed data and the generation of an accurate estimate of a space trajectory was discussed in the previous monograph (Reference 1.3). (See also References 1.4 and 1.5).

Under these two assumptions, techniques for handling the equations and for integrating the equations numerically are discussed. This step is required since the complexity of the equations precludes an analytic solution and is

acceptable since a tabular array of positions and velocities can be substituted for a functional representation.

The monograph divides these discussions, referred to generally as special perturbations, into two basic sections. The first section deals with the formulation of the problem to effect the most accurate and/or the most efficient solution. The second section treats the problem of generating the best tabular array representing the solution.

The three currently prevalent techniques for formulating the equations of motion for numerical or approximate analytic solution are Cowell's method, Encke's method, and the variation of parameters. Cowell's method is the straightforward numerical integration of equation (0.1) to obtain position and velocity data. This method will, therefore, not be discussed at this time. Encke's method is based on the observation that for the case where the perturbing forces are small, the solution closely approximates conic motion. This observation allows the derivation from the conic reference to be integrated independent of predominant motion in the interest of accuracy and efficiency. The position and velocity in the conic can be found analytically and added to the perturbed position and velocity to obtain the true position and velocity. The third method, variation of parameters, is based on the assumption that elliptic orbital elements can be found and updated at each integration step so that the osculating (instantaneous) elements describe the position and velocity vectors of the true solution. These methods are well documented in the open literature (Reference 1.4), and are thus discussed in a straightforward expository manner. The strengths and weaknesses of the methods as well as the types of trajectories for which each method is suited will be discussed.

Having presented the formulation of the problem, the monograph will then cover specific numerical integration techniques that can be applied to the equations. The section on numerical integration is divided into four topics:

1. Quadrature
2. The Runge-Kutta family
3. Predictor-corrector methods
4. Starting procedures for predictor-corrector methods.

The first section treats quadrature or the numerical integration of a definite integral; an equivalent definition is that quadrature is the numerical integration of a first-order ordinary differential equation whose right-hand side is a function of the independent variable alone. The well-known Newton-Cotes and Gaussian formulas (Reference 1.7) are developed in this section. These discussions are followed with the discussion of the Runge-Kutta family of integration logics. This family differs from the above in that the right-hand side of the differential equation can be a function of both the dependent and independent variable. The third section discusses predictor-corrector methods. Predictor-corrector methods are generally multi-step methods of the form

$$y_{n+1} = a_0 y_0 + \ldots + a_n y_n + b_0 y_0' + b_n = y_n' + b_{n+1} y_{n+1}'$$

where the formula is "open" if $b_{n+1} = 0$ and "closed" otherwise. In the predictor-corrector technique one formula is used to predict y_{n+1} and another used to correct it. The corrector formula yields a significantly higher accuracy than the predictor formula alone. (It is noted that open formulas are used as predictors and closed formulas are used as correctors.) The methods discussed range from Adam's first predictor-corrector methods (Reference 1.8) to later and more sophisticated methods such as Nordsieck (Reference 1.9) and Lanczos (Reference 1.10). The discussions of numerical integration concludes with a presentation of starting procedures. This presentation is essential to the discussions of this monograph since predictor-corrector methods are generally not self-starting. Important integration formulas from each family are discussed in detail; the equations are derived; the methods of implementing the equations, and the accuracy that can be expected are covered.

Finally, the numerical integration techniques are evaluated and compared, and specific methods are recommended.

2.0 STATE OF THE ART

2.1 THE N-BODY PROBLEM

The n-body problem considers the motion of n bodies in inertial space interacting through their mutual gravitational attractions. Historically, in Astronomy, one of the bodies, (e.g., a comet) has been of negligible mass, and the motion of the other n-1 bodies has been known to a good degree. Thus, the motion of the only one of the n-bodies need be determined. Modern applications involving the motion of spacecraft are simple extensions of this material differing primarily in the nature of the forces (need not be conservative). Thus, in current usage, the n-body problem has come to mean the problem of determining the motion of a spacecraft under the influence of gravitational forces (including oblateness), atmospheric forces, thrust, and solar radiation pressure.

2.1.1 FORCES CONSIDERED

The forces considered in the n-body problem can be many and varied and are not necessarily restricted to gravitational attractions. The following paragraphs will present discussions of some of the more important accelerations. The resultant acceleration is then defined as:

$$\ddot{\vec{r}} = \sum \vec{A}_i$$

where each $\vec{A}_i$ is the acceleration per unit mass acting on the spacecraft. The form of some of the more common of these accelerations will be discussed in the following paragraphs.

2.1.1.1 Gravitational Accelerations

Generally, the origin of the reference system will be fixed at the center of the primary attracting body. The acceleration of the spacecraft due to the primary body is

$$\vec{A} = -\mu \frac{\vec{r}}{r^3}$$

where μ = gravitational constant of primary body

$\vec{r}$ = position of the spacecraft

$r = |\vec{r}|$

If the attractions due to other bodies are to be considered, let $\vec{r}_B$ and μ_B represent the position vector and the gravitational constant of the perturbing body. Then, its effect on the spacecraft can be written

$$A = -\mu_B \frac{(\vec{r} - \vec{r}_B)}{|\vec{r} - \vec{r}_B|^3}$$

Other perturbing bodies are treated similarly.

Often, it is desirable to include the effect of a non-spherical Earth (or other body); the gravitational potential is generally written as an expansion in Legendre polynomials (see Reference 2.8). Vinti's potential function for the Earth is:

$$\ddot{x} = -\frac{\mu x}{r^3} + \left[J_2 \left(\frac{R}{r}\right)^2 \frac{3}{2} \left(1 - 5\frac{z}{r^2}\right) \right.$$
$$+ J_3 \frac{z}{r} \left(\frac{R}{r}\right)^3 \frac{5}{2} \left(3 - 7\frac{z^2}{r^2}\right)$$
$$+ J_4 \left(\frac{R}{r}\right)^4 \frac{5}{8} \left(-3 + 42\frac{z^2}{r^2} - 63\frac{z^4}{r^4}\right)$$
$$\left. + J_5 \left(\frac{R}{r}\right)^5 \frac{z}{r} \frac{1}{8} \left(-693\frac{z^4}{r^4} + 630\frac{z^2}{r^2} - 105\right) + \ldots \right]$$

$$\ddot{y} = \ddot{x} \frac{y}{x}$$

$$\ddot{z} = -\frac{\mu z}{r^5} \left[J_2 \left(\frac{R}{r}\right)^3 \frac{3}{2} \left(3 - 5\frac{z}{r}\right) \right.$$
$$+ J_3 \left(\frac{R}{r}\right)^3 \frac{r}{z} \frac{3}{2} \left(-1 + 10\frac{z^2}{r^2} - \frac{35}{3}\frac{z^4}{r^4}\right)$$
$$+ J_4 \left(\frac{R}{r}\right)^4 \frac{5}{8} \left(-15 + 70\frac{z^2}{r^2} - 63\frac{z^4}{r^4}\right)$$
$$\left. + J_5 \left(\frac{R}{r}\right)^5 \frac{r}{z} \frac{1}{8} \left(15 - 315\frac{z^2}{r^2} + 945\frac{z^4}{r^4} - 693\frac{z^6}{r^6}\right) + \ldots \right.$$

where the J_i's are harmonic coefficients.

2.1.1.2 Atmospheric Lift and Drag Forces

The expressions for lift and drag are taken from References 2.0 and 2.2.

$$\ddot{x} = D_o^2 \dot{s}^2 \Big\{ \Big[\mu' \gamma(v)\, \sigma(H)\, \gamma(\sigma)\, v \; \frac{v_x}{\dot{s}^2}$$

$$- \mu' \Big(A + \frac{B}{\dot{s}} \Big) \; \frac{f(r)}{D_o^2} Q_x \Big]$$

$$- \mu' \gamma(v)\, \sigma(H)\, \gamma(\sigma) \; \frac{V^2}{\dot{s}^2} \; \frac{C_L}{C_{D_o}} \Big[\Big(\hat{r} \times \frac{\vec{v}}{v} \Big)_x \sin \xi$$

$$+ \Big\{ \frac{\vec{v}}{v} \times \Big(\hat{r} \times \frac{\vec{v}}{v} \Big) \Big\}_x \cos \xi \Big] \Big\} \quad \ddot{x} \to \ddot{y}, \ddot{z}$$

where the vehicle velocity relative to a rotating atmosphere with cross winds is given by

$$v_x = \dot{x} + y\Omega_e + q(\cos\alpha \sin\phi' \cos\beta + \sin\alpha \sin\beta)$$

$$v_y = \dot{y} - x\Omega_e + q(\sin\alpha \sin\phi' \cos\beta - \cos\alpha \sin\beta)$$

$$v_z = \dot{z} - q \cos\phi' \cos\beta$$

where A = constant fitted to the Mach number variation of the drag coefficient with a mean sonic speed = 1

A_0 = initial projected frontal area of the vehicle

B = constant fitted to Mach number variation of the drag coefficient with a mean sonic speed

$$\cong C_s \Big(\frac{C_{D_T}}{C_{D_o}} - 1 \Big)$$

C_{D_0} = reference (hypersonic continuum) value of the drag coefficient (0.92 for a sphere, 1.5 for a typical entry capsule)

C_L = lift coefficient

C_s = local sonic speed in terms of surface circular satellite speed

$D_0^2 = C_{D_0} A_0 \; {}_0 V_{C0}{}^2 / 2 g_0 m_0$

$f(r) = \mu_0 D_0^2 \sigma \gamma(\sigma)$

g_0 = acceleration of gravity at unit distance (surface of Earth)

H = altitude above an oblate Earth = r - 1

$$+ f \sin^2\phi' + \frac{f^2}{2} \Big(\frac{1}{\pi} - \frac{1}{4} \Big) \sin^2 2\phi' + \ldots$$

where the flattening f = $\frac{1}{298.3}$ (units of Earth radii)

m = mass of space vehicle

$\vec{Q}$ = unit vector in the orbit plane perpendicular to the line of apsides

q = speed of the cross wind measured in a system rotating with Earth's angular rate (units of surface circular satellite speed V_{CO})

r = radius from the geocenter to the vehicle

$\dot{s}$ = speed of the vehicle with respect to an inertial frame, directed along Q

V_{CO} = surface speed for circular orbit

x,y,z = equatorial coordinates in units of equatorial Earth radii

α = right ascension of the vehicle (radians)

β = azimuth of the direction from which the wind is coming

$\gamma(v)$ = C_D $(v/C_s)/C_{D_0}$, the drag coefficient variation with Mach number

$\gamma(\sigma)$ = C_D $(\alpha)/C_{D_0}$, the drag coefficient variation in the transitional regime

Ω_e = constant relating to the rotational rate of the Earth, 0.058834470

μ' = m_0/m

ξ = bank angle

ρ = atmospheric density, kg/m^3

ρ_o = "sea level" atmospheric density, 1.225 kg/m^3

$\sigma = \frac{\rho}{\rho_o}$

ϕ' = geocentric latitude, radians

2.1.1.3 Solar Radiation Pressure

The solar radiation pressure may become significant on large area to mass vehicles or on missions with long flight times.

The magnitude of acceleration due to solar pressure acting on some surfact of the satellite is given (Reference 1.2) as

$$|\vec{A}_s| = \frac{P}{C}(1+R)\cos^2\alpha\,\frac{A}{m}$$

where α = angle of incidence with respect to surface normal

C = speed of light

R = reflection coefficient (=1 for specular reflection and = 0 for complete absorption)

P = power of the incident light

A = area of surface

m = mass of spacecraft

2.1.1.4 Thrust

The thrusting acceleration is handled very simply

$$\vec{A} = \frac{1}{m}(T_x, T_y, T_z$$

where T_x, T_y, T_z are the x, y, and z components of the thrust

m = mass of the spacecraft.

2.1.1.5 Electromagnetic Forces

As a satellite moves through a partly ionized medium, the incident flux of electrons on the satellite surface is larger than the ion flux, so that the satellite acquires a negative potential. On the day-side of the Earth, this effect is opposed by the photoejection of electrons. Jastrow (Reference 2.3) estimates that the satellite potential may approach -60 volts on the day-side and will not be greater than -10 volts on the night-side.

In addition to the potential acquired by ionic collision, the motion of a conducting satellite through the magnetic field of the Earth causes the satellite to acquire a potential gradient which is proportional to the strength of the magnetic field and the velocity of the satellite. The interaction of the electric currents thus induced in the satellite skin with the magnetic field causes a magnetic drag to act upon the satellite; this drag is proportional to the cube of the satellite dimensions.

If these forces are found not to be negligible, they can be included directly by the use of Maxwell's equations or indirectly by use of an atmospheric model which takes the effects into account.

2.1.1.6 Relativistic Effects

Perturbations caused by relativity are of the order $\alpha = \frac{V_0^2}{c^2} = \frac{\mu}{rc^2}$, where c is the speed of light. Since α is a very small quantity and any measurable deviations occur only after a long period of time, relativistic effects can usually be ignored in the case of Earth satellites. A modification of Newton's law, as a consequence of the theory of relativity, can be found in Danby (Reference 2.4).

2.1.2 BREAKDOWN OF ANALYTIC APPROACH

Consider a system of n point masses, m_i, at $\vec{r}_i$, where $i = 1, 2, \ldots, n$ and the $\vec{r}_i$ are measured in an inertial reference frame. Define r_{ij} to be

$$r_{ij} = \left| \vec{r}_i - \vec{r}_j \right| = \left| \vec{r}_j - \vec{r}_i \right|$$

Then the equation of motion of m_i is

$$m_i \ddot{\vec{r}}_i = -k^2 m_i \sum_{j \neq i} m_j \frac{(\vec{r}_i - \vec{r}_j)}{r_{ij}^3} \tag{1.1}$$

where k^2 = universal gravitational constant. This is the classic n-body problem; only gravitational forces are considered and, therefore, the force field is conservative. There exists ten constants or integrals of the above system of equations as will be shown below. Suppose that all of the equations (1.1) above are added. Then the right-hand sides vanish, and the result is a sum

over i

$$\sum_{i=1}^{n} m_i \ddot{\vec{r}}_i = 0$$

which can be immediately integrated to give

$$\sum_{i=1}^{n} m_i \dot{\vec{r}}_i = \vec{A}t + \vec{B}$$

where $\vec{A}$ and $\vec{B}$ are constant vectors.

This result shows that the center of mass of the system moves with respect to the inertial system of reference in a straight line with constant speed. The origin can, therefore, be set at the center of mass. The constant vectors $\vec{A}$ and $\vec{B}$ provide six constraints of motion.

Next, cross each of the vectors $\vec{r}_i$ into (1.1) and add the n resulting equations. Again, the terms on the right-hand side cancel, i.e.,

$$\sum_{i=1}^{n} m_i \vec{r}_i \times \ddot{\vec{r}}_i = 0$$

Thus, upon integration

$$\sum_{i=1}^{n} m_i \vec{r}_i \times \dot{\vec{r}}_i = \vec{h}$$

where $\vec{h}$ is a constant vector; $\vec{h}$ provides three more constants of motion. The tenth constant comes from the total energy of the system.

Define U, the force function, by

$$U = k^2 \sum_{i<j=1}^{n} \frac{m_i m_j}{r_{ij}}$$

Then

$$\frac{\partial U}{\partial x_i} = k^2 m_i \frac{\partial}{\partial x_i}\left[\sum_{j=1}^{n} \frac{m_j}{r_{ij}}\right]$$

$$= -k^2 m_i \sum_{j=1}^{n} m_j \left(\frac{x_i - x_j}{r_{ij}^3}\right)$$

Therefore, equation (1.1) can be written

$$m_i \ddot{\vec{r}}_i = \vec{\nabla}_i U \tag{1.2}$$

where

$$\vec{\nabla}_i = \hat{i}\frac{\partial}{\partial x_i} + \hat{j}\frac{\partial}{\partial y_i} + \hat{k}\frac{\partial}{\partial z_i}$$

where $\hat{i}$, $\hat{j}$, $\hat{k}$ are unit vectors in the directions of the x, y, and z coordinate axes. -U is the total potential energy of the system where the zero level of potential energy is attained when all of the particles are infinitely separated.

Now, if each $\dot{\vec{r}}_i$ is dotted into equation (1.2) and all of the n equations are added

$$\sum_{i=1}^{n} m_i \dot{\vec{r}}_i \cdot \ddot{\vec{r}}_i = \sum_{i=1}^{n} \dot{\vec{r}}_i \cdot (\vec{\nabla}_i U) = \frac{dU}{dt}$$

Integration gives

$$\frac{1}{2} \sum_{i=1}^{n} m_i |\dot{\vec{r}}_i|^2 = U + \text{constant}$$

The left-hand term is the kinetic energy T. The equation can be written $T = U + C$, where C is the tenth constant of the problem. Bruns and Poincare proved that there exist no other <u>algebraic</u> integrals of the n-body problem. A general solution, thus exists only for the two-body proboem. Investigations continue, however, to determine if new functions can be found which will yield additional constants of the motion; no such constants or functions have as yet been found.

2.2 SPECIAL PERTURBATION METHODS

2.2.1 COWELL'S METHOD

2.2.1.1 Discussion

Perhaps the simplest technique for the solution of n-body motion is the direct-numerical integration of the equations of motion; this method is referred to as Cowell's method.

Given the initial position and velocity vectors at a specified time, the system of three second-order equations can be integrated to give the position and velocity vectors at any subsequent time. In this form, the equations are very simple; and the integration time per step is nominal.

The Gauss-Jackson, or Σ^2, method of numerical integration has been frequently referred to in the literature as "Cowell's Method"; however, in this monograph, the ambiguity will be avoided, and Cowell's method will always refer to a means of formulating the equations of motion.

2.2.1.2 Derivation of Equations

There is no derivation. The equations are simply an expression of Newton's second law.

$$\ddot{x} = \frac{1}{m} \sum_{i=1}^{n} f_i$$

2.2.1.3 Method

The extreme simplicity of the equations is one of the strengths of Cowell's method. All that is required is the computation of the accelerations of the problem in some appropriate coordinate system and the recourse to numerical integration for stepping the function. The accelerations are integrated once to obtain velocities. Then the position vector can be obtained either by integrating the velocities or by integrating the accelerations using a second-order integration technique.

2.2.1.4 Advantages and Disadvantages

Because the total accelerations are being integrated, the attractions change rapidly with time so that small-integration steps are required to maintain accuracy. In addition, care must be exercised in choice of the number of significant figures maintained, or the effect of smaller accelerations will be lost with the resultant loss in accuracy in the solution.

Because of loss of numerical significance problems during integration and

because of the fact that Encke's method and the variation of parameters method are so much superior for those problems in which the perturbing forces are small, Cowell's method has been largely relagated to applications where the non-two-body accelerations change rapidly with time (e.g., during boost, or reentry).

2.2.2 ENCKE'S METHOD

2.2.2.1 Discussion

Encke's method is based on the observation that many n-body motions differ only slightly from two-body motion (which can be solved in closed-form). Thus, Encke, a German astronomer, proposed that the difference in the total acceleration and that of an approximately selected reference be integrated. The position and velocity in the reference conic are then obtained analytically and added to the integrated position and velocity vectors to get the total position and velocity vectors.

It should be carefully noted that absolutely nothing in the formulation requires the reference motion to be conic. Historically, a conic reference motion has usually been selected; however, any other convenient motion could be choses. (That is, secular and periodic motions resulting from the Earth's oblateness can be included in the reference to magnify the significance of a special set of perturbing forces.)

In the derivation presented in the following section, the perturbing force is expanded in an infinite series about the reference motion. If the force is expressed in this way (it need not be) then rectification of the reference motion (i.e., a new reference motion is calculated which is much closer to the actual motion and generally coincides with the true position and velocity at the instant of rectification) is performed whenever the actual motion deviates significantly from the reference motion so that accuracy in the evaluation of the force is retained. Such rectification would not be necessary if the force were expressed in a closed form. The series expansion, however, allows calculation of the force to a larger number of places when working with a fixed-word length computational procedure.

2.2.2.2 Derivation of Equations

In this section, the equations will be discussed for the case where the reference motion is conic.

The validity of the Encke approach follows from the linearity property of the integral. If u, v, ans w are continuous function of t on the interval $a \leq t \leq b$; and if $u(t) + v(t) = w(t)$ for $a \leq t \leq b$, then

$$\int_a^b w(t)\,dt = \int_a^b u(t)\,dt + \int_a^b v(t)\,dt$$

Thus, the accelerations experienced by the vehicle can be integrated in two separate parts and added to give the total position and velocity.

The subscript E will be used to denote the reference, or Encke, motion; thus, (x, y, z) denote the actual position of the vehicle; and (x_E, y_E, z_E) denote the position on the reference orbit. Let all of the non-two-body accelerations be lumped in the expression $\sum_i \ddot{x}_i$. Then, the equations of motion in an inertial system can be written

$$\ddot{x} = \frac{-\mu x}{r^3} + \sum_i \ddot{x}_i \qquad x \rightarrow y, z$$

where

$$r = \sqrt{x^2 + y^2 + z^2}$$

μ = gravitational constant of central body

Now, using the linearity property of the integral and noting that the first term in the acceleration is the equation of motion in the conic reference orbit

$$\ddot{x}_E = \frac{-\mu x_E}{r_E^3}$$

then, the deviations from the reference orbit are

$$\xi = x - x_E$$

$$\eta = y - y_E$$

$$\zeta = z - z_E$$

The accelerations actually integrates are thus

$$\ddot{\xi} = \ddot{x} - \ddot{x}_E \qquad \begin{cases} x \rightarrow y, z \\ \xi \rightarrow \eta, \zeta \end{cases}$$

$$= \mu\left(\frac{x_E}{r_E^3} - \frac{x}{r^3}\right) + \sum \ddot{x}_i$$

The bracketed acceleration is generally called the Encke acceleration, and the quantities within the bracket are very nearly equal; therefore, some rearranging is desirable to avoid loss of significance.

$$\ddot{\xi} = \frac{\mu}{r_E^3}\left[x_E - x\left(\frac{r_E}{r}\right)^3\right] + \sum_i \ddot{x}_i$$

$$= \frac{\mu}{r_E^3}\left\{x\left[1 - \left(\frac{r_E}{r}\right)^3\right] - \xi\right] + \sum_i \ddot{x}_i$$

Attention is now concentrated on the term

$$\left(\frac{r_E}{r}\right)^3$$

The term $\left(\frac{r}{r_E}\right)^2$ can be expanded as

$$\left(\frac{r}{r_E}\right)^2 = \frac{1}{r_E^2}\left[(x_E+\xi)^2+(y_E+\eta)^2+(z_E+\zeta)^2\right]$$

$$= \frac{1}{r_E^2}\left[r_E^2+2\xi x_E+\xi^2+2\eta y_E+\eta^2+2\zeta z_E+\zeta^2\right]$$

$$\approx 1+\frac{2}{r_E^2}\left[\left(x_E+\frac{\xi}{2}\right)\xi+\left(y_E+\frac{\eta}{2}\right)\eta+\left(z_E+\frac{\zeta}{2}\right)\zeta\right]$$

If the second term is now defined to be $2q$, i.e.,

$$q \triangleq \frac{1}{r_E^2}\left[\left(x_E+\xi/2\right)\xi+\left(y_E+\eta/2\right)\eta+\left(z_E+\zeta/2\right)\zeta\right]$$

$$= \frac{1}{r_E^2}\left[\hat{r}_E\cdot\Delta\hat{r}+\frac{\Delta\bar{r}\cdot\Delta\bar{r}}{2}\right]$$

then

$$\left(\frac{r}{r_E}\right)^2 = 1+2q$$

Therefore,

$$\left(\frac{r_E}{r}\right)^3 = (1+2q)^{-3/2}$$

Further, define

$$fq = 1-(1+2q)^{-3/2}$$

and expand fq in a binomial series

$$fq = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{(2k+1)!}{2^k(k!)^2}q^k \qquad , -\tfrac{1}{2}<q<\tfrac{1}{2}$$

and the Encke equation can be written

$$fq = 3q-\frac{3\cdot5}{2}q^2+\frac{5\cdot7}{2}q^3-\frac{5\cdot7\cdot9}{2^3}q^4+\frac{7\cdot9\cdot11}{2^3}q^5-\frac{7\cdot9\cdot11\cdot13}{3\cdot2^4}q^6+\frac{9\cdot11\cdot13\cdot15}{3\cdot2^4}q^7+\cdots$$

$$\ddot{\xi} = \frac{\mu}{r_E^3}(xfq-\xi)+\sum_i \ddot{\nu}_i \qquad \begin{matrix} x\rightarrow y,z \\ \xi\rightarrow\eta,\zeta \end{matrix}$$

When the deviations from the reference orbit are small, the series expression provides greater accuracy. In addition, a criterion for rectification can be developed from the series expression. Normally, a certain number of terms, say six, will be carried in the Encke series. The next term, the seventh,

could be monitored to determine when it becomes large enough to affect the final digit of the word, that is, when it becomes significant. When this occurs, the reference orbit should be rectified.

2.2.2.3 Method

The procedure followed in the computation of position and velocity histories via Encke's method is illustrated in the accompanying flow chart.

2.2.2.4 Advantages and Disadvantages

Since the deviations from the reference orbit are small, it is evident that integrating steps which are larger than those suitable for Cowell's method can be taken. In addition, greater accuracy should be possible when working with a fixed-word length because round-off and truncation errors produce small errors in the deviations; these errors are then "lost" when the deviations are added to the reference motion. Although the round-off error is less, Encke's method, as generally employed, involves expressions that are much more complicated and ofter less symmetric than Cowell's simple formulas. In addition, both the necessity of solving the two-body formulas at every step and the possible need for rectification introduces additional sources of error.

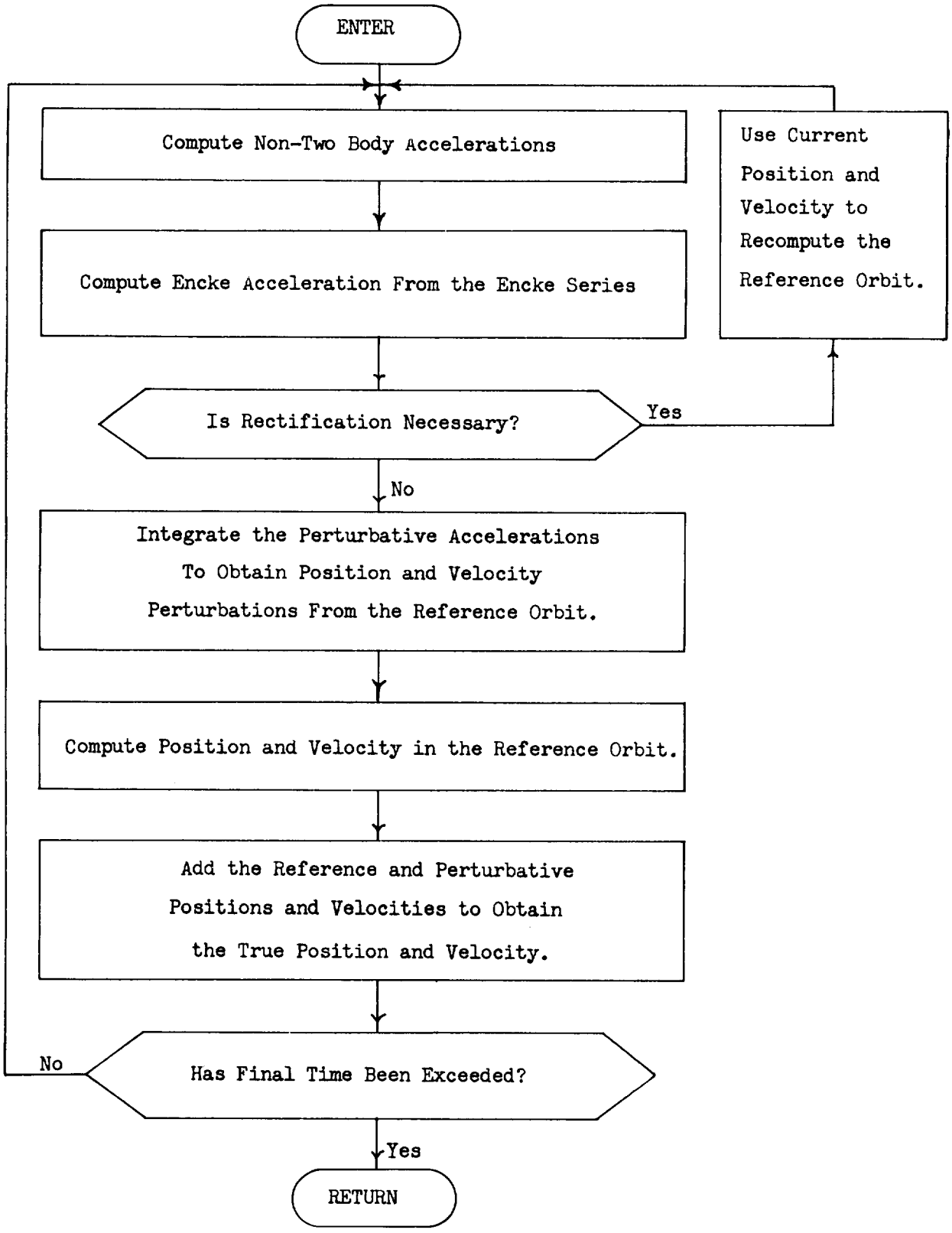
2.2.3 VARIATION OF PARAMETERS

2.2.3.1 Discussion

In a two-body problem, six orbital elements are sufficient to determine the subsequent motion. If non-two-body forces are acting, then an analytic expression for the motion cannot be found and the concept of "orbital elements" may have little meaning. It is possible, however, to use the analytic two-body equations to describe the motion if the existance of time-varying orbital elements is admitted. In this case, the orbital elements at a particular time describe a motion which would take place if the perturbing forces were removed at that particular instant. This instantaneous two-body orbit is tangent to the actual motion at all times and is often called the osculating orbit. The variation of parameters method expresses the time rate of change of the orbital elements as functions of the perturbing force. These derivatives are then integrated numerically to solve for the actual motion. Note the similarity to Encke's method in that only the perturbing forces are integrated. In the Encke method, however, there is only one reference orbit with particular orbital elements while in the variation of parameters there is a continuous set of elements for the reference motion.

In Section 2.2.3.3.1 the equations for the variations of the six parameters a, e, i, Ω, w and M_o are developed. These six elements are not the only ones which may be used for the variation of parameters technique. Indeed, the components of the initial position and velocity vectors can be employed in a manner similar to that presented in a previous monograph (Reference 2.8). In fact, any six independent constants of motion can be used. Furthermore, the motion need not be referenced to conic motion. An inspection of the final equations in Section 2.2.3.3.1 will reveal that for the parameters above the

Figure 2.1 Computational Logic for Encke's Method



equations becomes singular for parabolic orbits. This fact illustrates a basic difficulty in the variation of parameters technique. This difficulty can be overcome by switching to a more suitable set of variables whenever one set starts to fail but only at the expense of more complex programming. The different sets of parameters used by various people have been chosen to avoid singularities of a particular type. Variables that are free of nearly all singularities have been proposed by Garafalo (Reference 2.5) and Newton (Reference 2.6). These methods break down only when the angular momentum vanishes. Wong (Reference 2.7) proposes a method which is singular only when the magnitude of the position vector goes to zero.

2.2.3.2 Derivation of Equations

The following derivation is taken from Reference 2.1.

For the variation-of-parameters method any six independent parameters could be chosen. In this derivation the parameters, a, e, i, Ω, w, and M_o will be used.

a = semi-major axis

e = orbital eccentricity

i = orbital inclination

Ω = right ascension of the ascending mode

w = argument of periapse

M_o = mean anomaly of epoch

If no force other than a central force were acting on a satellite, its orbit would be a conic section and would be completely described by this act of six quantities; further, these quantities would be constants. If a non-central force acts on the satellite, then the six orbital elements will not be constant but will vary with time. The set of differential equations which govern the time history of the elements is equivalent to the Newton or Lagrange set.

In order to derive the equations, a coordinate system in which the non-two-body forces are expressed and must be chosen. The coordinate system which will be used will have unit vectors $\hat{n}_R$, $\hat{n}_S$, and $\hat{n}_\omega$ with the center of the system at the active focus of the instantaneous two-body orbit. The vector $\bar{R}$ is directed away from the focus along the instantaneous radius vector, $\bar{S}$ is normal to R in the orbit plane in the direction of increasing true anomaly (i.e., $\bar{v} \cdot \hat{S} > 0$), and $\bar{W}$ is normal to the orbit plane (along the angular-momentum vector) completing the right-handed set. The perturbing force vector can now be expressed in this system as:

$$\bar{F} = m(R\bar{n}_R + S\bar{n}_S + W\bar{n}_\omega)$$

If $\bar{V}$ is now used to denote the instantaneous velocity vector

$$\bar{V} = \dot{r}\,\bar{n}_r + r\dot{\theta}\,\bar{n}_v = \dot{\theta}\left(\frac{dr}{d\theta}\bar{n}_r + r\bar{n}_s\right)$$

If an expression for $\dot{\theta}$ is obtained as a function of the radius from the conservation of the angular momentum

$$h = r^2\dot{\theta} = \sqrt{\mu p} = na^2\sqrt{1-e^2} \tag{2.1}$$

where
$$\therefore \quad \dot{\theta} = \frac{na^2}{r^2}\sqrt{1-e^2} \tag{2.2}$$

$$n = \sqrt{\frac{\mu}{a^3}} \tag{2.3}$$

and if an expression for $dr/d\theta$ is obtained by differentiating the equation of a conic in polar form

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

$$\frac{dr}{d\theta} = \frac{p\sin\theta}{(1+e\cos\theta)^2} = \frac{r\sin\theta}{1+e\cos\theta} \tag{2.4}$$

the expression for velocity becomes [from (2.2) and (2.4)]

$$\bar{V} = \frac{na^2\sqrt{1-e^2}}{r^2}\left(\frac{re\sin\theta}{1+e\cos\theta}\,\bar{n}_r + r\bar{n}_s\right).$$

The time rate of change of energy per unit mass (ϵ) is found from the dot product of the perturbing force per unit mass vector $\bar{F}$ and $\bar{V}$.

$$\frac{d\epsilon}{dt} = \frac{\bar{F}}{m}\cdot\bar{V} \tag{2.5}$$

But, a second expression for energy per unit mass is obtained by the addition of the kinetic energy and potential energy as

$$\epsilon = \frac{-\mu}{2a} \tag{2.6}$$

Thus, differentiation of (2.5) and substituting (2.6) gives an expression for the time rate of change of the semi-major axis. This equation is the first in the set describing the variation of the "constants" of integration (osculating elements):

$$\frac{da}{dt} = \frac{2e\sin\theta}{n\sqrt{1-e^2}}R + \frac{2a\sqrt{1-e^2}}{nr}S \tag{2.7}$$

To derive the changes in the other orbital elements, it is necessary to know the rate at which the angular momentum vector $\bar{h}$ changes. The rate of change of $\bar{h}$ is known to be equal to the summation of the external moments acting on the body in orbit, i.e.,

$$\frac{d\bar{h}}{dt} = \frac{1}{m}(\bar{r} \times \bar{F})$$
$$= r n_r \times (R\bar{n}_r + S\bar{n}_s + W\bar{n}_w)$$
$$= rS\bar{n}_w - rW\bar{n}_s \qquad (2.8)$$

But the time derivative of a vector may be expressed as

$$\frac{d\bar{h}}{dt} = \frac{dh}{dt}\bar{n}_w + \frac{h\,d\alpha}{dt}\bar{n}_s \qquad (2.9)$$

where d α is the angle through which the angular momentum vector is rotated in time dt. Comparing (2.8) and (2.9) it is seen that

$$\frac{dh}{dt} = rS \quad \text{and} \quad \frac{-rW}{h} = \frac{d\alpha}{dt} \qquad (2.10)$$

Now, the eccentricity of the orbit may be expressed in terms a and h through equations (2.1) and (2.3) as

$$e = \left(1 - \frac{h^2}{\mu a}\right)^{1/2} = (1 - p/a)^{1/2}$$

Thus, differentiation yields the following

$$\frac{de}{dt} = \frac{-h}{2\mu a e}\left(\frac{2\,dh}{dt} - \frac{h}{a}\frac{da}{dt}\right)$$
$$= \frac{-\sqrt{1-e^2}}{2na^2e}\left(2\frac{dh}{dt} - na\sqrt{1-e^2}\,\frac{da}{dt}\right) \qquad (2.11)$$

Upon substituting equations (2.7) and (2.10) for da/dt and dh/dt, equation (2.11) takes the final form

$$\frac{de}{dt} = \frac{\sqrt{1-e^2}\sin\theta}{na}R + \frac{\sqrt{1-e^2}}{na^2e}\left[\frac{a^2(1-e^2)}{r} - r\right]S \qquad (2.12)$$

The motion of the node is the same as the motion of the projection of $\bar{h}$ on the equatorial plane (See Figure 2.3.)

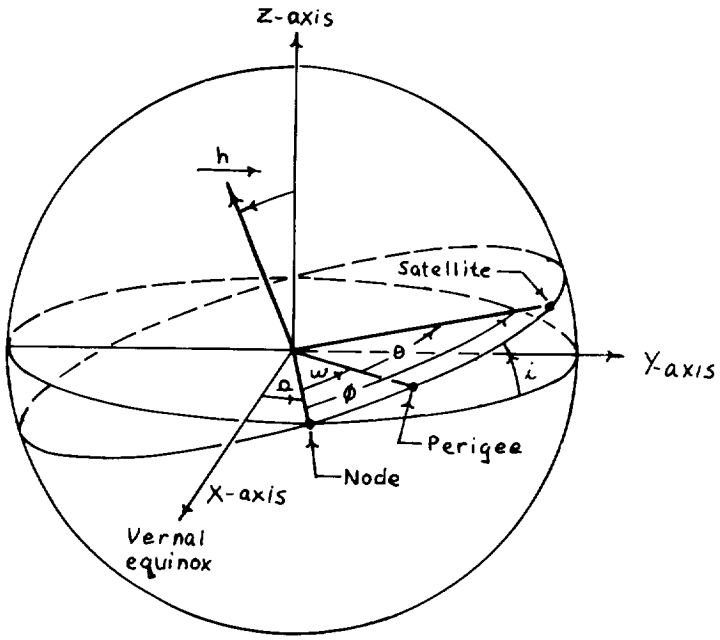


Figure 2.2 Orbital Orientation

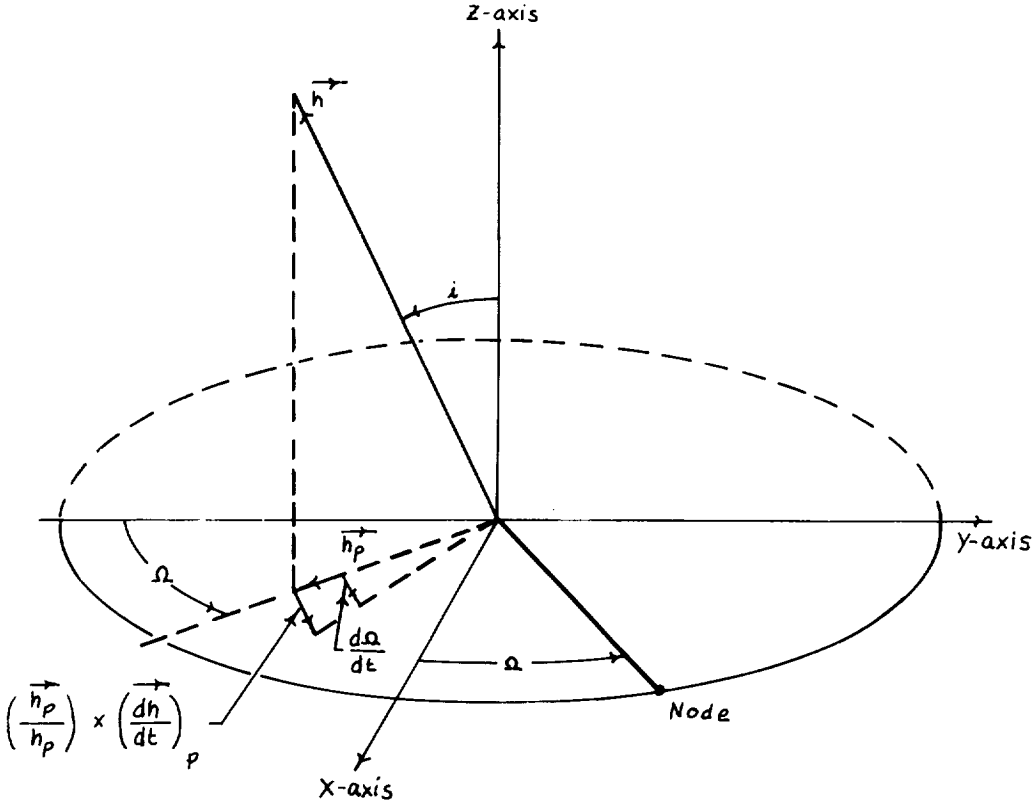


Figure 2.3 Orbital Plane Geometry

Let the subscript p denote the projection of any vector on the equatorial plane. Then it can be seen that

$\bar{h}_p$ = protection of $\bar{h}$ on the equatorial plane

$\left(\frac{d\bar{h}}{dt}\right)_p$ = projection of dh/dt on the equatorial plane

$\left(\frac{\bar{h}_p}{h_p}\right) \times \left(\frac{d\bar{h}}{dt}\right)$ = component of $\left(\frac{d\bar{h}}{dt}\right)_p$ which is normal to $\bar{h}_p$

By referring to Figure (2.3), it can be seen that

$$\frac{d\Omega}{dt} = \frac{\left| \frac{\bar{h}_p}{h_p} \times \left(\frac{d\bar{h}}{dt}\right)_p \right|}{h_p}$$

or using equations (2.9) and (2.11)

$$\frac{d\Omega}{dt} = \frac{\left| \bar{h}_p \times \left(\frac{h\, d\alpha}{dt} \bar{n}_s\right)_p \right|}{h_p^2} \tag{2.13}$$

But

$$\bar{h}_p = h \sin i \left(\hat{i} \sin\Omega - \hat{j} \cos\Omega\right)$$

where $\hat{i}$ and $\hat{j}$ are unit vectors along the X and Y axes, respectively, and

$$\left(h \frac{d\alpha}{dt} \bar{n}_s\right)_p = -rw \left[\hat{i} \left(-\cos\phi \cos i \sin\Omega - \sin\phi \cos\Omega\right) + \hat{j} \left(-\sin\phi \sin\Omega + \cos\phi \cos i \cos\Omega\right) \right]$$

Thus, on performing the cross product, equation (2.13) becomes

$$\frac{d\Omega}{dt} = \frac{rw \sin\phi}{ma^2 \sqrt{1-e^2} \sin i}$$

The change in orbital inclination is related to the change in the node. This fact can be seen by referring to Figure (2.4) in which

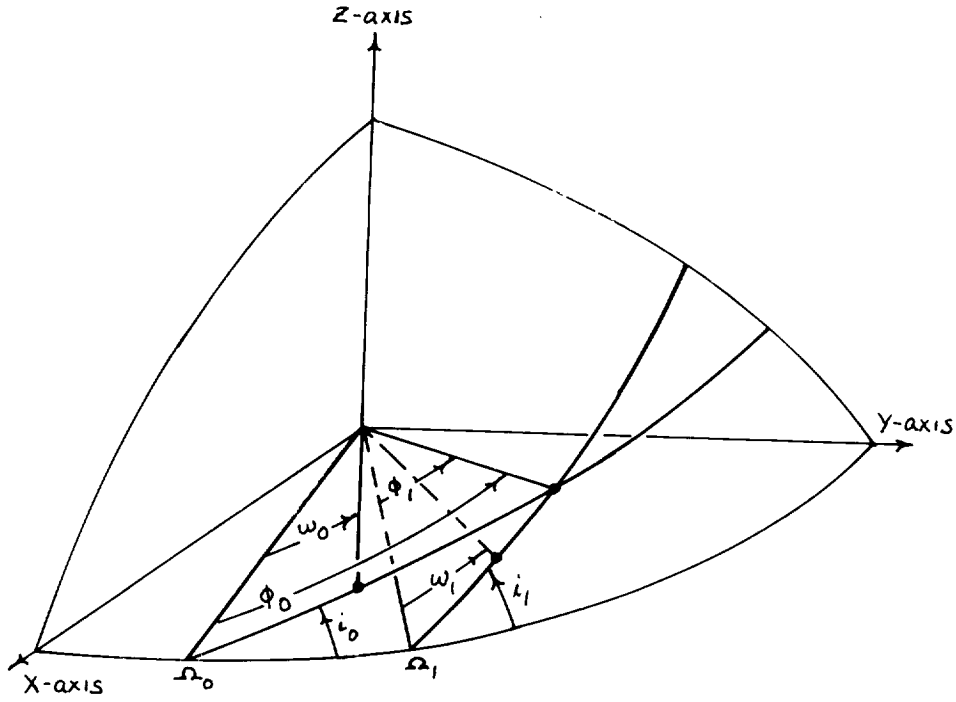


Figure 2.4. Variation of Node

two positions of the node, Ω_0 and Ω_1, are shown with

$$\Delta\Omega = \Omega_1 - \Omega_0$$

and

$$\Delta i = i_1 - i_0$$

But these angles are related by spherical trigonometry as

$$\sin\Delta i = \sin i_1 \cos i_0 - \sin i_0 \cos i_1$$

$$= \frac{\sin i_0}{\sin\phi_1}\left[\cos i_0 \sin\phi_0(1-\cos\Delta\Omega) + \cos\phi_0 \sin\Delta\Omega\right]$$

Differentiation of this equation and investigation of the limit as $\Delta\Omega \to 0$ yields the following equation

$$\frac{di}{dt} = \frac{\sin i}{\sin\phi}\cos\phi\frac{d\Omega}{dt}$$

Therefore,

$$\frac{di}{dt} = \frac{rW\cos\phi}{na^2\sqrt{1-e^2}}$$

The change in the argument of perigee, ω, arises from two sources. One is the motion of perigee caused by the forces in the orbital plane tending to rotate the ellipse in its plane. The other change occurs because ω is measured from the moving node (Figure 2.4). To evaluate the latter changes, assume that the in-plane perturbing forces are zero. Then the change in ω equals the change in ϕ. According to the relations in a spherical triangle,

$$\cos\phi_1 = \cos\Delta\Omega\cos\phi_0 + \sin\Delta\Omega\sin\phi_0\cos i_0$$

Differentiating this expression and taking the limit as $\Delta\Omega \to 0$ now yields

$$\frac{d\phi}{dt} = \left(\frac{dw}{dt}\right)_W = -\cos i\frac{d\Omega}{dt} = \frac{-r\sin\phi\cot i}{na^2\sqrt{1-e^2}}W$$

where the subscript w means that this is the change in the argument of periapse contributed by the nodal motion which is caused by the component of the perturbing acceleration normal to the orbit plane.

The change in the argument of periapse caused by the in-plane components, R and S, is denoted by $(dw/dt)_{R,S}$. The effect of these in-plane forces is to change the instantaneous velocity vector which must, at every instant, remain tangent to the instantaneous osculating ellipse. This ellipse will, therefore, have a changing perigee position. The resulting rate of change of the argument of perigee will be

$$\left(\frac{dw}{dt}\right)_{R,S} = -\frac{d\theta}{dt} \qquad (2.14)$$

Here $d\theta/dt$, the rate of change of the true anomaly caused by the perturbing force, must not be confused with $\dot{\theta}$ which is the rate of change of θ in an upper-turbed Kepler orbit. To evaluate $d\theta/dt$, refer to Figure 2.3. After the force $m(R\bar{n}_r + S\bar{n}_s)$ has been applied for the time dt; the velocity vector is changed from $\bar{V}$ to $\bar{V} + d\bar{V}$, th- true anomaly from θ to $\theta + d\theta$, and the angle, γ, between $\bar{n}_s$ abd v_s is changed from γ to $\gamma +$ d . The expression from the angular momentum

$$h = rV\cos\gamma$$

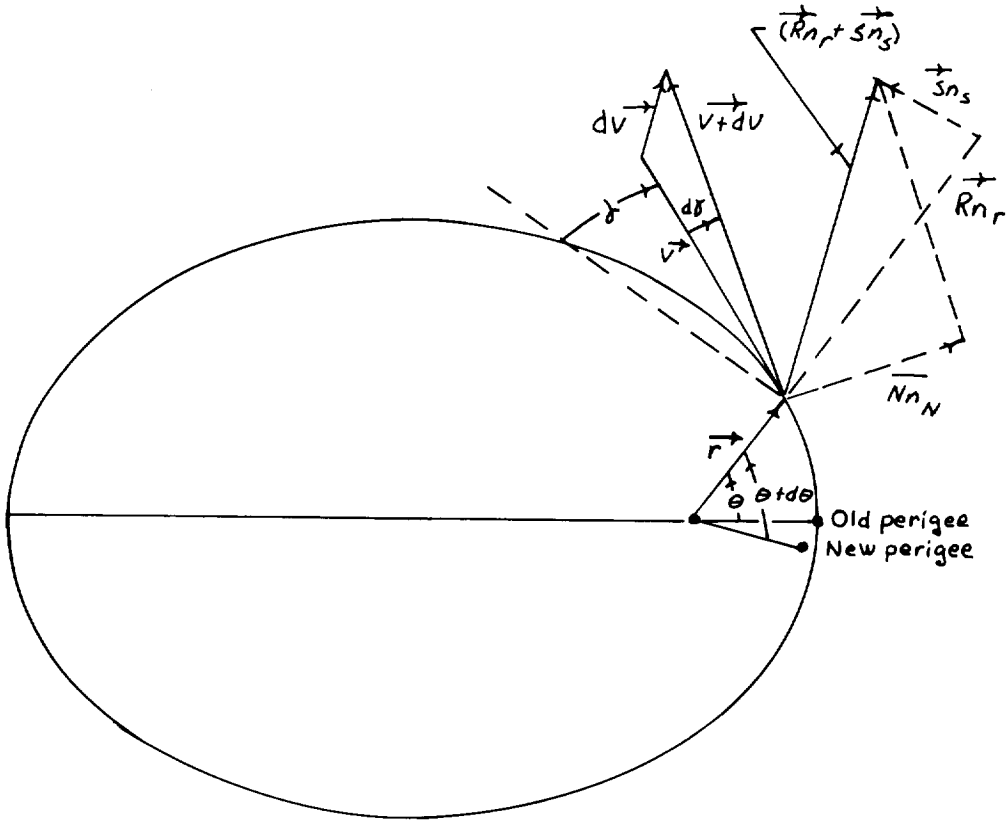


Figure 2.5. Variation of True Anomaly ($\frac{d\theta}{dt}$)

Since $h = r^2\dot{\theta}$ and $v = \sqrt{\dot{r}^2 + r^2\dot{\theta}^2}$, it follows that

$$\cos\gamma = \left(1 + \frac{1}{r^2}\left(\frac{dr}{d\theta}\right)^2\right)^{-1/2}$$

Computing $dr/d\theta$ from equation (2.4) yields

$$\cos\gamma = \frac{1 + e\cos\theta}{\sqrt{1+e^2+2e\cos\theta}}$$

and

$$\sin\gamma = \frac{e\sin\theta}{\sqrt{1+e^2+2e\cos\theta}} \qquad (2.15)$$

Differentiating equation (2.15) with respect to time and using equation (2.14), it is found that

$$\left(\frac{d\omega}{dt}\right)_{R,S} = \left[\frac{1+e^2+2e\cos\theta}{e(e+\cos\theta)}\left(\frac{\sin\theta}{1+e^2+2e\cos\theta}\frac{de}{dt} - \frac{d\gamma}{dt}\right)\right]$$

Now, if N is defined to be the component of the force normal to $\bar{V}$, then

$$d\gamma = \frac{N\,dt}{V}$$

But

$$V = R\cos\gamma - S\sin\gamma$$

and

$$V = \frac{h}{r}\frac{\sqrt{1+e^2+2e\cos\theta}}{1+e\cos\theta}$$

Therefore,

$$\frac{d\gamma}{dt} = \left\{\frac{r(1+e\cos\theta)}{h(1+e^2+2e\cos\theta)}\left[R(1+e\cos\theta) - (e\sin\theta)S\right]\right\} \qquad (2.16)$$

Equation (2.16) along with equation (2.12) for de/dt, yields

$$\left(\frac{d\omega}{dt}\right)_{R,S} = \frac{\sqrt{1-e^2}}{nae}\left[-(\cos\theta)R + \sin\theta\left(1+\frac{1}{1+e\cos\theta}\right)S\right]$$

The total rate of change of the argument of perigee is

$$\frac{d\omega}{dt} = \left(\frac{d\omega}{dt}\right)_{\omega} + \left(\frac{d\omega}{dt}\right)_{R,S}$$

The final element, mean anomaly at epoch, defines the position of the satellite in the osculating orbit at any time also has a time rate. This relationship is obtained directly from Kepler's equation

$$M_0 = E - e\sin E - nt$$

and can be found by using the equations already obtained for de/dt and dθ/dt, with the relationship between E and θ given by

$$\cos\theta = \frac{\cos E - e}{1 - e\cos E}$$

$$\sin\theta = \frac{\sqrt{1-e^2}\sin E}{1 - e\cos E}$$

The result is

$$\frac{dM_0}{dt} = \frac{-1}{na}\left(\frac{2r}{a} - \frac{1-e^2\cos\theta}{e}\right)R - \frac{(1-e^2)}{nae}\left[1 + \frac{r}{a(1+e^2)}\right](\sin\theta)S - t\frac{dn}{dt}$$

The complete set of equations is summarized below.

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

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

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

$$\frac{di}{dt} = \frac{r\cos\phi}{na^2\sqrt{1-e^2}}W \qquad (2.17)$$

$$\frac{dw}{dt} = \frac{r\sin\phi\cos i}{na^2\sqrt{1-e^2}}W - \frac{\sqrt{1-e^2}\cos\theta}{nae}R + \frac{\sqrt{1-e^2}}{nae}\left(1 + \frac{1}{1+e\cos\theta}\right)\sin\theta\, S$$

$$\frac{dM_0}{dt} = -\frac{1}{na}\left(\frac{2r}{a} - \frac{1-e^2}{e}\cos\theta\right)R - \frac{(1-e^2)}{nae}\left(1 + \frac{r}{a(1-e^2)}\right)\sin\theta\, S - t\frac{du}{dt}$$

2.2.3.3 Method

The sequence of operations to be performed to use the variation of parameters method is listed below. A block diagram depicting this sequence is also included:

1. From the initial conditions, i.e., a position and velocity vector at $t = t_o$, compute the orbital elements a, e, i, Ω, w, M_o.

2. Compute the force in the RSW coordinate system. This computation will generally involve computing a transformation to the RSW system from an inertial system in which the forces are expressed.

3. Using the force components from (2) and the current values of the orbital elements, compute the time rate of change of the orbital elements as given by equation (2.17) in Section 2.1.3.3.2.

4. Obtain a solution for the r changes in the orbital elements by numerical integration time $(t_o + \Delta t)$.

5. Obtain new orbital elements by addition of changes in the orbital elements to the previous values.

6. From the orbital emements at $(t_o + \Delta t)$ compute the position and velocity vector at $t_o + \Delta t$.

7. Repeat steps (2 thru 6) until the desired end conditions are reached.

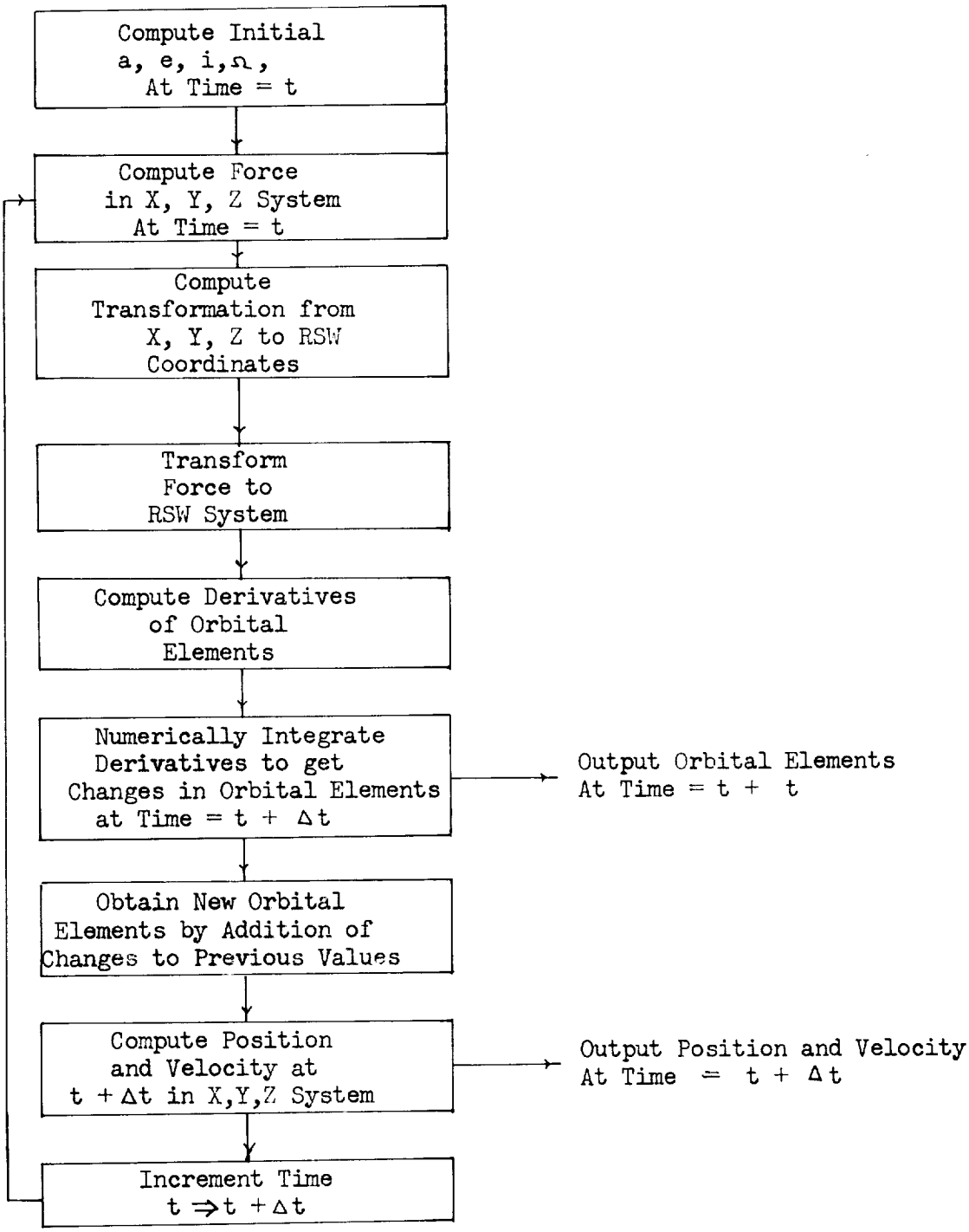


Figure 2.6. Computational Logic for Variation of Parameters Method

2.2.3.4 Advantages and Disadvantages

The essential characteristic of this method is that the integration is carried out on parameters which are changing much more slowly than either the rectangular coordinates or the perturbations relative to a fixed reference trajectory. Further, since they vary slowly, the error accumulation from the calculation of the derivative is, for a long time, far beyond the last significant digit of the initial calculation. Thus, it is expected that truncation error would appear only for very large integration steps. Simultaneously, the round-off problem is reduced since the number of steps is reduced with the next result being that much larger integration steps can be taken for a given accuracy. A disadvantage is that the programming and numerical analysis involved in this method are the most complicated of the three methods discussed. Because of this, the computation time per integration step is at least twice as long as for a Cowell method. As was mentioned in another section, any particular choice of variables will exhibit singularities for certain conditions and, therefore, each case requires special consideration thus detracting from the usefulness of parameter methods as a general integration technique. The variation of parameters method is primarily applicable to missions in which small perturbations act throughout the orbit, e.g., microthrust transfer.

2.2.3.5 Comparison

The three most common mathematical formulations in current use are: Cowell's method, Encke's method, and the variation of parameters.

Cowell's method is the straightforward integration of the equations of motion while Encke's method and the variation of parameters involve a mathematically more sophisticated approach. The integration times per step are smaller for this approach than for either of the other approaches because of the simplicity of the formulation. However, because of the total accelerations are being integrated and because the attractions change rapidly with small time changes, a small integration step is required to maintain accuracy and a large number of significant figures must be carried to prevent loss of numerical significance.

For many types of trajectories, the Cowell method requires about ten times as many integration steps as the Encke or variation of parameter methods; and although computing time per step is approximately one-half, the overall computing time may be considerably greater. In addition, since round-off errors accumulate as some power of the number of steps, the Cowell method can be expected to be more susceptible to accuracy degradation as a result of accumulation of round-off error. Thus, for many orbits, including lunar orbits, which can be closely approximated by two-body motion, the Cowell method is the least accurate and least efficient computationally of the three methods.

However, there are some trajectories for which the Cowell method is well suited; in particular, those in which the perturbation accelerations are changing rapidly. This type of trajectory occurs, for example, during boost

and reentry. In addition, this method is universally applicable to all types of orbits and presents no fundamental difficulties in the exceptional cases of nearly circular, nearly parabolic, or hyperbolic orbits.

Encke's method has recently been successful in the computation of Earth satellite trajectories. Because only perturbations from a reference motion are integrated, larger integrating steps than with Cowell's method are possible. Moreover, greater accuracy is possible when working with a fixed-word length because truncation errors effect only the deviations from the reference trajectory. (These errors are generally lost when adding the two words representing the nominal and perturbed components of the state.) The advantages of Encke's method have been particularly marked on lunar flights where the deviations from two-body motions are small.

The variation of parameters or variation-of-elements method differs from the Encke method in that there is a continuous set of elements for the reference orbit. The reference motion of the satellite can be represented by a set of parameters that, in the absence of perturbative forces, would remain constant with time. In the presence of perturbative forces, the reference orbit is being continuously rectified. Some formulations have inherent singularities, and care should be exercised when choosing a set of parameters for the problem at hand. The variation of parameters method is primarily applicable to missions in which small perturbations act throughout the orbit, e.g., microthrust transfer.

2.3 NUMERICAL INTEGRATION

2.3.1 Numerical Integration Discussion

In previous sections, the classical n-body equations of motion and the corresponding equations of motion in a non-conservative force field have been shown to be unyielding to an analytic approach; therefore, in this section, recourse is taken to numerical integration techniques.

The basic problem is to solve the ordinary differential equation

$$\frac{dy}{dx} = f(x,y) \tag{3.1}$$

with the initial condition

$$y(x_o) = y_o \ .$$

later, the solution to higher order differential equations will be discussed. At that time, multiple integrals will be developed; and as a secondary effort, it will be shown that an ordinary differential equation of any order can be reduced to an equivalent system of first-order equations. Thus for simplicity, only one first-order equation will be discussed. Numerical integration seeks to construct a tabular array of the dependent variable (y) in a stepwise process from values of the function evaluated at discrete values (generally, but not necessarily, evenly spaced) of the independent variable, x.

There are several prevalent methods of numerical integration which will be discussed. Accordingly, the section has been divided to display these approaches: (1) quadrature methods, (2) the Runge-Kutta family, (3) predictor-corrector methods, and (4) starting procedures for predictor-corrector methods. In the context of this monograph, quadrature refers to the numerical integration of a definite integral. An equivalent definition is that quadrature is the numerical solution of a first-order ordinary differential equation whose right-hand side is a function of the independent variable only. Thus, equation (3.1) becomes

$$\frac{dy}{dx} = y' = f(x)$$

with

$$y(x_o) = y_o$$

In this case, the solution can be written

$$y(x) = y_o + \int_{x_o}^{x} f(t)\,dt \tag{3.2}$$

and as noted, the problem reduces to the evaluation of the integral. The general result, as given by quadrature (Reference 2.10) can be written

$$\int_{x_0}^{X} f(t)\,dt \simeq a_0 f(x_0) + a_1 f(x_1) + \ldots + a_n f(x_n)$$

where

$$x_0 < x_1 < \ldots < x_n \leq X$$

If $x_n = X$, the formula is referred to as "closed", and if $x_n < X$, the formula is said to be "open." Although quadrature, per se, is seldom used in connection with the n-body problem, its formulas permeate numerical analysis; and some of the more sophisticated methods draw heavily upon quadrature.

The Runge-Kutta family is a collection of numerical integration methods which replace a Taylor series with derivative values computed from points within the step. This family is well known since several of its members have been used extensively in trajectory work. Currently, however, most of the adherents of Runge-Kutta methods are turning toward multistep predictor-corrector methods because of their greater speed and efficiency. These predictor-corrector methods commonly employ two steps, a "predictor" and a "corrector," although some methods (infrequently used) require that the corrector be used more than once. Predictor-corrector methods take the general form (Reference 2.9):

$$y_{n+1} = a_0 y_0 + a_1 y_1 + \ldots a_n y_n + b_0 y_0' + b_1 y_1' + \ldots b_{n+1} y_{n+1}' \qquad (3.3)$$

where the formula is "open" if $b_{n+1} = 0$ and "closed" otherwise. Open formulas can be used as predictors because they do not involve knowledge of the state at the $n+1$ point. Whereas, closed formulas, which require the value of y'_{n+1} and hence y_{n+1}, must be used as correctors. A more descriptive definition of an "open" integration formula is that it does not depend on the derivatives of the function at the point where the integral is desired. An example of an open integration would be a Taylor series, i.e., if $y' = f(x, y)$ then

$$y_{n+1} = y_n + h f(x_n, y_n) + \ldots$$

A closed integration formula is one that explicity contains the derivative at the point at which the integral is desired. An example of this type of integration is the trapezoidal rule

$$y_{n+1} = \frac{[f(x_n, y_n) + f(x_{n+1}, y_{n+1})]}{2} h$$

It is clear that predictor-corrector methods of the form (3.3) are not self-starting and that a matrix of antecedent values of y and y' are required. Predictor-corrector methods can be illustrated using very simple formulas. The illustrative formulas chosen, because of their extreme simplicity, are self-starting; however, one should be cautioned that they also require a singularly small step size. All currently operational predictor-corrector methods, being more sophisticated and more efficient, require a starting procedure. Euler's Methods, one of the oldest and most straightforward of numerical integration methods, is used as a predictor; and the

trapezoidal rule serves as a corrector. Let ΔX be the distance between successive values of the independent variable, X, and then Euler's formula is

$$y^p_{n+1} = y_n + (f'_n)(\Delta X)$$

which is simply the first term in the Taylor series expansion. The familiar trapezoidal rule is

$$y^c_{n+1} = y_n + \frac{(\Delta X)}{2}(f'_n + f'_{n+1})$$

It is again emphasized that the predictor formula does not depend on the value of y at the $n+1$ point, while the corrector formula does (necessarily) depend only at the point. To illustrate the use of Euler's method, consider the ordinary differential equation

$$y' = x + y$$

with the initial condition

$$y(0) = y_0 = 1$$

A predictor-corrector cycle will be illustrated using a step size $h = .1$.

(1) compute f at $x = 0$

$$f_0 = f(x_0, y_0) = x_0 + y_0 = 0 + 1 = 1.$$

(2) predict y_1

$$y_1^p = y_0 + (h)(f_0) = 1 + (.1)(1) = 1.1$$

(3) compute f at $X = .1$ using predicted

$$f_1 = f(x_1, y_1^p) = x_1 + y_1^p = .1 + 1.1 = 1.2$$

(4) correct y_1

$$y_1^c = y_0 + \frac{h}{2}(f_0 + f_1) = 1 + \frac{.1}{2}(1 + 1.2) = 1.11$$

At this point, the predictor-corrector cycle is complete. The equation chosen can be solved in closed form to give a check on the accuracy of the example. The solution of the given equation and initial condition is

$$y(x) = 2e^x - 1 - x$$

and for $X = .1$

$$e^{.1} = 1.1052$$

$$\therefore y(.1) = 2(1.1052) = 1.00 - .10 = 1.1104$$

The error, ε_p, between the predicted value of y_1 and the exact value of y_1 is

$$\varepsilon_p = y_1^p - y(.1) = 1.1 - 1.11052 = -.01052$$

and the error, ε_c, between the corrected and exact values of y_1 is

$$\varepsilon_c = y_1^c - y(.1) = 1.11 - 1.11052 = -.00052$$

Thus, in this example, the corrector cycle has considerably improved the accuracy of the solution.

2.3.2 QUADRATURE

2.3.2.1 Introduction

Numerical integration is the procedure of computing the value of a definite integral from a set of numerical values of the integrand. The process is called quadrature when it is applied to the integration of a function of a single variable. This method is useful since position and velocity are functions of time (i.e., $x = \phi(t)$, $x' = x(t)$) so that an acceleration written as $x'' = f(x, \dot{x}, t)$ is really a function of a single variable, time (i.e., $x'' = F(t)$). However, for trajectory problems, $F(t)$ is not known analytically; and quadrature formulas must be used in a corrector cycle using a tabulated $F(t)$ developed by a predictor formula. In general, quadrature formulas are developed by representing the integrand by an interpolation formula and then integrating this formula between the specified limits. Thus, to find the value of the integral, $\int_a^b f(x)\,dx$, the function $f(x)$ is replaced by an interpolation formula, often one involving differences, and integrated between the limits a and b. In this manner, quadrature formulas can be developed for the approximate integration of any function for which numerical values are known.

The quadrature formulas are presented here in two categories: those having even spacing, the Newton-Cotes formulas; and those having uneven spacing, Gaussian quadrature formulas.

2.3.2.2 Newton-Cotes Formulas

2.3.2.2.1 Discussion. Newton-Cotes formulas use equidistant spacing and comprise the well-known formulas of Euler, Simpson, Lobatto, and others (see Reference 2.10). The formulas are of two types: (1) open; and (2) closed. If the integral is represented as

$$I = \int_a^b f(x)\,dx$$

then the formula obtained is closed if it depends on $f(b)$ and open if it does not.

2.3.2.2.2 Closed Newton-Cotes Formulas. The problem of determining the area under a given curve is frequently referred to as "mechanical quadrature"; and the oldest method of approximating this area is that of inscribed polygons known as the trapezoidal rule. In this method, the ordinates are connected by straight lines; and the area under the curve is approximated by the area under polygons. In a hand calculation, the ordinate of the curve would not

necessarily be used; rather the functional value would be chosen so as to minimize the error in the approximating polygon. However, procedures based upon the use of the ordinates can be easily mechanized, whereas minimizing the error cannot.

If the ordinates $y = y_0, y_1, \ldots, y_n$ correspond to the abscissa values $x = x_0, x_1, \ldots, x_n$, the elementary form for the area of the trapezoid gives the following result for the area under the polygons:

$$A \approx \tfrac{1}{2}(y_0+y_1)(x_1+x_0) + \ldots \tfrac{1}{2}(y_{n-1}+y_n)(x_n-x_{n-1})$$
$$= \tfrac{1}{2}[y_0(x_1-x_0)+y_1(x_2-x_0)+y_2(x_3-x_1)$$
$$+ \ldots + y_{n-1}(x_n-x_{n-2})+y_n(x_n-x_{n-1})] .$$

For equidistant spacing

$$x_1 - x_0 = x_2 - x_1 = \ldots = x_n - x_{n-1} \equiv h$$

the formula becomes

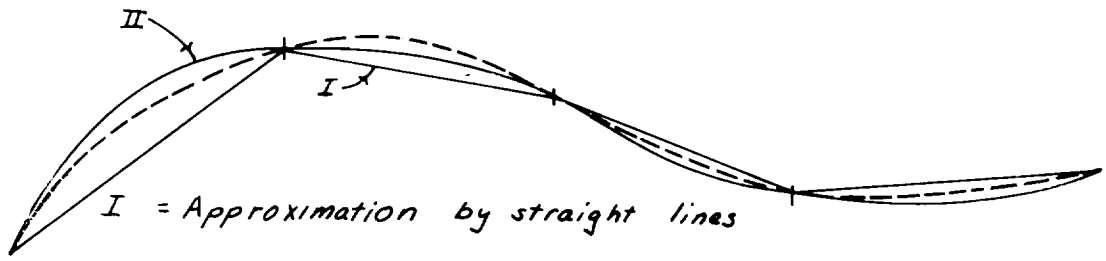
$$A \approx h(\tfrac{1}{2}y_0 + y_1 + y_2 + \ldots + y_{n-1} + \tfrac{1}{2}y_n) ,$$

Now, from the fundamental theorem of integral calculus, the area under the curve $y = f(x)$ can be defined as the limit (if it exists) to which A tends as approaches zero. Therefore,

$$A = \int_a^b f(x)dx = \lim_{h \to 0} h(\tfrac{1}{2}y_0 + y_1 + \ldots + y_{n-1} + \tfrac{1}{2}y_n)$$

For theoretical purposes, this limit process is quite satisfactory; and from the viewpoint of practical computation, the trapezoidal rule given quite good results if a significant number or ordinates is utilized. However, straight lines can be too rigid for a really satisfactory approximation of curves. If a curve is to be approximated by a succession of straight lines, a great many small lines are needed for this purpose. Hence, it is generally more efficient to approximate the curve with higher-order polynomials.

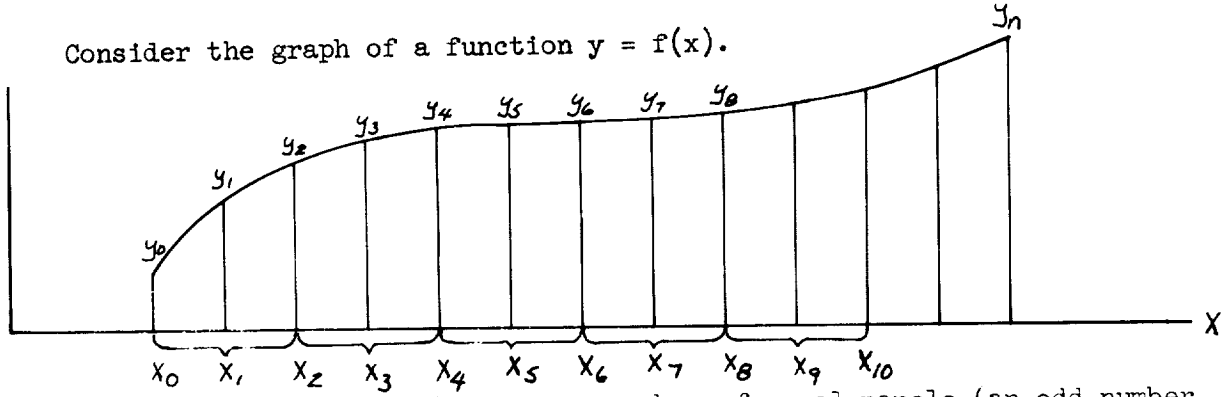
Polynomials of the second-order are much better for this approximation; and with the help of such parabolas, a curve with a higher degree of accuracy can be generated. A large number of very short straight line segments is often replaced by a smaller number of parabolic sections connecting three consecutive ordinates. This method yields a much more accuracy solution because it produces a patched curve which approximates the actual curve much better than the straight line approximation.



I = Approximation by straight lines

II = Approximation by parabolic arcs

Consider the graph of a function y = f(x).



The total area is divided into an even number of equal panels (an odd number of ordinates); and for the sake of convenience, the width (h) of each panel is considered to be 1. In each double panel, (i.e., two intervals), the curve is approximated by a polynomial of second-order.

Letting the first double panel be composed of the ordinates, y_0, y_1, y_2, $y = f(x)$ can be expanded around the point $x = 1$ into a local power series, making use of the method of differences. Stirling's formula as a power series (see Reference 2.11) then yields:

$$f(1+t) = f(1) + \delta f(1)t + \delta^2 f(1) \frac{t^2}{2}$$

where

$$f(1) = y_1$$
$$\delta f(1) = 1/2 (y_1 - y_0)$$
$$\delta^2 f(1) = y_2 - 2y_1 + y_0 \ .$$

Here, the curve between $x = 0$ and $x = 2$ is approximated by a parabola which coincides with the actual curve at the three points of interpolation $x = x_0$, x_1, x_2 (in this particular case, let $x_0 = 0$, $x = 1$, $x_2 = 2$). The area under the approximating parabola can be obtained by integrating between the points $x = 0$ and $x = 2$.

$$A_{02} \approx \int_{-1}^{1} f(1+t)\,dt$$
$$= f(1)t + \delta f(1) t^2/2 + \delta^2 f \frac{(1)}{2} t^3/3 \Big|_{-1}^{+1}$$
$$= 2f(1) + 1/3\ \delta^2 f(1)$$
$$= 2y_1 + 1/3 (y_2 - 2y_1 + y_0)$$
$$= 1/3\, y_0 + 4/3\, y_1 + 1/3\, y_2 \ .$$

Now generalizing for non unit intervals between x_0, x_1, and x_2

$$A \approx h/3 (y_0 + 4y_1 + y_2)$$

which is familiarly known as Simpson's rule. The process can be repeated for the remaining panels A_{24}, A_{46}, ... until the total area is exhausted. Summing these areas, leads to the parabolic rule where

$$A = h/3 \left[y_0 + 4y_1 + 2y_2 + 4y_3 + 2y_4 + \ldots + 4y_{n-1} + y_n \right] \ .$$

In this formula, the even and odd ordinates are separated and applied with

different weights instead of the same weights as in the trapezoidal method. This discrimination of weights greatly increases the accuracy of the result.

The necessity of an even number of panels is sometimes an inconvenient limitation for the use of Simpson's rule; however, other formulas can be developed which consider integrating over any number of panels. In order to obtain such formulas for the approximate evaluation of an integral of the form $\int_a^b f(x)\,dx$, it is necessary to introduce the change of variables $x = a + ((b-a)/n)s$ where n is an integer, to obtain the relation

where $$\int_a^b f(x)\,dx = \frac{b-a}{n}\int_0^n F(s)\,ds$$

$$F(s) = f\left(a + \frac{b-a}{n}s\right)$$

Assuming that $f(x)$ can be approximated over (a,b) (the open interval (a,b) is defined to be the set of x such that $a < x < b$) by a polynomial which agrees with it at $n+1$ equally spaced points in (a,b) (yielding n panels), the approximating formula

$$\int_0^n F(s)\,ds \approx \sum_{K=0}^{n} C_K^{(n)} F(K)$$

may be obtained where

$$C_K^{(n)} = \int_0^n \frac{s(s-1)\cdots(s-k+1)(s-k-1)\cdots(s-n)}{k(k-1)\cdots(k-k+1)(k-k-1)\cdots(k-n)}\,ds \qquad (3.4)$$

Letting $h=(b-a)/n$ and $x_j=a+jh$, the result established can be put in the more explicit form

$$\int_{x_0}^{x_n} f(x)\,dx \approx h\sum_{K=0}^{n} C_K^{(n)} f(x_K)$$

This error involved in the process can now be expressed as

$$E_n = \frac{h^{n+2} f^{(n+1)}}{(n+1)!}(E)\int_0^n s(s-1)\cdots(s-n)\,ds \qquad (n \text{ odd})$$

$$E_n = h^{n+3} f^{(n+2)}(E)\int_0^n \left(s-\frac{n}{2}\right)s(s-1)\cdots(s-n)\,ds \qquad (n \text{ even})$$

Hence, the following formulas known as the closed Newton-Cotes formulas may be obtained where $f^{(n)}(\xi)$ is the error term, $x_0 \le \xi \le x_n$. They are classified as closed formulas since they are characterized by the fact that the integral extends from the first to the last (n) ordinate used in the formula but never beyond.

Simpson's rule:

$$\int_{x_0}^{x_1} f(x)\,dx = \frac{h}{2}(f_0 + f_1) - \frac{h^3}{12}f''(\xi)$$

$$\int_{x_0}^{x_2} f(x)\,dx = \frac{h}{3}(f_0 + 4f_1 + f_2) - \frac{h^5}{90}f^{IV}(\xi)$$

$$\int_{x_0}^{x_3} f(x)\,dx = 3h/8\,(f_0 + 3f_1 + 3f_2 + f_3) - 3h^5/80\, f^{IV}(\xi)$$

$$\int_{x_0}^{x_4} f(x)\,dx = 2h/45\,(7f_0 + 32f_1 + 12f_2 + 32f_3 + 7f_4) - 8h^7/945\, f^{VI}(\xi)$$

$$\int_{x_0}^{x_5} f(x)\,dx = 5h/288\,(19f_0 + 75f_1 + 50f_2 + 50f_3 + 75f_4 + 19f_5) - 275h^7/12096\, f^{VI}(\xi)$$

$$\int_{x_0}^{x_6} f(x)\,dx = \frac{1}{140}(41f_0 + 216f_1 + 27f_2 + 272f_3 + 27f_4 + 216f_5 + 41f_6) - \frac{9h^9}{1400} f^{VIII}(\xi)$$

$$\int_{x_0}^{x_7} f(x)\,dx = \frac{7h}{17280}(751f_0 + 3577f_1 + 1323f_2 + 2989f_3 + 2989f_4 + 1323f_5 + 3577f_6 + 751f_7) - \frac{8183}{518400} f^{VIII}(\xi)$$

$$\int_{x_0}^{x_8} f(x)\,dx = \frac{4h}{14175}(989f_0 + 5888f_1 - 928f_2 + 10496f_3 - 4540f_4 + 10496f_5 - 928f_6 + 5888f_7 + 989f_8) - \frac{2368h^{9}}{467775} f^{X}(\xi)$$

Since all the integral formulas of this type must be exact if $f(x)$ is a constant, the sum of the weight coefficients must equal the length of the interval, i.e., $h\sum_{i=1}^{n} c_i = b - a$. Also, an inspection of the error terms reveals that a formula involving an odd number, $n+1 = 2m+1$, of points would yield exact results if $f(x)$ were a polynomial of degrees $n+1$ or less, whereas one involving an even number $n+1 = 2m$, of points would be exact only if $f(x)$ were a polynomial of degree n or less. Thus, the two formulas involving $2m$ and $2m-1$ ordinates have the same order of accuracy so that, generally, no great advantage is gained by advancing from a formula involving an odd number of ordinates to one involving one more ordinate.

2.3.2.2.3 Open Newton-Cotes Formulas. Often, in trying to obtain solutions of differential equations, it is necessary to be able to integrate ahead, beyond the last value for which the function is known. This may be done by employing an open-type quadrature formula where the integration extends beyond the ordinates employed in the system.

The first few such formulas which do not involve the ordinates at the end of the interval may be expressed as follows:

$$\int_{x_0}^{x_3} f(x)\,dx = 3h/2\,(f_1 + f_2) + 3h^3/4\, f''(\xi)$$

$$\int_{x_0}^{x_4} f(x)\,dx = 4h/3\,(2f_1 - f_2 + 2f_3) + 14h^5/45\, f^{IV}(\xi)$$

$$\int_{x_0}^{x_5} f(x)\,dx = 5h/24\,(11f_1 + f_2 + f_3 + 11f_4) + 95h^5/144\, f^{IV}(\xi)$$

$$\int_{x_0}^{x_6} f(x)\,dx = 3h/10\,(11f_1 - 14f_2 + 26f_3 - 14f_4 + 11f_5) + 41h^7/140\, f^{VI}(\xi)$$

where in this case, the weighting function (n)

$$C_k^{(n)} = \int_0^n \frac{(s-1)\cdots(s-k+1)(s-k-1)\cdots(s-n-1)}{(k-1)\cdots(k-k+1)(k-k-1)\cdots(k-n-1)}\,ds \quad . \tag{3.5}$$

Notice that although these open formulas do not include f_o and f_n, the limits of integration extend from x_o to x_n. Also, equation (3.5) is the same as equation (3.4) except for the elimination of the terms s, $s-n$, k, and $k-n$.

2.3.2.2.4 Difference Table Method. When the data are tabulated for uniformly spaced abscissas with spacing, h, it is conventional to define the differences of the tabular data and employ quadrature formulas in terms of these differences. In these cases, the forward differences $\Delta f(x_o)$ are defined as

$$\Delta f(x_o) = f(x_o + h) - f(x_o)$$

and

$$= f_1 - f_o$$

$$\Delta f_{r-1} = f_r - f_{r-1} \quad . \tag{3.6}$$

Similarly, the differences of these first differences, called second differences, are denoted by $\Delta^2 f_o$ or $\Delta^2 f(x_o)$ and defined as

$$\Delta^2 f_o = \Delta f_1 - \Delta f_o = f_2 - 2f_1 + f_o \quad .$$

The third and higher order are obtained by continuing in the same manner.

The following difference table shows how these values are arranged for convenient usage. Notice that the subscript remains constant along each forward diagonal.

x	f	Δf	$\Delta^2 f$	$\Delta^3 f$	$\Delta^4 f$	$\Delta^5 f$	$\Delta^6 f$	$\Delta^7 f$	$\Delta^8 f$
x_0	f_0								
		Δf_0							
x_1	f_1		$\Delta^2 f_0$						
		Δf_1		$\Delta^3 f_0$					
x_2	f_2		$\Delta^2 f_1$		$\Delta^4 f_0$				
		Δf_2		$\Delta^3 f_1$		$\Delta^5 f_0$			
x_3	f_3		$\Delta^2 f_2$		$\Delta^4 f_1$		$\Delta^6 f_0$		
		Δf_3		$\Delta^3 f_2$		$\Delta^5 f_1$		$\Delta^7 f_0$	
x_4	f_4		$\Delta^2 f_3$		$\Delta^4 f_2$		$\Delta^6 f_1$		$\Delta^8 f_0$
		Δf_4		$\Delta^3 f_3$		$\Delta^5 f_2$		Δf_1	
x_5	f_5		$\Delta^2 f_4$		$\Delta^4 f_3$		$\Delta^6 f_2$		
		Δf_5		$\Delta^3 f_4$		$\Delta^5 f_3$			
x_6	f_6		$\Delta^2 f_5$		$\Delta^4 f_4$				
		Δf_6		$\Delta^3 f_5$					
x_7	f_7		$\Delta^2 f_6$						
		Δf_7							
x_8	f_8								

For calculation near the end of a tabulated range, the notation ∇ for backwards differences if often more convenient. In this case,

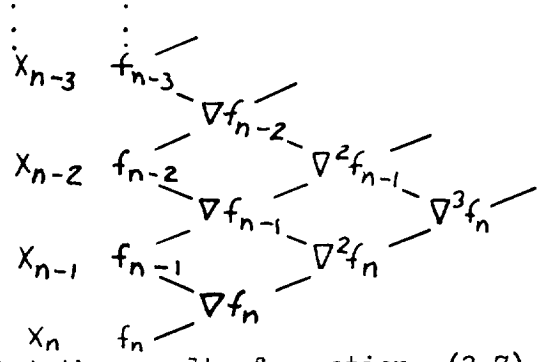
$$\nabla f(x) = f(x) - f(x-h)$$

$$\nabla f_s = f_s - f_{s-1} \tag{3.7}$$

and

$$\nabla^{k+1} f(x) = \nabla^k f(x) - \nabla^k f(x-h)$$

where the subscripts of the difference tables are changed to read



so that the result of equation (3.7) is the same as in the previous equation (3.6) but denoted by a different subscript, and the subscript remains constant along each backward diagonal.

For remaining calculations, the notation of central differences is usually most convenient. If the calculation is to be done near a certain interior tabular point, it is convenient to number that abscissa as x_0. In central difference notation

$$\delta f(x) = f(x + \tfrac{1}{2}h) - f(x - \tfrac{1}{2}h)$$

and

$$\delta^{r+1} f(x) = \delta^r f(x + \tfrac{1}{2}h) - \delta^r f(x - \tfrac{1}{2}h) \quad .$$

It is seen that $\delta f(x)$ does not necessarily involve tabulated ordinates. However, the second difference

$$\delta^2 f(x_k) = \delta f(x_k + \tfrac{1}{2}h) - \delta f(x_k - \tfrac{1}{2}h)$$

$$= f_{k+1} - 2f_k + f_{k-1}$$

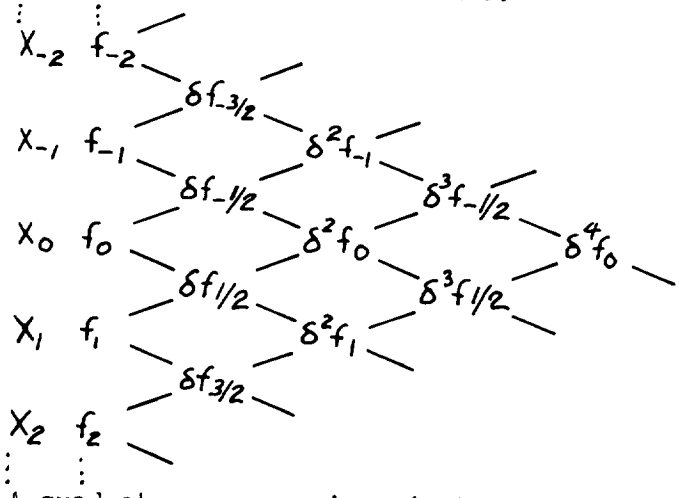
does involve tabular entries as do all central differences of even order. Thus, it may be written that

$$\delta f_{1/2} = f_1 - f_0$$

$$\delta f_1 = \delta f_{3/2} - \delta f_{1/2} \quad .$$

The central difference table around x_0 (the zero subscript in this case indicates the center of the interval) illustrates that the subscript remains

constant along the horizontal lines.



A quadrature expression, in terms of differences, may now be obtained by integrating Stirling's formula (Reference 2.12) from $x = x_0 - h$ to $x = x_0 + h$ ($u = -1$ to $u = 1$) yielding

$$I = \int_{-h}^{h} \phi(x)\,dx = h\int_{-1}^{1} \Big(y_0 + u\,\frac{\Delta y_{-1} + \Delta y_0}{2} + \frac{u^2}{2}\Delta^2 y_{-1} + \frac{u(u^2-1)\,\Delta^3 y_{-2} + \Delta^3 y_{-1}}{2} + \frac{u^2(u^2-1)}{4!}\Delta^4 y_{-2}$$

$$+ \frac{u(u^2-1)(u^2-4)\,\Delta^5 y_{-3}}{5!} + \frac{\Delta^5 y_{-2}}{2} + \frac{u^2(u^2-1)(u^2-4)}{6!}\Delta^6 y_{-3} + \dots \Big)\,du$$

$$= h\Big[2y_0 + \tfrac{1}{3}\Delta^2 y_{-1} + \tfrac{2}{24}\Big(\tfrac{1}{5} - \tfrac{1}{3}\Big)\Delta^4 y_{-2} + \frac{2}{720}\Big(\frac{1}{7} - 1 + \frac{4}{3}\Big)\Delta^6 y_{-3}\Big]$$

$$= 2h\Big[y_0 + \frac{1}{6}\Delta^2 y_{-1} - \frac{1}{180}\Delta^4 y_{-2} + \frac{1}{1512}\Delta^6 y_{-3} + \dots\Big]\,.$$

By advancing the subscripts of the y's by one unit, the value of the integral I_0^2 from $x = x_0$ to $x = x_0 + 2h$ is obtained. The integrals $I_2^4, I_4^6, \dots, I_{n-2}^n$ are likewise seen to be

$$I_0^2 = 2h\Big[y_1 + \frac{1}{6}\Delta^2 y_0 - \frac{1}{180}\Delta^4 y_{-1} + \frac{1}{1512}\Delta^6 y_{-2}\Big]\,.$$

$$I_2^4 = 2h\Big[y_3 + \frac{1}{6}\Delta^2 y_2 - \frac{1}{180}\Delta^4 y_1 + \frac{1}{1512}\Delta^6 y_0\Big]\,,$$

$$I_4^6 = 2h\Big[y_5 + \frac{1}{6}\Delta^2 y_4 - \frac{1}{180}\Delta^4 y_3 + \frac{1}{1512}\Delta^6 y_2\Big]\,,$$

. . . . . .

$$I_{n-2}^n = 2h\Big[y_{n-1} + \frac{1}{6}\Delta^2 y_{n-2} - \frac{1}{180}\Delta^4 y_{n-3} + \frac{1}{1512}\Delta^6 y_{n-4}\Big]\,.$$

Adding all these separate integrals yields

$$I_0^n = 2h\Big[y_1 + y_3 + y_5 + \cdots + y_{n-1} + \frac{1}{6}(\Delta^2 y_0 + \Delta^2 y_2 + \cdots + \Delta^2 y_{n-2}) \qquad (3.8)$$

$$- \frac{1}{180}(\Delta^4 y_{-1} + \Delta^4 y_1 + \Delta^4 y_3 + \cdots + \Delta^4 y_{n-3}) + \frac{1}{1512}(\Delta^6 y_{-2} + \Delta^6 y_0 + \Delta^6 y_2 + \cdots + \Delta^6 y_{n-4})\Big],$$

where

$$I_0^n = \int_{x_0}^{x_0 + nh} f(x)\,dx$$

and n is even.

By neglecting fourth and fifth differences and replacing the second differences by their values in terms of the y's, it is found that equation (3.8) reduces to Simpson's rule. Thus, it may be seen that the formula represents Simpson's rule with correctional

2.3.2.2.5 Numerical Example. As an example, using difference tables, the approximate value of the integral

$$I = \int_{1}^{2} \frac{dx}{x}$$

is computed. This example will serve to illustrate the integration technique; however, the sample problem is not similar to the trajectory problem since the derivative for this sample problem is a known function and does not suffer from cumulative round-off error resulting from the previous integration steps. Taking $h = .01$ a difference table of $y = 1/x$ is set up at one-tenth unit intervals from $x = 0.7$ to $x = 2.3$ (note that the differences are referenced to the last place in the function). Substituting into equation (3.8) the appropriate differences from the table yields

$$I = 0.1\left[6.93771403 - 1/2(7594744) + 11/720(593339)\right.$$

$$\left. - 191/60480(136810)\right]$$

The correct value is $\ln 2 = 0.693147181$

Table 3.1. DIFFERENCE TABLE FOR y = 1/x

	X	y	Δy	$\Delta^2 y$	$\Delta^3 y$	$\Delta^4 y$	$\Delta^5 y$	$\Delta^6 y$
	0.7	1.42857143						
			-17857143					
	0.8	1.25000000		3968254				
			-13888889		-1190476			
	0.9	1.11111111		2777778		432900		
			-11111111		- 757576		-180375	
	1.0	1.00000000		2020202		252525		83252
			-9090909		- 505051		- 97123	
	1.1	0.9090909 1*		1515151*		155402*		41618*
Y_0 =			-7575758		- 349649		-55505	
	1.2	0.83333333		1165502		99897		22212
			-6410256		- 249752		-33293	
	1.3	0.76923077*		915750*		66604*		12472*
			-5494506		-183148		-20821	
	1.4	0.71428571		732602		45783		7362
			-4761904		-137365		-13459	
	1.5	0.66666667*		595237*		32324*		4478*
			- 4166667		-105041		- 8981	
	1.6	0.62500000		490196		23343		2834
			-3676471		- 81698		- 6147	
	1.7	0.58823529*		408498*		17196*		1855*
			-3267973		- 64502		- 4292	
	1.8	0.55555556		343996		12904		1215
			-2923977		- 51598		- 3077	
	1.9	0.52631579*		292398*		9827*		843*
Y_N =			- 2631579		-41771		- 2234	
	2.0	0.50000000		250627		7593		589
			-2380952		-34178		- 1645	
	2.1	0.47619048		216449		5948		
			-2164503		-28230			
	2.2	0.45454545		188219				
			-1976284					
	2.3	0.43478261						

* Denotes values used in example

2.3.2.3 Gaussian Quadrature

2.3.2.3.1 Discussion. Until the time of Gauss, quadrature formulas assumed that the chosen points $X = X_n$ were equidistantly spaced by h units. Gauss conceived the idea that much greater accuracy could be gained by relaxing the constraint of equidistant spacing, and instead, requiring only that the points be symmetrically spaced with respect to the midpoint of the interval of integration. The formulas so obtained are termed Gaussian quadrature formulas.

2.3.2.3.2 Gaussian Quadrature Formulas. To derive the Gaussian quadrature formulas, let $I = \int_a^b f(x)dx$ denote the integral to be computed and change the variable by substituting

$$x = (b-a)u + \frac{a+b}{2}$$

so that the limits to integration become $-\frac{1}{2}$ and $\frac{1}{2}$. The function $f(x)$ is now

$$f(x) = f\left[(b-a)u + \frac{a+b}{2}\right] \equiv \phi(u)$$

Then, since $dx = (b-a)du$, the integral becomes $I = (b-a)\int_{-\frac{1}{2}}^{1/2} \phi(u)\,du$.

Gauss proposed that the integral be written

$$\int_{-1/2}^{1/2} \phi(u)du = R_1\phi(u_1) + R_2\phi(u_2) + \ldots + R_n\phi(u_n) \,, \qquad (3.9)$$

where $u_1, u_2, \ldots, u_n$ are the points of subdivision of the interval $u = -1/2$ to $u = 1/2$. The corresponding values of x are, therefore

$$x_1 = (b-a)u_1 + (a+b)/2$$

$$x_2 = (b-a)u_2 + (a+b)/2$$

and the value of the integral $\int_a^b f(x)\,dx$ is

$$I = \int_a^b f(x)dx = (b-a)\left[R_1\phi(u_1) + R_2\phi(u_2) + \ldots + R_n\phi(u_n)\right] \qquad (3.10)$$

The values $u_1, u_2, \ldots, u_n$ and $R_1, R_2, \ldots, R_n$ can be found by assuming that $\phi(u)$ can be expanding in a convergent power series in the interval $u = -1/2$ to $u = 1/2$. Hence,

$$\phi(u) = a_0 + a_1u + a_2u^2 + \ldots + a_mu^m + \ldots \qquad (3.11)$$

Integrating equation (3.11) between the limits $-\frac{1}{2}$ and $\frac{1}{2}$ yields

$$I = \int_{-1/2}^{1/2} \phi(u)du = \int_{-1/2}^{1/2} (a_0 + a_1u + \ldots + a_mu^m + \ldots)\,du \qquad (3.12)$$

$$= a_0 + 1/12\,a_2 + 1/80\,a_4 + 1/448\,a_6 + 1/2304\,a_8 + \ldots \,.$$

Note that the odd terms disappear when the integral is evaluated at both the upper and lower limits. Also, from equation (3.11)

$$\phi(u_1) = a_0 + a_1 u_1 + a_2 u_1^2 + a_3 u_1^3 + a_4 u_1^4 + \cdots + a_m u_1^m + \cdots$$

$$\phi(u_2) = a_0 + a_1 u_2 + a_2 u_2^2 + a_3 u_2^3 + a_4 u_2^4 + \cdots + a_m u_2^m + \cdots$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\phi(u_n) = a_0 + a_1 u_n + a_2 u_n^2 + a_3 u_n^3 + \cdots + a_m u_n^m + \cdots .$$

Substituting these values of $\phi(u_1), \phi(u_2), \ldots, \phi(u_n)$ and into equation (3.9) rearranging them, J is obtained as

$$\begin{aligned} J = {} & a_0 (R_1 + R_2 + R_3 + \cdots + R_n) \\ & + a_1 (R_1 u_1 + R_2 u_2 + \cdots + R_n u_n) \\ & + a_2 (R_1 u_1^2 + R_2 u_2^2 + \cdots + R_n u_n^2) \\ & \cdots\cdots\cdots\cdots\cdots \\ & + a_m (R_1 u_1^m + R_2 u_2^m + \cdots + R_n u_n^m) \end{aligned} \qquad (3.13)$$

Now if the integral J in equation (3.13) is to be identical with equation (3.12), <u>regardless of the form of $\phi(u)$</u>, then corresponding coefficients of $a_0, a_1, a_2, \ldots$ in equations (3.13) and (3.13) must be equal. This requirement yields

$$\begin{aligned} & R_1 + R_2 + R_3 + \cdots + R_n = 1 , \\ & R_1 u_1 + R_2 u_2 + R_3 u_3 + \cdots + R_n u_n = 0 , \\ & R_1 u_1^2 + R_2 u_2^2 + R_3 u_3^2 + \cdots + R_n u_n^2 = \frac{1}{12} , \\ & R_1 u_1^3 + R_2 u_2^3 + R_3 u_3^3 + \cdots + R_n u_n^3 = 0 , \\ & R_1 u_1^4 + R_2 u_2^4 + R_3 u_3^4 + \cdots + R_n u_n^4 = \frac{1}{80} , \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ & \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned} \qquad (3.14)$$

Solving these equations using algebraic means is tedious. However, it can be shown (Reference 2.12) that if $\phi(u)$ is a polynomial of degree not higher than $2n-1$, then $u_1, u_2, \ldots, u_n$ are the zeros of the Legendre polynomial $P_n(u)$ or the roots of $P_n(u) = 0$. These roots are found from the equation

$$\frac{d^n}{du^n}\left[u^2 - (1/2)^2\right]^n = 0 \qquad (3.15)$$

The n roots $u_1, u_2, \ldots, u_n$ of this equation are all real. Thus, substitution into equation (3.14) defines the R's.

In solving equation (3.15) for any n, it is to be noted that the y's are symmetrically placed with respect to the midpoint of the interval of integration and that the R's are equal for each symmetric pair of u's.

The numerical values of the u's and corresponding R's for $n=2$ to $n=10$ are given in the following table (Reference 2.10) where the notation for the form $u_{\pm K} = N$ means $u_K = N$ and $u_{-K} = -N$ and u_0 designates the midpoint.

Hence, the value of the integral may be determined by substituting these values into equation (3.9).

	u	R
$n = 2$	$u_{\pm 1} = 0.2886751346$	$R = \frac{1}{2}$
$n = 3$	$u_0 = 0$	$R = 4/9$
	$u_{\pm 1} = 0.3872983346$	$R = 5/18$
$n = 4$	$u_{\pm 1} = 0.1699905218$	$R = 0.3260725774$
	$u_{\pm 2} = 0.4305681558$	$R = 0.1739274226$
$n = 5$	$u_0 = 0$	$R = 64/225$
	$u_{\pm 1} = 0.2692346551$	$R = 0.2393143352$
	$u_{\pm 2} = 0.4530899230$	$R = 0.1184634425$
$n = 6$	$u_{\pm 1} = 0.1193095930$	$R = 0.2339569673$
	$u_{\pm 2} = 0.3306046932$	$R = 0.1803807865$
	$u_{\pm 3} = 0.4662347571$	$R = 0.0856622461 9$
$n = 7$	$u_0 = 0$	$R = 0.2089795918$
	$u_{\pm 1} = 0.2029225757$	$R = 0.1909150253$
	$u_{\pm 2} = 0.3707655928$	$R = 0.1398526957$
	$u_{\pm 3} = 0.4745539562$	$R = 0.0647424830 8$
$n = 8$	$u_{\pm 1} = 0.0917173212$	$R = 0.1813418917$
	$u_{\pm 2} = 0.2627662050$	$R = 0.1568533229$
	$u_{\pm 3} = 0.3983332387$	$R = 0.1111905172$
	$u_{\pm 4} = 0.4801449282$	$R = 0.0506142681 5$
$n = 9$	$u_0 = 0$	$R = 0.1651196775$
	$u_{\pm 1} = 0.1621267117$	$R = 0.1561735385$
	$u_{\pm 2} = 0.3066857164$	$R = 0.1303053482$
	$u_{\pm 3} = 0.4180155537$	$R = 0.0903240803 5$
	$u_{\pm 4} = 0.4840801198$	$R = 0.0406371941 8$
$n = 10$	$u_{\pm 1} = 0.0744371695$	$R = 0.1477621124$
	$u_{\pm 2} = 0.2166976971$	$R = 0.1346333597$
	$u_{\pm 3} = 0.3397047841$	$R = 0.1095431813$
	$u_{\pm 4} = 0.4325316833$	$R = 0.0747256745 8$
	$u_{\pm 5} = 0.4869532643$	$R = 0.0333356721 5$

2.3.2.3.3 Computational Procedure and Example. Because Gaussian quadrature formulas do not involve equidistant spacing of the independent variable, a difference table set-up cannot be used. However, the application of the formulas in the ordinate form is straightforward and can be illustrated by example.

Compute the interval

$$I = \int_5^{12} \frac{dx}{x}$$

Solution: Define the function u as

$$x = (b-a)u + \frac{a+b}{2} = 7u + 8.5$$

$$y = \frac{1}{x} = \frac{1}{7u+8.5} = \phi(u) .$$

Now assuming $n = 5$

$$y_0 = \phi(u_0) = \frac{1}{8.5} = 0.117647059$$

$$y_1 = \phi(u_1) = \frac{1}{7u_1+8.5} = \frac{1}{10.3846426} = 0.096296043 9$$

$$y_{-1} = \phi(u_{-1}) = \frac{1}{7u_{-1}+8.5} = \frac{1}{6.61535741} = 0.151163412$$

$$y_2 = \phi(u_2) = \frac{1}{7u_2+8.5} = \frac{1}{11.67162946} = 0.085677839 9$$

$$y_3 = \phi(u_{-2}) = \frac{1}{7u_{-2}+8.5} = \frac{1}{5.32837054} = 0.187674636$$

Substituting these values into (3.10) with the corresponding R's for $n = 5$ yields

$$I = 7\left[\frac{64}{225} \times 0.117647059 + 0.239314335 2 (0.151163412 + 0.096296043 9) + 0.118463442 5 (0.187674636 + 0.085677839 9)\right] = 0.875468458 .$$

In contrast, the true value of the integral is

$$I = \int_5^{12} \frac{dx}{x} = \ln \frac{12}{5} = \ln 2.4 = 0.875468737$$

2.3.2.3.4 Remarks. Hildebrand (Reference 2.12) offers a discussion of other quadrature methods with the weighting function

$$w(x) = e^{-x}, e^{-x^2}, \frac{1}{1-x^2}, (1-x)^{\alpha} (1+x)^{\beta}$$

Generally, it should be noted that Gauss's formula gives an exact result when $f(x)$ is a polynomial of the $(2n-1)$ degree or lower. However, though, Gauss's method is of great accuracy and theoretically sound, it has the disadvantage of being laborous in its application since the limits of the integral must be transformed $-\frac{1}{2}$ and $\frac{1}{2}$. Also, if the values of y are to be computed from a formula, the numerical value of u to be substituted

must have as many significant figures as is desired in the y's. After determining the y's to the desired number of significant figures, they must be multiplied by the R's having at least as many significant figures.

This requirement compels use of a large number of significant figures in every step to achieve the accuracy that the formula is capable of providing; however, Gaussian quadrature is definitely applicable when the determination of ordinates needed for the conventional formulas would involve either direct calculation, physical measurement, or interpolation, since these values are unlikely to be equidistantly spaced.

2.3.3 RUNGE-KUTTA FAMILY

2.3.3.1 Introduction

The name Runge-Kutta is given to a family of numerical integration techniques which approximate a Taylor series extrapolation of a function by several evaluations of the first derivative at points within the interval of extrapolation. The family has many members some of which have names of specific individuals (e.g. Huen) and some of which are designated by an order number. The order of a particular Runge-Kutta solution is the order of the highest power of the step size retained in the equivalent Taylor expansion. A solution for $y(x_n + h)$ is obtained from a knowledge of the current solution point, $y(x_n)$, and the first derivative of y at points in the interval $x_n \leq x \leq x_n + h$. Values of y at previous points are not necessary for continuing the solution. The method is, therefore, a single step method (i.e. a single step method requires only the current solution point as opposed to multi-step methods which require that previous solution points be available before the process of obtaining the solution point ahead can begin) and as such has the advantage of being self starting. An associated disadvantage is that the information lost in discarding previous points must be regenerated (in a different form) at each point along the solution with a resulting waste of time compared to multi-step methods. The Runge-Kutta method follows the solution curve well, is capable of good accuracy, and can easily accommodate any change in step size although a criterion for changing step size is not available as a part of the solution calculation as is the case with some multi-step methods.

2.3.3.2 Third Order Methods

2.3.3.2.1 Discussion. The third order Runge-Kutta equations are characterized by approximating a Taylor series solution for y to terms of order h^4. As will be seen in the next section, the general equations contain two arbitrary parameters. Several people have chosen particular values for these free parameters and their names are generally appended to the Runge-Kutta name when reference is made to a particular formula (c.g., Runge-Kutta-Heun). However, it is not the intent to display all possible third order formulas in this section. Indeed such a feat would be impossible because there exists a two-fold infinity of possible formulas. Rather, only some of the better known or more useful formulas are listed.

2.3.3.2.2 Deviation of General Third Order Runge-Kutta Formula. The derivation presented here will be for the general third-order Runge-Kutta formulation. The differential equation which is to be solved is

$$y'(x) = f(x, y) \tag{3.16}$$

with the initial condition

$$y(x_0) = y_0$$

The deriviation of the equations begins by considering the integral of equation (3.16) from x_o to $(x_o + h)$.

$$\Delta y = y(x_o + h) - y(x_o) = \int_{x_o}^{x_o + h} f(x, y(x))\, dx \tag{3.17}$$

It is known, from the mean value theorem (Reference 2.26), that a solution to equation (3.17) of the form

$$\Delta y = h f(\xi, y(\xi)) = h y'(\xi)$$
$$x_o \leq \xi \leq x_o + h \tag{3.18}$$

exists. The value, ξ , is however not generally known prior to obtaining the solution to the problem. The Runge-Kutta approach to the solution of(3.16) is essentially of the form (3.18): but since is not known, a weighted average of three Δy_i's is calculated at points in the interval $x_o \leq x_i \leq x_o + h$. This weighted average is

$$\Delta y \approx \sum_{i=1}^{N} a_i \Delta y_i \tag{3.19}$$

where

$$\Delta y_i = h y'(x_i)$$

The weights, a_i, will be chosen so that the value Δy calculated from equation (3.19) will agree with a Taylor series expansion of $y(x)$ about the point (x_o, y_o) to terms of order four. A point to be brought out here is that when the proper values of a_i have been chosen Δy will have been determined as a third order Taylor series expansion about (x_o, y_o) even though no derivatives of y higher than the first will have been evaluated. To determine the appropriate values of the a_i's, three Δy_i's must be calculated. The first Δy_i will be evaluated at $x = x_o$.

$$\Delta y_1 = h y'(x_o) = h f(x_o, y_o) \tag{3.20}$$

The second Δy_i will be evaluated at (x_1, y_1) where $x_1 = x_o + mh$, $x_o \leq x_1 \leq x_o + h$.

$$\Delta y_2 = h f(x_1, y_1)$$

At this point it appears that a problem exists because y_1 cannot be found until the solution to the original problem is known. An approximate value for y_1 can, however, be found from a first-order Taylor series about x_o.

$$y_1 = y_0 + mkf(x_0, y_0) \tag{3.21}$$

It will be seen by following the derivation that such a choice for y_1 can be made without relaxing the requirement that equation (3.19) agrees with three terms of the Taylor series for y. Using equation (3.21), the expression for Δy_2 becomes

$$\Delta y_2 = kf(x_0 + mk,\ y_0 + mkf(x_0, y_0))$$

or

$$\Delta y_2 = kf(x_0 + mk, y_0 + m\Delta y) \tag{3.22}$$

The third Δy_i will be calculated at $x = x_2 = x_0 + \lambda h$. The value of y_2 will again be approximated by a linear Taylor series, but in two steps. The derivative of y at (x_0, y_0) will be used to evaluate y (x) at some arbitrary x, say at $x' = x_0 + (\lambda - \rho)h$, between x_0 and $x_0 + \lambda h$. Next the change in y from $x = x'$ to $x = x_0 + \lambda h$ is evaluated using the derivative at the point (x_1, y_1). An examination of Figure 3.1 may aid in understanding the above description. The relation for y_2 is

$$y_2 = y_0 + (\lambda - \rho)\, kf(x_0\ y_0) + \rho\, kf(x_1, y_1) \tag{3.23}$$

and Δy_3 is

$$\Delta y_3 = kf(x_0 + \lambda k,\ y_0 + (\lambda - \rho)\, kf(x_0, y_0) + \rho\, kf(x_1, y_1))$$

or

$$\Delta y_3 = kf(x_0 + \lambda k,\ y_0 + (\lambda - \rho)\Delta y_1 + \rho \Delta y_2) \tag{3.24}$$

At this point, the general form of the Runge-Kutta equations has been developed. These equations are summarized below

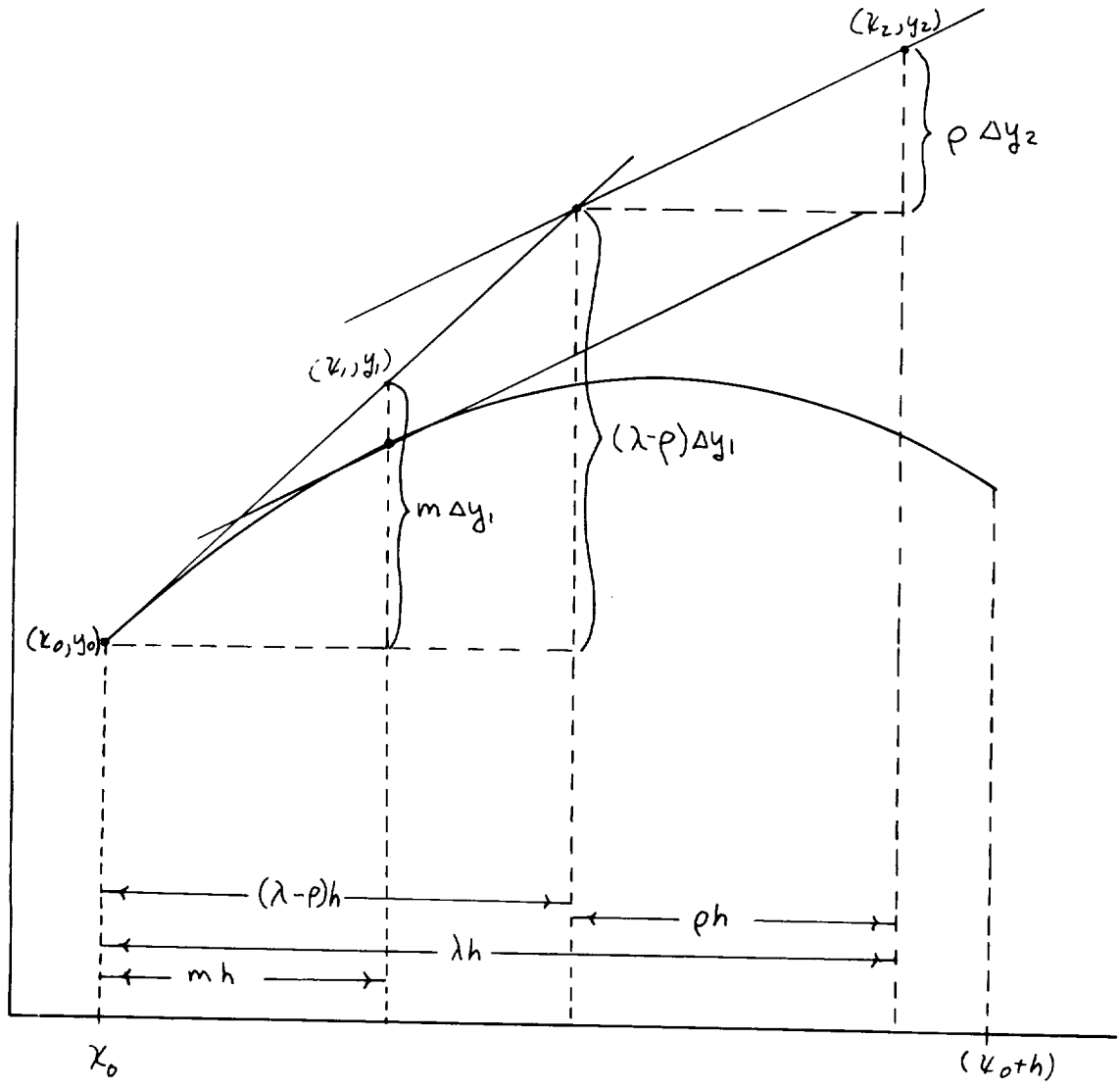


Figure 3.1. Illustration of y_2 Calculation

$$
\left.
\begin{aligned}
&y(x_0+h) - y(x_0) = a_1 \Delta y_1 + a_2 \Delta y_2 + a_3 \Delta y_3 \\
&\Delta y_1 = h\, f(x_0, y_0) \\
&\Delta y_2 = h f(x_0 + mh, y_0 + m \Delta y_1) \\
&\Delta y_3 = h f(x_0 + \lambda h, (\lambda - \rho)\Delta y_1 + \rho \Delta y_2)
\end{aligned}
\right\} \qquad (3.25)
$$

It is now necessary to evaluate the constants a_1, a_2, a_3, m, λ, and ρ. To accomplish this objective, three Δy_i's of equations (3.20), (3.22), and (3.24) are now expanded in a two dimensional Taylor series about the point (x_0, y_0). In the following equations, it will be assumed that the function $f(x, y)$ and its partial derivative $f_x(x, y)$, $f_y(x, y)$, $f_{xy}(x, y)$, etc., are evaluated at the point (x_0, y_0) and the arguments will therefore be omitted for brevity.

$$\Delta y_1 = hf \qquad (3.26)$$

$$\Delta y_2 = hf + mh^2(f_x + ff_y) + \tfrac{1}{2} m^2 h^2 (f_{xx} + 2ff_{xy} + f^2 f_{yy}) + \cdots \qquad (3.27)$$

$$\Delta y_3 = hf + \lambda h^2 (f_x + ff_y) + \tfrac{1}{2} h^3 \left[\lambda^2 (f_{xx} + 2f_{xy} f + f_{yy} f^2) + 2m\rho (f_x + f_y f) f_y\right] + \cdots \qquad (3.28)$$

If Δy is expanded in a Taylor series about (x_0), there is obtained

$$\Delta y = f h + \tfrac{1}{2}(f_x + f_y f) + \tfrac{1}{6} h^3 \left[f_{xx} + 2f_{xy} f + f_{yy} f^2 + (f_x + f f_y)\right] + \cdots \qquad (3.29)$$

The relation between Δy and the Δy_i's of (3.26), (3.27), and (3.28) is given by equation (3.19). Thus, if equation (3.26) is multiplied by a_1, equation (3.21) is multiplied by a_2 equation (3.18) is multiplied by a_3 and these results added, Δy is approximately obtained. Comparing like powers of h in the resultant series for Δy and that obtained in equation (3.29) the following set of equations is obtained if equality to terms of the order h^4 are desired

$$
\left.
\begin{aligned}
&a_1 + a_2 + a_3 = 1 \\
&a_2 m + a_3 \lambda = \tfrac{1}{2} \\
&a_2 m^2 + a_3 \lambda^2 = \tfrac{1}{2} \\
&a_3 \rho m = \tfrac{1}{6}
\end{aligned}
\right\} \qquad (3.30)
$$

This system of equations is underdetermined; thus, the choice for two of the six parameters is arbitrary. Several choices are discussed in the next sections. It is noted here only that there are some choices which are not allowable. If either a_3, ρ , or m is zero, the last equation of (3.30) cannot be satisfied; if both a_1 and a_2 are zero, the first three equations of (3.30) cannot be satisfied.

From a cursory examination of the derivation leading to equation (3.30) it may seem that the underdetermined nature of that set of equations could be eliminated by requiring equality to the Taylor series for y to terms of order h^5. If this were done, there would result two additional equations when the coefficients of h^4 and h^5 are compared, thus yielding an even determined system and reducing the error to the order of h^6. The difficulty with such a procedure is that equality of the coefficient of h^4 and h^5 cannot be attained without imposing severe restrictions on f(x, y).

2.3.3.2.3 Heun's Formulas. It is clear that some simplifications of equation (3.25) can be made by suitable choices of the two free parameters. Henn considered a family of equations in which $\lambda = \rho$. If the further simplification $a_2 = 0$ is made Heuns formulas are obtained

$$\left.\begin{aligned} \Delta y &= \tfrac{1}{4}\Delta y_1 + \tfrac{3}{4}\Delta y_3 \\ \Delta y_1 &= f(x_0, y_0)\,h \\ \Delta y_2 &= f(x_0 + \tfrac{1}{3}h,\ y_0 + \tfrac{1}{3}\Delta y_1)\,h \\ \Delta y_3 &= f(x_0 + \tfrac{2}{3}h,\ y_0 + \tfrac{2}{3}\Delta y_2)\,h \end{aligned}\right\}$$

2.3.3.2.4 Kutta-Simpson Third-Order Rule. This set of equations reduces to Simpson's rule for the evaluation of an integral for the case where f(x,y) is a function of x alone

$$\left.\begin{aligned} \Delta y &= \tfrac{1}{6}\Delta y_1 + \tfrac{2}{3}\Delta y_2 + \tfrac{1}{6}\Delta y_3 \\ \Delta y_1 &= f(x_0, y_0)\,h \\ \Delta y_2 &= f(x + \tfrac{1}{2}h,\ y_0 + \tfrac{1}{2}\Delta y_1)\,h \\ \Delta y_3 &= f(x + h,\ y_0 - \Delta y_1 + 2\Delta y_2)\,h \end{aligned}\right\} \qquad (3.31)$$

If f(x,y) = f(x), then the first equation of (3.31) reduces to

$$\Delta y = \frac{h}{6}\left[y'(x_0) + \Delta y'(x_0 + \frac{h}{2}) + y'(x+h)\right]$$

which is the expression for the evaluation of the integral

$$\Delta y = \int_{x_0}^{x_0+h} f(x)\,dx$$

by Simpson's rule.

2.3.3.2.5 Runge-Kutta-Ralston. In section 2.2.4.2.7, an expression for the truncation error committed at each step is developed. Ralston (Reference 2.13) developed this expression and then chose the parameters so that the truncation error would be minimized

$$\Delta y = 2/9\,\Delta y_1 + 1/3\,\Delta y_2 + 4/9\,\Delta y_3$$

$$\Delta y_1 = h f(x_0, y_0)$$

$$\Delta y_2 = h f(x_0 + 1/2\,h,\ y_0 + 1/2\,\Delta y_1)$$

$$\Delta y_3 = h f(x_0 + 3/4\,h,\ y_0 + 3/4\,\Delta y_2)$$

2.3.3.2.6 Computational Procedure. A sample problem and description of the computational procedure for a fourth order method is presented in section 2.2.4.3. This method is sufficiently general that it may easily be extended to third order formulas. Therefore, no computational procedure is discussed here and the reader is directed to section 2.2.3.3.7, for a representative discussion.

2.3.3.2.7 Error Analysis. A bound for the truncation error committed at each step can be determined for the general third order equation by writing the equations for Δy_1, Δy_2, Δy_3 and Δy as Taylor series with remainders in place of the infinite Taylor series of equations (3.26), (3.27), (3.28), and (3.29)

$$\Delta y_1 = (hf)_{x=x_0}$$

$$\Delta y_2 = \left\{ \left[hf + mh^2(f_x + ff_y) + \frac{1}{2}m^2h^3(f_{xx} + 2ff_{xy} + f^2f_{yy}) \right]_{x=x_0} + \frac{m^3h^4}{6}\left[f_{xxx} + 3ff_{xxy} + 3f^2f_{yyx} + f^3f_{yyy} \right]_{x=\xi} \right\}$$

$$\Delta y_3 = \Bigg\{ \Big[hf + \frac{\lambda h^2}{2}(f_x + f f_y) + \frac{h^3}{2}\big(\lambda^2 [f_{xx} + 2 f_{xy} f + f^2 f_{yy}] + 2 m\rho (f_x + f f_y) f_y\big) \Big]_{x = x_0}$$

$$+ \frac{h^4}{6}\Big[\lambda^3 (f_{xxx} + 3 f f_{xxy} + 3 f^2 f_{yyx} + f^3 f_{yyy}) + 3\rho m^2 f_y (f_{xx} + 2 f f_{xy} + f^2 f_{yy})$$

$$+ \rho m \lambda (f_x f_{xy} + f f_x f_{yy} + f f_y f_{xy} + f^2 f_y f_{yy}) \Big]_{x = \xi_1} \Bigg\}$$

$$\Delta y = \Bigg\{ \Big[hf + \frac{h^2}{2}(f_x + f f_y) + \frac{h^3}{6}(f_{xx} + 2 f f_{xy} + f^2 f_{yy} + f f_x + f^2 f_y) \Big]_{x = x_0}$$

$$+ \frac{h^4}{24}\Big[f_{xxx} + 3 f_{xxy} + 3 f_{xxy} + 3 f^2 f_{yyx} + f^3 f_{yyy} + f_y f_{xx} + 2 f f_y f_{xy}$$

$$+ f_y f^2 f_{yy} + f_y^2 f_x + f_y^2 f f_y + 3 (f_x f_{xy} + f f_x f_{yy} \qquad (3.32)$$

$$+ f f_y f_{xy} + f^2 f_y f_{yy}) \Big]_{x = \xi_2} \Bigg\}$$

Equation (3.32) is the exact solution for Δ y. The approximate Runge-Kutta solution is given by the weighted sum

$$\Delta y_{RK} = a_1 \Delta y_1 + a_2 \Delta y_2 + a_3 \Delta y_3 \qquad (3.33)$$

The truncation error is the difference between equations (3.32) and (3.33)

$$E = \Delta y_{RK} - \Delta y$$

The constants a_1, a_2, a_3, λ, ρ, and m have been chosen such that the coefficients of h, h^2, and h^3 in equation (3.32) match exactly those of equation (3.33) so that error E will contain only the h^4 term

$$E = h^4 \Big\{ \frac{a_2 m^3}{6}\Big[f_{xxx} + 3f f_{xxy} + 3f^2 f_{yyx} + f^3 f_{yyy}\Big]_{x=\xi}$$

$$+ \frac{a_3}{6}\Big[\lambda^3 (f_{xxx} + 3f f_{xxy} + 3f^2 f_{yyx} + f^3 f_{yyy})$$

$$+ 3\rho m^2 f_y (f_{xx} + 2f f_{xy} + f^2 f_{yy}) + \rho m \lambda (f_x f_{xy}$$

$$+ f f_x f_{yy} + f f_y f_{xy} + f^2 f_y f_{yy})\Big]_{x=\xi_1} - \frac{1}{24}\Big[f_{xxx} + 3f_{xxy} \qquad (3.34)$$

$$+ 3f_{xyy} + 3f_{yyx} + f^3 f_{yyy} + f_y f_{xx} + 2f f_y f_{xy} + f_y f^2 f_{yy}$$

$$+ f_y^2 f_x + f_y^2 f f_y + 3(f_x f_{xy} + f f_x f_{yy} + f f_y f_{xy} + f^2 f_y f_{yy})\Big]_{x=\xi_2}\Big\}$$

If, in the region where a solution is desired, the bounds M and L defined as follows

$$|f(x,y)| \le M \qquad (3.35)$$

$$\left|\frac{\partial^{i+j} f}{\partial x^i \, \partial y^j}\right| \le \frac{L^{i+j}}{M^{j-1}} \qquad (3.36)$$

exist, then, a bound for E can be found. Substituting (3.35) and (3.36) in (3.34) the error bound is

$$E \le h^4 M L^3 \Big[\frac{4}{3} a_2 m^3 + \frac{a^3}{6}(8\lambda^3 + 12\rho m^2 + 24\rho m \lambda) - \frac{26}{24}\Big] \qquad (3.37)$$

If a method of determining truncation error is to be used as a means of determining step size equation (3.37) may be of little value because the bounds M and L may not be known before hand.

An approximate value for the truncation error can be found (Reference 2.14) from the results of two integrations one of which has step size h and the other has step size 2h. If y(x) is the true value, $y_1(x)$ is the value obtained for y(x) using step size h, and $y_2(x)$ is the value obtained using step size 2h then using (3.34)

$$y(x) = y_1(x) + C_1 (2h)^4$$

$$y(x) = y_2(x) + 2C_2 h^4$$

The approximate error, assuming $C_1 \approx C_2$ is

$$E = y - y_2 = \frac{y_2 - y_1}{2^r - 1} \tag{3.38}$$

Some care should be taken, however, if equation (3.38) is to be used as a criterion for changing step size. If the step size is to be changed, it is probably because the solution curve changes rapidly and a constant step size is inadequate. If this is the case then the assumption that $C_1 \approx C_2$ will most likely not be satisfied.

In Reference (2.15), it is shown that for a stable differential equation the accumulated error after a number of steps is proportional to the maximum error committed at each step divided by the step size h. The constant of proportionality does not vary with the number of the step so that a bound for the accumulated error can be found which is independent of the number of steps taken. Thus, the Runge-Kutta solution is stable in the sense that for a stable differential equation the error will be bounded and the truncation error will go to zero as the step size h, goes to zero. For more detailed information, see References (2.14), (2.16), and (2.17).

2.3.3.3 Fourth Order Methods

2.3.3.3.1 Discussion. Fourth order Runge-Kutta methods reduce the error by one order over third order methods at the expense of an additional derivative evaluation. It should be noted that while third order methods require three derivative evaluations, and fourth order requires four derivative evaluation methods of fifth order require six evaluations and methods of sixth order require eight evaluations. It is probably for this reason that the fourth order methods are the most popular; i.e., they provide a good compromise between step size and number of derivative evaluations.

2.3.3.3.2 Derivation of General Fourth Order Runge-Kutta Formulas. The derivation of the fourth-order Runge-Kutta equations is similar to the derivation of the third order equations presented in section 2.2.4.2.2 and the bulk of the derivation is not repeated. For the fourth order case, it is necessary to compute a Δy_4.

$$\Delta y_4 = h f(x_3, y_3)$$

The value of y_3 is determined in a fashion similar to the determination of y_2 in equation (13) of section 2.2.4.2.2. Three steps are used in determining y_3 instead of the two used for y_2 with the derivatives at the points (x_0, y_0), (x_1, y_1), and (x_2, y_2) being used over the arbitrary intervals α, β, σ.

The general form of the fourth order equations is given in the following equation:

$$
\left.
\begin{aligned}
\Delta y &= a_1 \Delta y_1 + a_2 \Delta y_2 + a_3 \Delta y_3 + a_4 \Delta y_4 \\
\Delta y_1 &= h f(x_0, y_0) \\
\Delta y_2 &= h f(x_0 + m h, y_0 + m \Delta y_1) \\
\Delta y_3 &= h f(x_0 + \lambda h, y_0 + (\lambda - \rho)\Delta y_1 + \rho \Delta y_2) \\
\Delta y_4 &= h f(x_0 + \gamma h, y_0 + (\gamma - \beta - \sigma)\Delta y_1 + \beta \Delta y_2 + \sigma \Delta y_3)
\end{aligned}
\right\}
$$

The relations between the parameters are given by equation (3.39)

$$
\left.
\begin{aligned}
& a_1 + a_2 + a_3 + a_4 = 1 \\
& a_2 m + a_3 \lambda + a_4 \gamma = \frac{1}{2} \\
& a_2 m^2 + a_3 \lambda^2 + a_4 \gamma^2 = \frac{1}{3} \\
& a_2 m^3 + a_3 \lambda^3 + a_4 \gamma^3 = \frac{1}{4} \\
& a_3 m \rho + a_4 (m\beta + \lambda\sigma) = \frac{1}{6} \\
& a_3 m^2 \rho + a_4 (m\beta + \lambda^2 \sigma) = \frac{1}{12} \\
& a_3 m \lambda \rho + a_4 (m\beta + \lambda\sigma) = \frac{1}{8} \\
& a_4 m \rho \sigma = \frac{1}{24}
\end{aligned}
\right\} \qquad (3.39)
$$

The set of equations (3.39) relating the parameters is, as was the case with the third order system, underdetermined and two of the ten parameters may be chosen arbitrarly.

2.3.3.3.3. Classical Runge-Kutta. This member of the family is the best known and it is often the form that is being thought of when the term Runge-Kutta is used. This form is also known as the Kutta-Simpson One-third Rule and, as might be expected, reduces to Simpson's one-third rule when $f(x,y) \equiv f(x)$

$$
\left.
\begin{aligned}
\Delta y &= \frac{1}{6}\Delta y_1 + \frac{1}{3}\Delta y_2 + \frac{1}{3}\Delta y_3 + \Delta y_4 \\
\Delta y_1 &= h f(x_0, y_0) \\
\Delta y_2 &= h f(x_0 + \frac{1}{2}h, y_0 + \frac{1}{2}\Delta y_1) \\
\Delta y_3 &= h f(x + \frac{1}{2}h, y_0 + \frac{1}{2}\Delta y_2) \\
\Delta y_4 &= h f(x + h, y_0 + \Delta y_3)
\end{aligned}
\right\}
$$

2.3.3.3.4 Kutta-Simpson Three-eights Rule. This form reduces to Simpson's three-eights rule when f(x, y) = f(x)

$$\left.\begin{aligned}
\Delta y &= \tfrac{1}{8}\Delta y_1 + \tfrac{3}{8}\Delta y_2 + \tfrac{3}{8}\Delta y_3 + \Delta y_4 \\
\Delta y_1 &= h f(x_0, y_0) \\
\Delta y_2 &= h f(x_0 + \tfrac{1}{3}h,\ y_0 + \tfrac{1}{3}\Delta y_1) \\
\Delta y_3 &= h f(x_0 + \tfrac{2}{3}h,\ y_0 - \tfrac{1}{3}\Delta y_1 + \Delta y_2) \\
\Delta y_4 &= h f(x_0 + h,\ y_0 + \Delta y_1 - \Delta y_2 + \Delta y_3)
\end{aligned}\right\}$$

2.3.3.3.5 Runge-Kutta-Ralston. This form provides minimum truncation error at each step

$$\Delta y = .17476028\,\Delta y_1 - .55148066\,\Delta y_2 + 1.20553560\,\Delta y_3 + .17118478\,\Delta y_4$$

$$\Delta y_1 = h f(x_0, y_0)$$

$$\Delta y_2 = h f(x_0 + .4h,\ y_0 + .4\,\Delta y_1)$$

$$\Delta y_3 = h f(x_0 + .45573725h,\ y_0 + .29697761\,\Delta y_1 + .15875964\,\Delta y_2)$$

$$\Delta y_4 = h f(x_0 + h,\ y_0 + .21810040\,\Delta y_1 - 3.05096516\,\Delta y_2 + 3.83286476\,\Delta y_3)$$

2.3.3.3.6 Runge-Kutta-Gill. Gill's formulas were developed specifically for use on high speed digital computers so that: 1) a minimum number of storage registers are used; 2) growth of round-off errors is controlled; and, 3) comparatively few instructions are required. [Reference (2.16) and (2.18)]. In the notation previously used, the Runge-Kutta-Gill formulas are

$$\left.\begin{aligned}
\Delta y &= \tfrac{1}{6}\Delta y_1 + \tfrac{1}{3}(1 - \sqrt{\tfrac{1}{2}})\Delta y_2 + \tfrac{1}{3}(1 + \sqrt{\tfrac{1}{2}})\Delta y_3 + \tfrac{1}{6}\Delta y_4 \\
\Delta y_1 &= h f(x_0, y_0) \\
\Delta y_2 &= h f(x_0 + \tfrac{h}{2},\ y_0 + \tfrac{1}{2}\Delta y_1) \\
\Delta y_3 &= h f(x_0 + \tfrac{h}{2},\ y_0 + (-\tfrac{1}{2} + \sqrt{\tfrac{1}{2}})\Delta y_1 + (1 - \sqrt{\tfrac{1}{2}})\Delta y_2) \\
\Delta y_4 &= h f(x_0 + h,\ y_0 + (-\sqrt{\tfrac{1}{2}})\Delta y_2 + (1 + \sqrt{\tfrac{1}{2}})\Delta y_3)
\end{aligned}\right\} \qquad (3.40)$$

In order to get a more efficient form equivalent to (3.40), the auxiliary quantities, q_i and y_i, are introduced and the set of equations become

$$
\left.
\begin{aligned}
\Delta y_1 &= hf(x_0, y_0) \\
\Delta y_2 &= hf\left(x_0 + \frac{h}{2}, y_1\right) \\
\Delta y_3 &= hf\left(x_0 + \frac{h}{2}, y_2\right) \\
\Delta y_4 &= hf(x_0 + h, y_3) \\
y_1 &= y_0 + \tfrac{1}{2}\Delta y_1 \\
y_2 &= y_1 + \left(1 - \sqrt{\tfrac{1}{2}}\right)(\Delta y_2 - q_1) \\
y_3 &= y_2 + \left(1 + \sqrt{\tfrac{1}{2}}\right)(\Delta y_3 - q_2) \\
y_4 &= y_3 + \tfrac{1}{6}(\Delta y_4 - 2q_3) \\
q_1 &= \Delta y_1 \\
q_2 &= (2 - \sqrt{2})\Delta y_2 + \left(-2 + 3\sqrt{\tfrac{1}{2}}\right) q_1 \\
q_3 &= (2 + \sqrt{2})\Delta y_3 + \left(-2 - 3\sqrt{\tfrac{1}{2}}\right) q_2
\end{aligned}
\right\} \qquad (3.41)
$$

Note that from these definitions

$$y_4 = \Delta y + y_0$$

At any stage in (3.41) the Δy_i, y_i, and q_i depend only on stored quantities of the previous stage and quantities computed during the current stage. As each quantity is computed, it can be stored in the same register where the corresponding quantity of the previous stage was stored. Since the quantity there is no longer required. Hence the overall process requires the storage of only three quantities at each stage and the same three registers can be used at each stage.

A further refinement to compensate for some round-off error can be made without increasing the complexity of the procedure with no increase in storage requirements. This refinement consists of introducing q_0 and q_4 with q initially zero. The quantity q_4 represents approximately three times the round-off error in y_4 accumulated in one step. To compensate for the accumulated round-off q_4 is used for q_0 in the next step. The final form of the Runge-Kutta-Gill equations is

$$\Delta y_1 = hf(x_0, y_0)$$

$$\Delta y_2 = hf(x_0 + \frac{h}{2}, y_1)$$

$$\Delta y_3 = hf(x_0 + \frac{h}{2}, y_2)$$

$$\Delta y_4 = hf(x_0 + h, y_3)$$

$$y_1 = y_0 + \frac{1}{2}(\Delta y_1 - 2q_0)$$

$$y_2 = y_1 + (1 - \sqrt{\frac{1}{2}})(\Delta y_2 - q_1)$$

$$y_3 = y_2 + (1 + \sqrt{\frac{1}{2}})(\Delta y_3 - q_2)$$

$$y_4 = y_3 + \frac{1}{6}(\Delta y_4 - 2q_3)$$

$$q_1 = q_0 + 3\left[\frac{1}{2}(\Delta y_1 - 2q_0)\right] - \frac{1}{2}\Delta y_1$$

$$q_2 = q_1 + \left[(1 - \sqrt{\frac{1}{2}})(\Delta y_2 - q_1)\right] - (1 - \sqrt{\frac{1}{2}})\Delta y_2$$

$$q_3 = q_2 + 3\left[(1 + \sqrt{\frac{1}{2}})(\Delta y_3 - q_2)\right] - (1 + \sqrt{\frac{1}{2}})\Delta y_3$$

$$q_4 = q_3 + 3\left[\frac{1}{6}(\Delta y_4 - 2q_3)\right] - \frac{1}{2}\Delta y_4$$

2.3.3.3.7 Computational Procedure. The following discussion and sample problem are taken from Reference (2.16). The Runge-Kutta-Gill formulation will be used for the sample problem. These equations are given in Section 2.3.4.3.6. The problem to be illustrated is the solution of the set of simultaneous differential equations

$$\left.\begin{aligned} y_1' &= \frac{1}{y^2} & y_1(0) &= 1 \\ y_2' &= -\frac{1}{y_1} & y_2(0) &= 1 \end{aligned}\right\} \qquad (3.42)$$

The analytic solutions to these equations are

$$y_1 = e^x$$

$$y_2 = e^{-x}$$

The error in the numerical calculation will be defined as

$$\epsilon_1 = e^x - y_{14}(x)$$

$$\epsilon_2 = e^{-x} - y_{24}(x)$$

For the Runge-Kutta calculation, the set of equations (3.42) will be written symbolically as

$$y_i' = f_i(y_0(x), y_1(x), y_2(x)) \qquad i = 1, 2$$

where

$$y_0(x) \triangleq x$$

and the subscript i denotes the first or second equation of the set (1). The values of the y_j's of equation (3.42) will be denoted by the second subscript j; i.e.,

$$y_i(x+h) = y_{i4}$$

The calculation proceeds in the following sequence

(1) Let j = 1
(2) Let i = 0
(3) Compute $\Delta Y_{i,j}$
(4) Repeat (3) until all value of i are used (i = 0, 1, 2)
(5) Set i = 0
(6) Compute $y_{i,j}$, and $q_{i,j}$
(7) Repeat (6) until all values of i are used
(8) Repeat steps (3) through (7) incrementing j each time until j = 4
(9) The desired answers are obtained as y_{14} and y_{24}

A flow chart depicting the above sequence is shown in Figure 3.2;

2.3.3.3.8 Error Analysis. The derivation of the truncation error is similar to that for the third order method and the reader is referred to Section 2.3.3.2.7. The bound for the truncation error is [Reference (2.13)]

$$E \le ML^4 \left\{ 16|b_1| + 4|b_2| + |b_2 + 3b_3| + |2b_2 + 3b_3| + |b_2 + b_3| + |b_3| + 8|b_4| \right.$$

$$\left. + |b_5| + |2b_5 + b_7| + |b_5 + b_6 + b_7| + |b_6| + |2b_6 + b_7| + |b_7| + 2|b_8| \right\}$$

where M and L are bounds on the derivatives over the region in which a solution is desired

$$M > |f(x, y)|$$

$$\frac{L^{i+j}}{M^{j-1}} \ge \left| \frac{\partial^{i+j} f}{\partial x^i \partial y^j} \right|$$

and

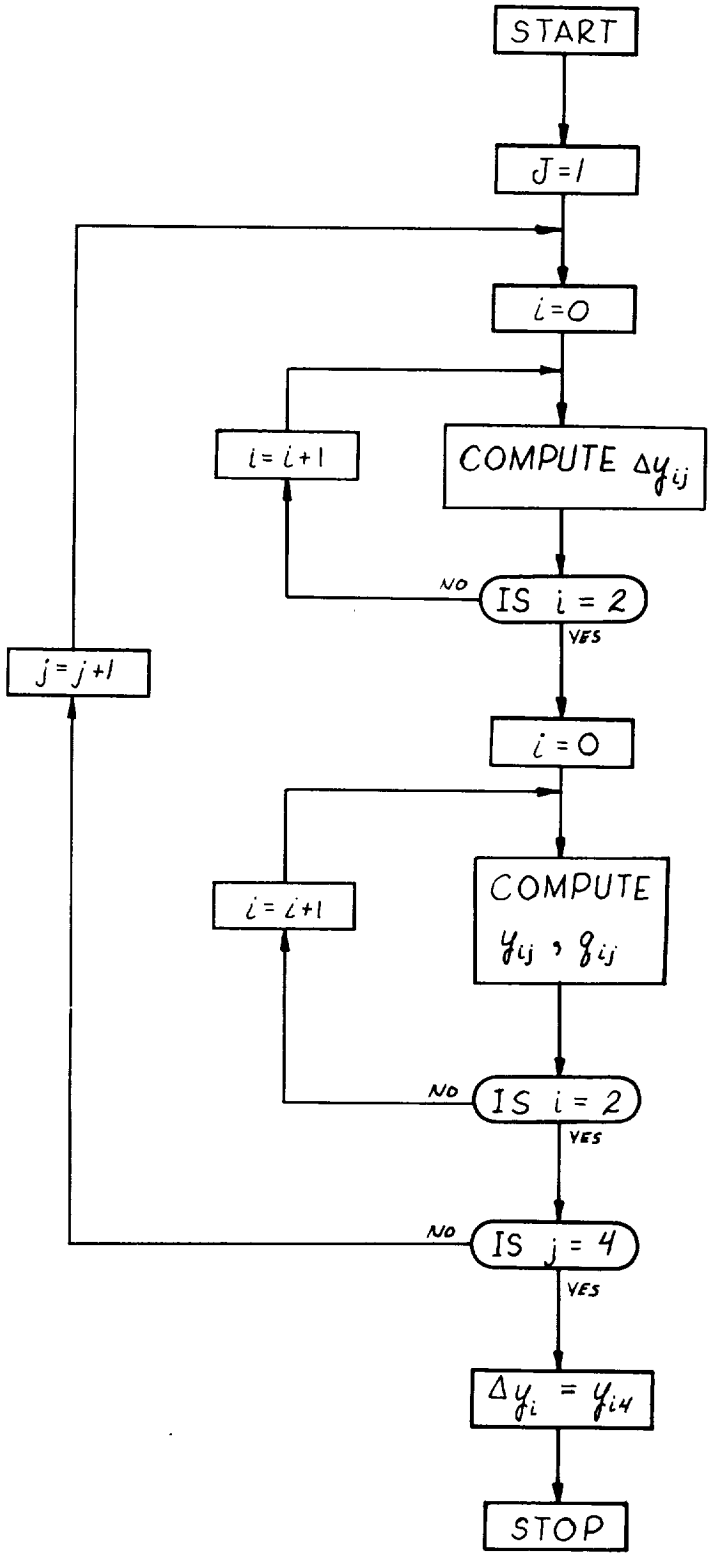


Figure 3.2. Flow Diagram of Runge-Kutta Calculation

Table 3.2 presents the numerical results of two steps in the calculation with h = .1.

Table 3.2. Intermediate Steps in Runge-Kutta-Gill Calculation

i	j	Y_{i_j}	Y_{i_j}	q_{i_j}
0		.10000000 x 10^{1}	.50000000 x 10^{-1}	.10000000 x 10^{-1}
1	1	.10000000 x 10^{1}	.10500000 x 10^{1}	.10000000 x 10^{1}
2		-.10000000 x 10^{1}	.95000000	-.10000000 x 10^{1}
0		.10000000 x 10^{1}	.50000000 x 10^{-1}	.70710678
1	2	.10526316 x 10^{1}	.10515415 x 10^{1}	.73793765
2		-.95238095	.95139473	-.67921219
0		.10000000 x 10^{1}	.99999999 x 10^{-1}	.50000000
1	3	.10510884	.11049997 x 10^{1}	.54736297
2		-.95098478	.90500025	-.44761411
0		.10000000 x 10^{1}	.99999999 x 10^{-1}	-.18189894 x 10^{-11}
1	4	.11049721 x 10^{1}	.11051705 x 10^{1}	-.36379788 x 10^{-11}
2		-.90497760	.90483776	-.90947470 x 10^{-12}
0		.10000000 x 10^{1}	.15000000	.10000000 x 10^{1}
1	1	.11051705 x 10^{1}	.11604290 x 10^{1}	.11051705 x 10^{1}
2		-.90483776	.85959587	-.90483776
0		.10000000 x 10^{1}	.15000000	.70710678
1	2	.11633374 x 10^{1}	.11621327 x 10^{1}	.81554692
2		-.86175025	.86085787	-.61457684
0		.10000000 x 10^{1}	.20000000	.50000000
1	3	.11616319 x 10^{1}	.12212131 x 10^{1}	.60492940
2		-.86048694	.81887839	-.40501815
0		.10000000 x 10^{1}	.20000000	-.18189894 x 10^{-11}
1	4	.12211825 x 10^{1}	.12214018 x 10^{1}	-.36379788 x 10^{-11}
2		-.81885791	.81873137	-.90949470 x 10^{-12}

Table 3.3 lists the final values at each step and the corresponding errors, with h = .1.

Table 3.3. Results of Sample Problem

x	$y_1(x)$	$y_2(x)$	$\varepsilon_1 x 10^7$	$\varepsilon_2 x 10^7$
0	$.10000000 \times 10^1$	$.10000000 \times 10^1$	---	---
.1	$.11051705 \times 10^1$	.90483776	4	- 3.4
.2	$.12214018 \times 10^1$	.81873137	10	- 6.2
.3	$.13498573 \times 10^1$	.74081905	15	- 8.3
.4	$.14918224 \times 10^1$	.67032105	23	-10.0
.5	$.16487181 \times 10^1$	.60653179	32	-11.3
.6	$.18221146 \times 10^1$	.54881286	42	-12.2
.7	$.20137474 \times 10^1$	.49658660	53	-13.0
.8	$.22255342 \times 10^1$	.44933030	67	-13.4
.9	$.24595947 \times 10^1$	.40657102	84	-13.6
1.0	$.27182715 \times 10^1$	.36788081	103	-13.7

$$b_1 = 1/120 - 1/24(m^4 a_2 + \lambda^4 a_3 + a_4)$$

$$b_2 = 1/20 - 1/2\,(m\lambda^2 \rho a_3 + a_4 m\beta + a_4 \lambda\sigma)$$

$$b_3 = 1/120 - 1/6(m^3 \rho a_3 + a_4 m^3 \beta + a_4 \lambda^3 \sigma)$$

$$b_4 = 1/30 - 1/2\,(m\lambda\rho a_3 + a_4 m^2\beta + a_4\lambda^2\sigma)$$

$$b_5 = 1/120 - 1/2\, m^2 \rho\sigma a_4$$

$$b_6 = 1/40 - 1/2\left[m^2\rho^2 a_3 + a_4(m\beta + \lambda\sigma)^2\right]$$

$$b_7 = 7/120 - m(1+\lambda)\rho\sigma a_4$$

$$b_8 = 1/120$$

2.3.3.4 Higher Order Methods

2.3.3.4.1 Discussion. Formulas for Runge-Kutta method of order higher that four are not widely used nor well known. These formulas have the disadvantage of requiring more evaluations of the derivative than the order of the formula; e.g., the fifth-order formula of Kutta-Nystrom requires six evaluations of the derivative while the sixth-order formula of Huta [Reference (2.19)] requires eight. The reason for requiring the extra derivative evaluations is that a contradiction exists if the relation between the parameters is written in the same form as was done for the third and fourth order methods. A different approach taken by Shanks [Reference (2.20)] is to require the equations which result in a contradiction to be only approximately true. Using this technique, Shanks develops equations through order eight which may be found in Reference (2.20).

2.3.3.4.2 Derivation. The derivation of the higher order equations is similar to that outlined for the third order method except that additional evaluations of the derivative are required in order to get a consistant set of equations relating the parameters. For the Shanks method, the contradictory set of equations obtained by n derivative equations for an nth order system is modified so that the conditions for order n-1 are met, and the remaining equations modified by addition of arbitrary parameters which represent small errors.

2.3.3.4.3 Runge-Kutta-Nystrom (fifth order).

$$\Delta y = \frac{23}{192}\Delta y_1 + \frac{125}{192}\Delta y_3 - \frac{81}{192}\Delta y_5 + \frac{125}{192}\Delta y_6$$

$$\Delta y_1 = hf(x_0\ y_0)$$

$$\Delta y_2 = hf\left(x_0 + \frac{h}{3},\ y_0 + \frac{\Delta y_1}{3}\right)$$

$$\Delta y_3 = hf\left(x_0 + \frac{2h}{5},\ y_0 + \frac{4}{25}\Delta y_1 + \frac{6}{25}\Delta y_2\right)$$

$$\Delta y_4 = hf\left(x_0 + h,\ y_0 + \frac{1}{4}\Delta y_1 - 3\Delta y_2 + \frac{15}{4}\Delta y_3\right)$$

$$\Delta y_5 = hf\left(x_0 + \frac{2h}{3},\ y_0 + \frac{6}{81}\Delta y_1 + \frac{90}{81}\Delta y_2 - \frac{50}{81}\Delta y_3 + \frac{8}{81}\Delta y_4\right)$$

$$\Delta y_6 = hf\left(x_0 + \frac{4h}{5},\ y_0 + \frac{6}{75}\Delta y_1 + \frac{36}{75}\Delta y_2 + \frac{10}{75}\Delta y_3 + \frac{8}{75}\Delta y_4\right)$$

2.3.3.4.4 Shanks Fifth Order Formula

$$\Delta y = \frac{105}{1134}\Delta y_1 + \frac{500}{1134}\Delta y_3 + \frac{448}{1134}\Delta y_4 + \frac{81}{1134}\Delta y_5$$

$$\Delta y_1 = hf(x_0\ y_0)$$

$$\Delta y_2 = hf\left(x_0 + \frac{h}{9000},\ y_0 + \frac{h}{9000}\right)$$

$$\Delta y_3 = hf\left(x_0 + \frac{3h}{10},\ y_0 - 404.7\Delta y_1 + 405\Delta y_2\right)$$

$$\Delta y_4 = hf\left(x_0 + \frac{3h}{4},\ y_0 + \frac{20241}{8}\Delta y_1 - \frac{20250}{8}\Delta y_2 + \frac{15}{8}\Delta y_3\right)$$

$$\Delta y_5 = hf\left(x_0 + h,\ y_0 - \frac{931041}{81}\Delta y_1 + \frac{931500}{81}\Delta y_2 - \frac{490}{81}\Delta y_3 + \frac{112}{81}\Delta y_4\right)$$

2.3.3.4.5 Runge-Kutta-Fehlberg. This technique will provide sixth order accuracy with only three derivative evaluations. The disadvantage is that it requires extensive preliminary analysis of the particular problem involved and therefore cannot be used as a general integration package. The procedure is to form an auxiliary differential equation and develop certain recursion relations. Durham [Reference (2.22)] has applied this technique to the restricted three-body problem and determined that it was superior to a single step Lie series method, a multi-step Cowell method and a multi-step Adams method. More information on this method can be found in References (2.21) and (2.22)

2.3.3.4.6 Error Analysis. Error bounds for higher order formulas could not be found although a derivation such as was done in section 2.3.3.2.7 could be performed to yield them.

2.3.3.4.7 Computational Procedure. The computational procedure is similar to that of other method, and section 2.3.3.3.7 illustrates the technique.

2.3.3.5 Comparison

One comparison which can be made between the various Runge-Kutta formulas is that in general for a given step size, the lower the order of the method the less computation time required. The reduced computation time is obtained at the expense of a larger truncation error. One exception to this rule is the Fehlberg technique which trades speed in the actual computation for extensive analysis of the particular equation whose solution is desired. If the Runge-Kutta technique is to be used for the entire solution, the technique such as the Runge-Kutta-Gill formulation which minimizes computer storage and controls round-off should be used. On the other hand, if the method is used to start a solution, the primary concern is with truncation error, and the Runge-Kutta-Ralston formulas should be used.

In summary, it can be stated that the large number of derivatives which must be evaluated at each step generally relegate the Runge-Kutta methods to applications requiring special starting procedures or to solutions where only a few points are calculated, and the slowness of the process if unimportant. Although each individual problem is more suited to a particular order, if a general recommendation is to be made, the fourth order methods will be chosen. In particular for starting the solution, the fourth order Runge-Kutta-Ralston should be used.

2.3.4 PREDICTOR CORRECTOR METHODS

2.3.4.1 Introduction

Single step methods, such as Runge-Kutta (section 2.2.3), compute the value of the functions, y_{n+1}, at x_{n+1}, by means of a formula which depends only on x_n, y_n, and the step size, h. After the first few points have been computed, there is available a large amount of information which is not being utilized; namely, the history of values of y. In an attempt to obtain greater efficiency and accuracy, formulas have been derived which depend not only on y_n, but on the prior values of the function; for example, y_{n-1} and y_{n-2}. These methods are called multi-step methods and have proved to be significantly faster, for the same accuracy, than single step methods. This advantage is obtained at the expense of some increase in complexity, for multi-step methods are not self-starting; and therefore, a special starting procedure is necessary.

Some multi-step methods develop two formulas - one for use as a predictor, and the other for use as a corrector. That is, one formula is used to predict the next value of y, y_{n+1}. This value of y_{n+1} is substituted into the differential equation, and y'_{n+1} is obtained. These values are used in the corrector formula to obtain another "corrected" value of y_{n+1}. The difference between the predicted and corrected values of y_{n+1} can be monitored to evaluate the effectiveness of the corrector cycle. Multi-step methods involving the use of two such formulas have naturally been called predictor-corrector methods.

The most popular predictor-corrector methods are discussed in this section as well as some of the more promising newer methods.

2.3.4.2 Adams Method

2.3.4.2.1 Discussion. The method of replacing the derivative of a function by a polynomial and integrating that polynomial over an interval to obtain a multi-step predictor formula was used by J. C. Adams as early as 1883. The technique was developed in a text co-authored by F. Bashforth (Reference 2.23); and probably for this reason, the method is sometimes called "Adams-Bashforth" and sometimes simply "Adams".

Although Adams did apply a corrective formula, the method bearing his name usually involves a predictor formula only and will be so discussed in this section.

2.3.4.2.2 Derivation of Equations. The integrand in the differential equation $y' = f(x, y)$ can be approximated with the polynomial formed by Newton's backward-difference formula (Reference 2.12) using the N+1 points $x_{n-N}, \ldots, x_{n-1}, x_n$

$$f_{n+s} = f_n + s\,\nabla f_n + s\,\frac{(s+1)}{2!}\nabla^2 f_n + \cdots + \frac{s(s+1)\ldots(s+N-1)\,\nabla^N f_n}{n!}$$

where $s = \dfrac{x - x_n}{h}$

$$\nabla f_n = f_n - f_{n-1}$$

$$\nabla^2 f_n = \Delta f_n - \Delta f_{n-1} = f_n - 2f_{n-1} + f_{n-2}$$

Integration of both sides of the differential equation from x_n to x_{n+1} gives, after changing the variable on the right-hand side from x to s.

$$y_{n+1} = y_n + h\int_0^1 f_{n+s}\,ds = y_n + h\sum_{k=0}^{N} a_k \nabla^k f_n$$

where

$$a_k = \int_0^1 \frac{s(s+1)\ldots(s+k-1)}{k!}\,ds$$

The leading terms of which are:

$$y_{n+1} = y_n + h\left(1 + \frac{1}{2}\nabla + \frac{5}{12}\nabla^2 + \frac{3}{8}\nabla^3 + \frac{251}{720}\nabla^4 + \frac{95}{288}\nabla^5 + \ldots\right) f_n$$

This is the standard Adams' predictor equation in difference notation.

2.3.4.2.3 Truncation Error. The truncation error corresponding to terminating the series with the Nth difference of f_n is

$$E = a_{N+1}\,h^{N+2}\,y^{(N+2)}(\xi) \qquad (x_{n-N} < \xi < x_{n+1})$$

Error estimation for variable step size integration requires the computation of an extra difference in order to estimate the higher derivative occurring in the truncation error expression.

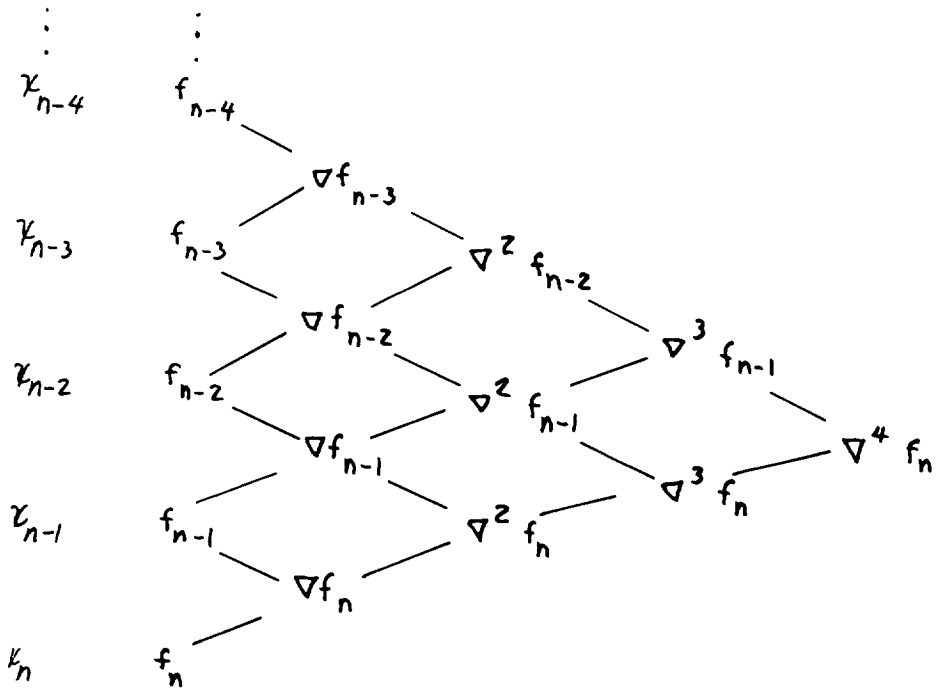
For example, when three differences are retained, the truncation error incurred at each step is

$$\frac{251}{720} h^5 y^v(\xi)$$

and is approximated by 251/720 h $\nabla^4 f_n$. The validity of replacing derivatives by differences is provided by the relation $y^{(n)}(\xi) = (\nabla^n y^{(x)})/h^n$.

2.3.4.2.4 Step Size Modification. Step size modification is generally limited to halving or doubling, which can be performed repeatedly if necessary. Doubling the step size is straightforward and involves only the use of alternate points, thus no new information need be generated. When the step size is halved, however, information not previously available will be required; that is, the value of y and y'=f halfway between the existing points.

The following discussion presents a general method of interpolation. The schematic:



shows that the subscript of the backward differences remains constant along the back diagonal. It is evident then that Adams' method is an end-point extrapolation, and once procedures have been specified for modifying the differences of the end point any integration step size may be performed.

Let ρ^h indicate the desired interpolation where $0 \leq \rho \leq 1$. The step size is halved when $\rho = \frac{1}{2}$. In what follows: ∇ denotes backward difference at spacing h and ∇_1 denotes backward differences at spacing ρ^h the general relationship for the rth difference for a new spacing ρ^h in terms of the differences at spacing h is:

$$(-\nabla_1)^r = \left[(1-\nabla)^\rho - 1\right]^r = \rho^r \left\{(-\nabla)^r + r\,\frac{(\rho-1)}{2}(-\nabla)^{r+1}\right.$$

$$+ \frac{r(\rho-1)}{24}\left[4(\rho-2) + 3(r-1)(\rho-1)\right](-\nabla)^{r+2}$$

$$\left. + \frac{r(\rho-1)}{48}\left[2(\rho-2)(\rho-3) + 4(r-1)(\rho-1)(\rho-2) + (r-1)(r-2)(\rho-1)^2\right](-\nabla)^{r+3} + \cdots\right\}.$$

Setting r = 1, 2, 3, and 4 provides the relationships:

$$\nabla_1 = \rho\nabla - \frac{\rho(\rho-1)}{2}\nabla^2 + \frac{\rho(\rho-1)(\rho-2)\nabla^3}{6} - \frac{\rho(\rho-1)(\rho-2)(\rho-3)\nabla^4}{24} + \cdots$$

$$\nabla_1^2 = \rho^2\nabla^2 - \rho^2(\rho-1)\nabla^3 + \rho^2\frac{(\rho-1)}{12}(7\rho-11)\nabla^4 + \cdots$$

$$\nabla_1^3 = \rho^3\nabla^3 - \frac{3}{2}\rho^3(\rho-1)\nabla^4 + \cdots$$

$$\nabla_1^4 = \rho^4\nabla^4 + \cdots$$

Setting $\rho = \frac{1}{2}$ provides the relationships for halving the step size:

$$\nabla_1 = \frac{\nabla}{2} + \frac{\nabla^2}{8} + \frac{\nabla^3}{16} + \frac{5\nabla^4}{128}$$

$$\nabla_1^2 = \frac{\nabla^2}{4} + \frac{\nabla^3}{8} + \frac{5}{64}\nabla^4$$

$$\nabla_1^3 = \frac{\nabla^3}{8} + \frac{3}{32}\nabla^4$$

$$\nabla_1^4 = \frac{\nabla^4}{16}$$

Setting $\rho = 2$ provides the relationships for doubling the step size: (direct computation of the differences using alternate points previously computed will also provide the modified differences)

$$\nabla_1 = 2\nabla - \nabla^2$$

$$\nabla_1^2 = 4\nabla^2 - 4\nabla^3 + \nabla^4$$

$$\nabla_1^3 = 8\nabla^3 - 12\nabla^4$$

$$\nabla_1^4 = 16\nabla^4$$

2.3.4.2.5 Starting and Restarting Procedure. The application of the Adams' formula employing N differences requires N beginning points in addition to the initial conditions.

The following is a possible method for generating the beginning values, but is not seriously recommended as it requires a very small step size h. The Runge-Kutta method will always provide adequate starting values, and obviates the small step size requirement.

$$y_1 = y_0 + h y_0'$$

$$y_2 = y_1 + h\left[y_1' + \tfrac{1}{2}\nabla y_1'\right]$$

$$y_3 = y_2 + h\left[y_2' + \tfrac{1}{2}\nabla y_2' + \tfrac{5}{12}\nabla^2 y_2'\right]$$

$$y_4 = y_3 + h\left[y_3' + \tfrac{1}{2}\nabla y_3' + \tfrac{5}{12}\nabla^2 y_3' + \tfrac{3}{8}\nabla^3 y_3'\right]$$

$$y_5 = y_4 + h\left[y_4' + \tfrac{1}{2}\nabla y_4' + \tfrac{5}{12}\nabla^2 y_4' + \tfrac{3}{8}\nabla^3 y_4' + \tfrac{251}{720}\nabla^4 y_4'\right]$$

2.3.4.2.6 Computational Procedure. The Adam's predictor formula employing N differences requires N eqidistant derivative values in addition to the initial conditions. Once these starting values are obtained, every application of the formula yields a new function value, the derivative value, of which is obtained by evaluating the differential equation. When the corresponding differences are formed, the predictor is again applied, and the process continues requiring just one derivative value for each integration step.

The estimate of truncation error is monitored at the completion of each integration step. If the error is large, the differences are modified and the step size is halved. Whereas if the truncation error is very small, the differences are modified and the integration step is doubled. The process continues with the step size unchanged when the truncation error tolerance is satisfied, but indicates no excess of significant figures.

An arbitrary integration step size, h*, may be effected for the purpose of satisfying constraints or printed output by modifying the differences such that $h^* = h$.

2.3.4.2.7 Advantages and Disadvantages. Advantages of the Adams' method are:

1) Step size modification is readily accomplished.
2) Requires only one derivative evaluation per integration step.
3) An estimate of the truncation error is obtainable.

Disadvantages of the method are:

1) Requires a special starting/restarting procedure.
2) An extra difference is required to estimate the truncation error.

2.3.4.3 Adams-Moulton Method

2.3.4.3.1 Discussion. Moulton in 1926 applied a corrector formula to the Adams' predictor formula and the resulting method - known as the Adams-Moulton multi-step method - has achieved widespread use. In fact, Adams-Moulton is probably the most commonly used integration method for trajectory problems.

2.3.4.3.2 Derivation of Equations. The equations will be derived first in difference form and then in ordinate form.

The differential equation is:

$$y' = f(x, y)$$

The corrector formula is developed by approximately the integrand with the polynomial formed by the N+1 equally spaced points

$$x_{n-N+1}, \cdots, x_n, x_{n+1}$$

$$h = x_{n+1} - x_n$$

where

$$f(x, y) = f_{n+s} = f_{n+1} + (s-1)\nabla f_{n+1} + \frac{(s-1)s}{2}\nabla^2 f_{n+1} + \cdots + \frac{(s-1)(s)(s+1)\cdots(s+N-2)}{N!}\nabla^N f_{n+1}$$

where

$$s = \left(\frac{x - x_n}{h}\right)$$

$$\nabla f_n = f_n - f_{n-1}$$

$$\nabla^2 f_n = \nabla f_n - \nabla f_{n-1} = f_n - 2f_{n-1} + f_{n-2}$$

Both sides of the differential equation are then integrated from x_n to x_{n+1}

$$\int_{x_n}^{x_{n+1}} y' dx = \int_{x_n}^{x_{n+1}} f_{n+s}\, dx$$

or

$$y_{n+1} - y_n = \int_{x_n}^{x_{n+1}} f_{n+s}\, dx$$

The polynomial expression for f_{n+s} is substituted into the right-hand side and the variable is changed from x to s.

$$y_{n+1} = y_n + \int_0^1 \left[f_{n+1} + (s-1)\nabla f_{n+1} + \frac{(s-1)s}{2}\nabla^2 f_{n+1} + \cdots + \frac{(s-1)s(s+1)\cdots(s+n-2)}{N!}\nabla^N f_{n+1}\right] ds$$

The integration yields the corrector formula

$$y_{n+1} = y_n + h\left[1 - \frac{1}{2}\nabla - \frac{1}{12}\nabla^2 - \frac{1}{24}\nabla^3 - \frac{19}{720}\nabla^4 - \frac{3}{160}\nabla^5 - \cdots\right] f_{n+1}$$

The predictor and corrector formulas, in difference forms, are

$$y_{n+1} = y_n + h\left[1 + \frac{1}{2}\nabla + \frac{5}{12}\nabla^2 + \frac{3}{8}\nabla^3 + \frac{251}{720}\nabla^4 + \frac{95}{288}\nabla^5 + \cdots\right] f_n$$

$$y_{n+1} = y_n + h\left[1 - \frac{1}{2}\nabla - \frac{1}{12}\nabla^2 - \frac{1}{24}\nabla^3 - \frac{19}{720}\nabla^4 - \frac{3}{160}\nabla^5 - \cdots\right] f_{n+1}$$

The corrector formula in ordinate form is obtained by substituting the ordinate definitions of the differences.

If differences through the third are retained and the ordinate form is introduced, the predictor and corrector formulas are:

$$y_{n+1}^{(P)} = y_n + \frac{h}{24}(55f_n - 59f_{n-1} + 37f_{n-2} - 9f_{n-3}) + \frac{251}{720}h^5 y^V(\xi_1)$$

$$y_{n+1}^{(C)} = y_n + \frac{h}{24}(9f_{n+1} + 19f_n - 5f_{n-1} + f_{n-2}) - \frac{19}{720}h^5 y^V(\xi_2)$$

2.3.4.3.3 Truncation Error. Hildebrand shows an estimate of the truncation error incurred in the step x_n to x_{n+1} to be:

$$E = \frac{1}{C}\left| y_{n+1}^{(C)} - y_{n+1}^{(P)} \right|$$

where $C = \frac{270}{19} \approx 14$ when third differences are retained

$C = \frac{502}{27} \approx 18$ when fourth differences are retained

The motivation for applying the corrector formula is that its coefficients multiplying the derivative values and step size-derivative factor in the truncation error expression are smaller than the corresponding coefficients in the associated predictor formula so that a greater accuracy is expected.

It is to be noted that the smallness of the residual formed by successive iterates of the corrector formula is a measure of only how well the difference equation employed in the approximate integration is satisfied, and in no way is a measure of difference between the iterate and the true solution of the differential equation.

It is probably for this reason that most applications of predictor-corrector type formulas apply the corrector only once per integration step.

2.3.4.3.4 Computational Procedures. The Adams-Moulton method requires starting values commensurate with the order of the formula being used. These values may be obtained by any procedure whatsoever, but are usually obtained by a Runge-Kutta method.

Once the starting values have been obtained, the predictor formula is applied and a new function value is determined. The corresponding derivative value is obtained by evaluating the differential equation. The previous values and this first estimate of the derivative are then substituted into the corrector formula, and a second or corrected function value is determined. The corresponding derivative value is then obtained and can be resubstituted into the corrector formula. This iteration of the corrector formula may be continued until there is no significant change in successive iterates. However, it is usual to apply the corrector only once per integration step. In any event, the local truncation error is proportional to the difference between the predicted and final corrected value, and provides an effective basis for variable step size integration.

The step size is diminished whenever the truncation error tolerance is exceeded, and increased when the truncation error is well within the specified tolerance. When the specified truncation error tolerance is satisfied but indicates no excess of significant figures, the process continues with the step size unchanged.

The procedure for interpolation of midpoints is available for obtaining a sequence of points at half the step size. Whereas forming a partition utilizing alternate values previously computed provide the necessary data for doubling the step size. The Runge-Kutta procedure is often used to generate a halved or doubled sequence of values necessary to insure an optimum step size throughout the computation. Since the procedure of interpolation or the selection of alternate values yields the same or better accuracy with far less computation, this use of Runge-Kutta is not recommended. The problem of obtaining output or function values corresponding to a non-integrated point of the interval is also optimally determined by interpolation.

2.3.4.3.5 Advantages and Disadvantages. Advantages of the Adams-Moulton method are:

1) Potentially more accurate than Adams' method.
2) An estimate of the truncation error is obtained.

Disadvantages are:

1) Requires a special starting/restarting procedure.
2) Requires at least two derivative evaluations per integration step.

2.3.4.4 Lanczos Method

2.3.4.4.1 Discussion. Most multi-step predictor methods make use of the function value at one previous point and the derivative values at several previous points (most commonly at 5). Lanczos (Reference 2.11), however, has derived a family of predictor formulas that make use of both function values and derivative values at the previous points. This form of solution allows development of what Lanczos calls "extrapolation of maximum efficiency" and "extrapolations of minimum round-off."

The Lanczos method, unlike some other methods (e.g., Adams-Moulton), does not have a long operational history from which to draw experience. Therefore, evaluation for trajectory problems is not complete.

2.3.4.4.2 Derivation of Equations. Generally, multi-step predictor methods take the following form

$$y_m = \sum_{k=1}^{m} \alpha_k y_{m-k} + h \sum_{k=1}^{m} \beta_k y'_{m-k}$$

However, Lanczos, in order to be completely general and include all possibilities, considered this problem in the light of Lagrangian interpolation in which every point of interpolation is a double point consisting of function and derivative. That is, instead of matching the interpolating polynomial at function values only, Lanczos matched both functional values and derivatives. In this way, he obtained

$$y_m = \sum_{k=1}^{m} (\alpha_k y_{m-k} + h \beta_k y'_{m-k}) + \frac{(m!)^2}{(2m)!} h^{2m} y^{(2m)}(\bar{x}) \qquad (x - mh < \bar{x} < x)$$

where

$$\alpha_k = \binom{m}{k}^2 \left[1 - 2k \left\{ \frac{1}{1-k} + \frac{1}{2-k} + \cdots + \frac{1}{m-k} \right\} \right]$$

and

$$\beta_k = \binom{m}{k}^2 k$$

and $\binom{m}{k}$ is the binomial coefficient

$$\binom{m}{k} = \frac{m!(m-k)!}{k!}$$

Identifying m with 1, 2, 3, 4 he obtains the following sequence of formulas called "extrapolations of maximum efficiency."

$m = 1$

$$y_1 = y_0 + h y'_0 + \frac{h^2}{2} y''(\bar{x})$$

$m = 2$

$$y_2 = -4y_1 + 5y_0 + h(4y'_1 + 2y'_0) + \frac{h^4}{6} y'''(\bar{x})$$

$m = 3$

$$y_3 = -18y_2 + 9y_1 + 10y_0 + h(9y'_2 + 18y'_1 + 3y'_0) + \frac{h^6}{20} y^{(6)}(\bar{x})$$

$m = 4$

$$y_4 = -\frac{128}{3} y_3 - 36 y_2 + 64 y_1 + \frac{47}{3} y_0 + h(16y'_3 + 72y'_2 + 48y'_1 + 4y'_0) + \frac{h^8}{70} y^{(8)}(\bar{x})$$

Inspection shows that the y ordinates are multiplied by factors larger than 1; therefore, whenever these formula are used in a repetitive fashion, small errors from a previous step will be subject to rapid amplification and quickly dominate the solution. However, the fact that these formulas are unstable does not preclude their use in special situations such as in starting/re-starting procedures.

Noting this feature of his solution, Lanczos then abandoned the concept of matching both function values and the derivatives at every point in favor of an approach which attempts to match only the derivatives at each point. As a portion of this process, he imposed conditions necessary for numerical stability by letting

$$\alpha_1 = \alpha_2 = \ldots = \alpha_m = \frac{1}{m}$$

and showed that the function contribution of the formula to be the arithmetic mean

$$\frac{1}{m}(y_0 + y_1 + \ldots y_{m-1})$$

minimized the effect of rounding errors (the best statistical averaging of random errors is obtained by the arithmetic mean). Another, more practiced, reason for choosing the α small is that β_i are given a greater role, and since these coefficients are not effected to the same degree with noise (the derivatives are determined by the differential equations) the accuracy of the solution is expected to be improved.

Subject to these conditions, the general predictor becomes:

$$y_m = \frac{1}{m}(y_0 + y_1 + \ldots y_{m-1}) + h\sum_{k=1}^{m} \beta_k y'_{m-k}$$

The remaining coefficients β_k can be uniquely determined by the method of undetermined factors. This family is called "extrapolations of minimum round-off" and follows for m=1 to m=5. $\underline{m=1}$

$$y_1 = M_1 + hy'_0 + \frac{h^2}{2}y''$$

$\underline{m=2}$

$$y_2 = M_2 + \frac{h}{4}(7y'_1 - y'_0) + \frac{3}{8}h^3y'''$$

$\underline{m=3}$

$$y_3 = M_3 + \frac{h}{6}(13y'_2 - 4y'_1 + 3y'_0) + \frac{13}{36}h^4y''''$$

$\underline{m=4}$

$$y_4 = M_4 + \frac{h}{48}(123y'_3 - 79y'_2 + 89y'_1 - 13y'_0) + \frac{95}{288}h^5y^{(5)}$$

$\underline{m=5}$

$$y_5 = M_5 + \frac{h}{240}(697y'_4 - 686y'_3 + 936y'_2 - 322y'_1 + 95y'_0) + \frac{461}{1440}h^6y^{(6)}$$

where

$$M_m = \frac{1}{m}(y_0 + y_1 + \ldots + y_{m-1})$$

2.3.4.4.3 Truncation Error. After applying a predictor formula to obtain y_m from a given set of y_{m-1} and y'_{m-1} data, an estimate of the truncation error is necessary to ensure the specified accuracy of the solution. Since the higher derivatives are not easily obtainable, an alternate approach would be to use the same set of y_{m-1} and y'_{m-1} data and extrapolate to y_{m+1}. This is a reference value only and will be denoted by $\bar{y}_{m+1}$. The difference (Δ) between the reference value $\bar{y}_{m+1}$ and y_{m+1} obtained from the predictor serves to estimate the truncation error (n) and provide as effective criterion for variable step size integration. This truncation estimate assumes $y^{(m+1)}(\bar{X})$ does not change drastically as the scheme proceeds from y_m to y_{m+1}. Lanczos lists the following reference value formulas for m=1 to m=5.

$m=1$

$$\bar{y}_2 = M_1 + 2hy'_0 + 2h^2 y''(\bar{X}) \qquad n = 0.33\Delta$$

$m=2$

$$\bar{y}_3 = M_2 + \frac{h}{4}(17y'_1 - 7y'_0) + \frac{55}{24}h^3 y'''(\bar{X}) \qquad n = 0.2\Delta$$

$m=3$

$$\bar{y}_4 = M_3 + \frac{h}{12}(79y'_2 - 72y'_1 + 29y'_0) + \frac{95.5}{36}h^4 y''''(\bar{X}) \qquad n = 0.16\Delta$$

$m=4$

$$\bar{y}_5 = M_4 + \frac{h}{48}(445y'_3 - 665y'_2 + 511y'_1 - 123y'_0) + \frac{4277}{1440}h^5 y^{(5)}(X) \qquad n = 0.125\Delta$$

$m=5$

$$\bar{y}_6 = M_5 + \frac{h}{360}(4411y'_4 - 9226y'_3 + 10272y'_2 - 5110y'_1 + 1093y'_0)$$

$$+ \frac{4738}{1440}h^6 y^{(6)}(\bar{X}) \qquad n = 0.11\Delta$$

2.3.4.4.4 Step Size Modification. As in previous methods, doubling the step size is straightforward since existing function values can be utilized if they are stored. When the step size is halved, Lanczos lists the following formulas in matrix notation for midpoint interpolation. The notation y_{01} refers to the midpoint value of y halfway between y_0 and y_1, similarly y_{12} to the midpoint value of y halfway between y_1 and y_2, and so on. It is tacitly understood that the arithmetic-mean $\bar{M}$ is added to the tabular products

m=2

	hy'_0	hy'_1	
y_{01} =	1	-1	÷8

m=3

	hy'_0	hy'_1	hy'_2	
y_{01}=	2	-11	-3	
				÷ 24
y_{12}=	3	11	-2	

m=4

	hy'_0	hy'_1	hy'_2	hy'_3	
y_{01}=	15	-205	-163	-31	
y_{12}=	31	147	-147	-31	÷ 384
y_{23}=	31	163	205	-15	

m=5

	hy'_0	hy'_1	hy'_2	hy'_3	hy'_4	
y_{01}=	348	-7518	-5454	-3930	-726	
y_{12}=	794	3178	-5178	-3794	-760	÷ 11520
y_{23}=	760	3794	5178	-3178	-794	
y_{34}=	726	3930	5454	7518	-348	

2.3.4.4.5 Computational Procedure. The initial conditions and the m equally spaced function and derivative values which the starting procedure provides, comprose m+1 equidistant function and derivative values necessary for evaluation of the predictor and reference value formulas.

The predictor formula is applied to the first m function and derivative values. This result is compared with the m+1 function value. If the specified accuracy is not met, a new starting sequence of m+1 function and derivative values must be obtained. If the specified accuracy is met, the reference value for the next integration step is obtained by evaluating the corresponding reference value formula again using the first m function and derivative values.

The successful completion of one integration step then yields a new function value from which the derivative may be evaluated, and a reference function value to be used for comparison at the completion of the next integration step.

Diminishing the integration step size when the truncation error exceeds the specified tolerance is accomplished by applying the midpoint interpolation formulas, thereby obtaining a partition at half the step size.

Increasing the integration step size when the truncation error is well within the specified tolerance is accomplished by deleting every other function and derivative value, thereby obtaining a partition at twice the step size.

2.3.4.4.6 Advantages and Disadvantages.

Advantages

1) Requires only one derivative evaluation per integration step.
2) Provides an estimate of the truncation error.
3) Makes available a matrix of local function and derivative history applicable to interpolation for the purpose of satisfying constraints or printed output without recourse to integration.

Disadvantages

1) Requires a special starting/restarting procedure.
2) Convenient step-size modification is limited to halving or doubling.

2.3.4.5 Nordsieck Method

2.3.4.5.1 Discussion. Nordsieck, by noting that all numerical methods of integration are equivalent to obtaining an approximating polynomial for $y(x)$, derives a family of methods that is based on the higher derivatives of the approximating polynomial rather than the usual past function and derivative values. The main advantage of such an approach is that the procedure is nearly interval-independent since the higher derivatives at the current point specify the same polynomial regardless of the interval size. Care should be exercised, however, in the use of the method because the effect of noise inherent in the calculation of the higher order differences which are matched by the approximating polynomial has not been resolved, at least in the literature reviewed for this monograph. Thus, while the method is theoretically accurate, practical restrictions in the mechanization of the solution may arise.

2.3.4.5.2 Derivation of Equations. Nordsieck first approximated the function $y(x)$ by a polynomial, denoted here by $P_5(x)$ (indicating fifth order). He then considered the Taylor series expansion

$$y(x+h) = y(x) + h\left\{f(x, y(x)) + \frac{h}{2!} P_5''(x) + \frac{h^2}{3!} P_5'''(x) + \frac{h^3}{4!} P_5'''(x) + \frac{h^4}{5!} P_5''''(x)\right\}$$

where for the fifth degree approximating polynomial $P_5(x) = y(x)$, $P_5'(x) = f(x, y(x))$.

The quantities

$$a(x) = \frac{h}{2!} P_5''(x) \qquad b(x) = \frac{h^2}{3!} P'''(x)$$

$$c(x) = \frac{h^3}{4!} P_5''''(x) \qquad d(x) = \frac{h^4}{5!} P'''''(x)$$

are chosen for a parameter set as they are reasonably controllable due to the step size h and their relation to a change of interval being simply a multiplication by a constant.

Suppose that the step size is changed from h to α h then the parameters a, b, c and d are modified as follows:

$$a_{\alpha h} = \alpha a_h$$

$$b_{\alpha h} = \alpha^2 b_h$$

$$c_{\alpha h} = \alpha^3 c_h$$

$$d_{\alpha h} = \alpha^4 d^h$$

The Taylor series expansion terminated at h^b then yields

$$y(x+h) = y(x) + h\left\{f[x, y(x)] + a(x) + b(x) + c(x) + d(x) + \tfrac{1}{6}\left[f(x+h, y(x+h)) - f^P\right]\right\}$$

where $f^P \equiv f[x, y(x)] + 2a(x) + 3b(x) + 4c(x) + 5d(x)$ (the predicted value of $f[x+h, y(x+h)]$)

A Taylor series expansion can also be written for $f[x, y(x)]$ in terms of $a(x)$, $b(x)$, $c(x)$ and $d(x)$

$$f[x+h, y(x+h)] = f[x, y(x)] + 2a(x) + 3b(x) + 4c(x) + 5d(x) + 1\left[f(x+h, y(x+h)) - f^P\right]$$

Similarly, expansions can be written for $a(x)$, $b(x)$, $c(x)$ and $d(x)$

$$a(x+h) = a(x) + 3b(x) + 6c(x) + 10d(x) + \frac{15}{6}\left[f(x+h, y(x+h)) - f^P\right]$$

$$b(x+h) = b(x) + 4c(x) + 10d(x) + \frac{20}{6}\left[f(x+h, y(x+h)) - f^P\right]$$

$$c(x+h) = c(x) + 5d(x) + \frac{15}{6}\left[f(x+h, y(x+h)) - f^P\right]$$

$$d(x+h) = d(x) + 1\left[f(x+h, y(x+h)) - f^P\right]$$

Nordsieck's analysis shows that the replacement of the six coefficients 1/6, 1, 15/6, 20/6, 15/6, 1 multiplying the [ ] term by the new coefficients 95/288, 1, 25/24, 35/72, 5/48, 1/120 provide an essentially optimum choice with respect to stability and minimum degradation of accuracy. This choice of coefficients then leads to a set of working equations, which Nordsieck recommends as the best choice of the possible lower and higher order versions of the method for a large-scale digital computer. It is to be noted that his experimentation and conclusions are dependent upon "fixed-point" (i.e., a method of calculation in which the operator must keep track of the decimal point as with a slide rule) arithmetic procedures; and, therefore, the relative merits of a "floating-point" (a method of calculation which automatically accounts for the location of the decimal point) version remain conjectural. However, it would appear that double precision for all computations except for derivative evaluations and convergence tests would provide the same results. This extended precision would undoubtedly be necessary in any event if high accuracy over a large integration interval is to be maintained.

Working equations corresponding to a fifth degree approximating polynomial for y(x) with corresponding truncation error estimate are:

$$y(x+h) = y(x) + h\left\{f[x, y(x)] + a(x) + b(x) + c(x) + d(x) + \frac{95}{288}\left[f[x+h, y(x+h)] - f^P\right]\right\}$$

$$f^P = f[x, y(x)] + 2a(x) + 3b(x) + 4c(x) + 5d(x)$$

$$a(x+h) = a(x) + 3b(x) + 6c(x) + 10d(x) + \frac{25}{24}\left\{f[x+h, y(x+h)] - f^P\right\}$$

$$b(x+h) = b(x) + 4c(x) + 10d(x) + \frac{35}{72}\left\{f[x+h, y(x+h)] - f^P\right\}$$

$$c(x+h) = c(x) + 5d(x) + \frac{5}{48}\left\{f[x+h, y(x+h)] - f^P\right\}$$

$$d(x+h) = d(x) + \frac{1}{120}\left\{f[x+h, y(x+h)] - f^P\right\}$$

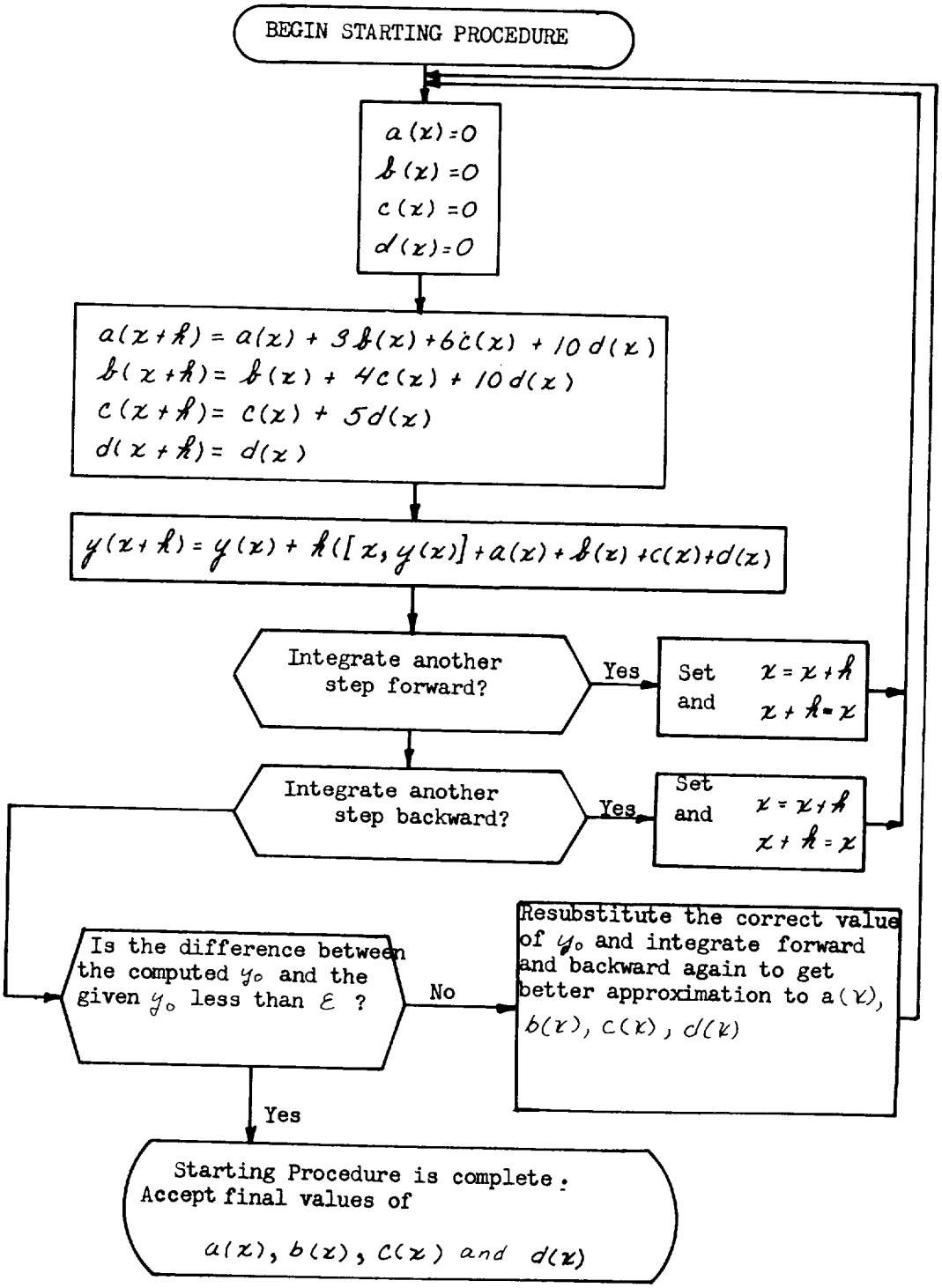
$$E_T = (863/12)\frac{h^7}{7!}y^{VII}$$

2.3.4.5.3 Starting. The starting problem, namely to determine y, f, a, b, c, d, at $X + h$ given only the initial conditions and the differential equation is a rather intricate procedure. The essential idea is to assume abnormal values of a, b, c, d (namely zero), in absence of any estimate as to their normal initial values; and to assume that after integrating a few steps that they will have approximately their normal values. The process Nordsieck outlines consists of integrating forward several steps then reversing the

sense of integration and integrating backwards, thereby arriving at the starting point with y (x) somewhat in error but with first approximations to a, b, c, and d. The correct value of y (x) (as specified by the initial conditions) is reinserted; then the process of integrating forward several steps, reversing the sense of integration, and integrating backwards is repeated. This process is continued until the computed value of y_o converges to the given value of y_o. The computational logic for starting is illustrated in the accompanying flow chart.

Figure 3.3

LOGIC FOR THE STARTING PROCEDURE IN NORDSIECK'S METHOD



2.3.4.5.4 Computational Procedure. The procedure for continuing the integration process, i.e., the solution of the working equations listed previously is not unique and should be performed as follows: A first-predicted value $y^{(1)}(x+h)$ is obtained by omitting the bracketed term, then a first approximation to the correct derivative is obtained by $f^{(1)}(x+h)=f\left[x+h, y^{(1)}(x+h)\right]$. The corrected value $y^{(2)}(x+h)$ is then obtained by including the bracketed term. The second and final approximation to the correct derivative value is obtained by $f\left[x+h, y^{(2)}(x+h)\right]$. The final corrected value $y^{(3)}(x+h)$ is then obtained by again including the bracketed term (see flow chart). Each cycle of the integration process includes the tests

$$\left| y_i^{(3)} - y_i^{(2)} \right|_{max} \leqq \frac{1}{\beta} \left| y_i^{(2)} - y_i^{(1)} \right|_{max}$$

and

$$\left| f_i(k+h) - f_i^{p} \right|_{max} \leqq \beta^{-e} / |h|$$

where e is a specifiable positive integer and "max" means the largest value obtained from the system of differential equations. The first test insures stability, in that the iteration error is overshadowed by the truncation error. The second test bounds the truncation error for a single integration step. The context of e here is then the local preservation of the e^{th} digit in the solution.

The integration step size selected is then the largest value that satisfies the above tests. Nordsieck includes four minor modifications of these tests in order to improve the usefulness and efficiency of the method and the smoothness of automatic interval control.

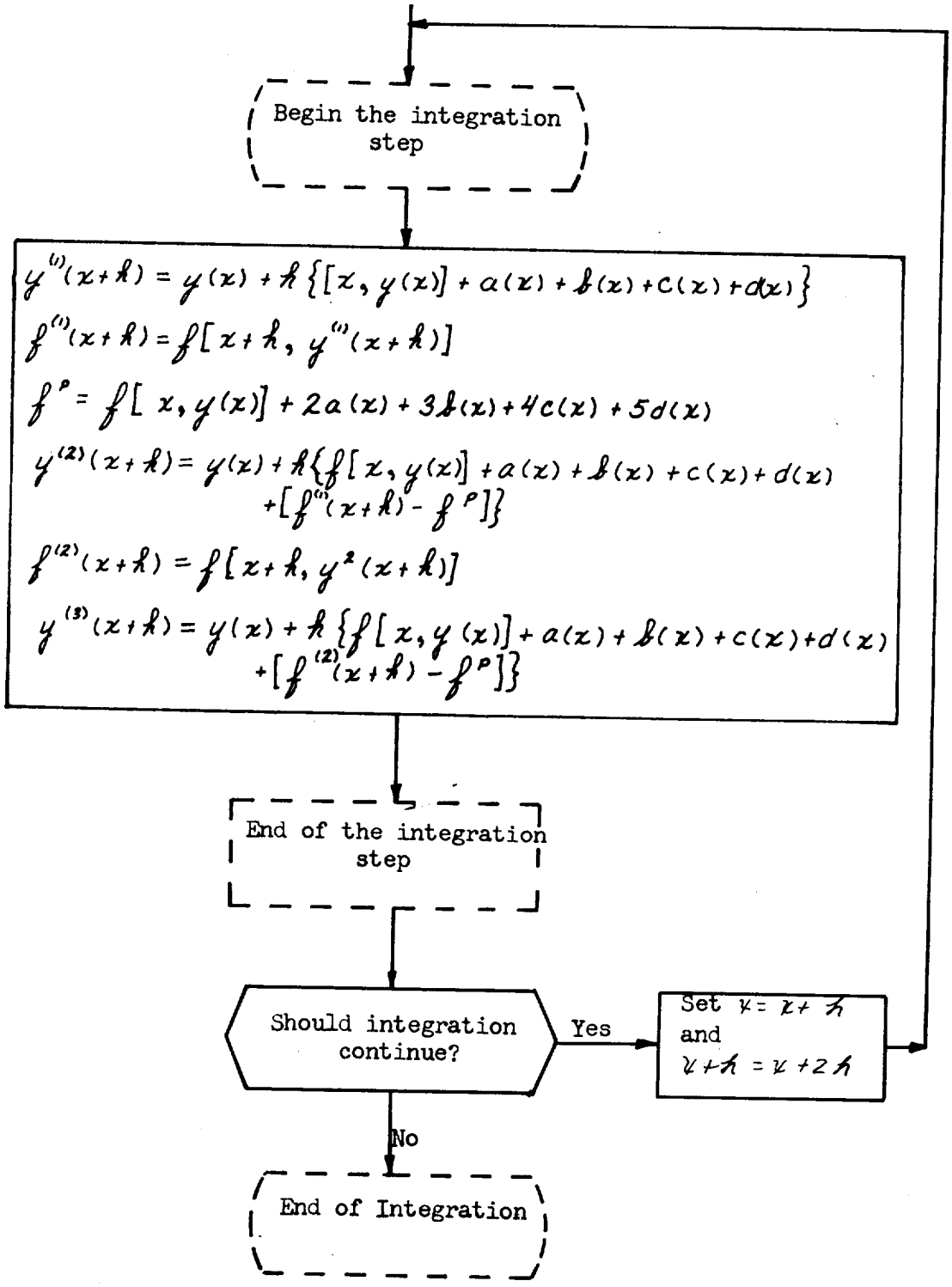
2.3.4.5.5 Advantages and Disadvantages. Advantages of the Nordsieck method are:

1. Step size modification is readily accomplished.
2. An estimate of the truncation error is obtained.
3. Step size is selected so as to insure stability.
4. Capable of high accuracy for a large number of integration steps.

Disadvantages are:

1. Requires a special starting/restarting procedure.
2. Requires two derivative evaluations per integration step.
3. The desirability of this method, when used with fixed-word length computers, is questionable since the higher orders differences are inherently noisy.

Table 3.4. Logic for the Continuation of Nordsieck's Integration



2.3.4.6 δ^2 Method

2.3.4.6.1 Discussion. The δ^2 method is developed by doubly integrating an interpolation formula involving finite differences (the finite difference operator is δ). The method is of interest because it provides a direct method for integrating second-order differential equations without reducing them to a system of first-order equations. If the force is conservative, as is the case with many astronomical and physical problems, then the first derivative need not be evaluated as part of a step by step integrating process; and the double integration performed directly by the δ^2 equations is simpler and faster than two single integrations by a δ formula (single integration of the interpolation formula) or related methods such as Adams-Bashforth. If the accelerations are a function of velocity, then it will be necessary to use the δ equations to obtain this quantity. However, it is still desirable to obtain position by the δ^2 formula rather than by two applications of the δ formula. (Reference 2.24).

2.3.4.6.2 Derivation of Equations. The derivation of δ^2 formulas begins by approximating the second-order function being integrated with a polynomial and by integrating the result twice to give the desired equations. The approximating polynomial for the second derivative is derived from formulas for interpolation or extrapolation. For this section, the δ^2 equation, which results from the use of Stirling central difference interpolation formula, is developed. It should be noted that if a different interpolation formula is used to approximate y", a different but equally valid δ^2 equation will result; for example, Scarborough (Reference 2.10) developed δ^2 equations using Newton's formula for backward interpolation.

Stirling's central difference formula can be found in any standard text on numerical analysis and is simply reproduced below

$$y'' = y_n'' + s\mu\delta y_n'' + \frac{s^2}{2}\delta^2 y_n'' + \ldots \tag{3.43}$$

where $s = \dfrac{x - x_n}{h}$

δ = central difference operator $(\delta x_n \triangleq x_{n+1/2} - x_{n-1/2})$

μ = averaging operator $[\mu x_n = 1/2(x_{n+1/2} + x_{n-1/2})]$

Integration of equation (3.43) with respect to x yields

$$y' = h\left(y_n''s + \frac{s^2}{2}\mu\delta y_n'' + \frac{s^3}{6}\delta^2 y_n'' + \ldots\right) + C$$

At this point, the value of the constant C is determined by substituting $x = x_n$ in equation (3.43). If $x = x_n$, then $y = y_n$ and the constant is seen to be $c = y_n'$. Thus, if C is replaced by y_n' and another integration performed, the result is

$$y = hs\,y_n' + h^2\left(\frac{s^2 y_n''}{2}\right) + \frac{s^3}{6}\mu\,\delta y_n'' + \frac{s^4}{24}\delta^2 y_n'' + \ldots) + c'$$

The constant c' is now evaluated by requiring that $y = y_n$ at $x = x_n$. The result is $c' = y_n$.

$$y = y_n + hs y_n' + h^2\left(\frac{s^2 y_0''}{2} + \frac{s^3}{6}\mu\,\delta y_n'' + \frac{s^4}{24}\delta^2 y_0'' + \ldots\right)$$

Values of y at $y = y_{n-1}$ and $y = y_{n+1}$ are now determined by evaluating the equation at the points $x = x_{n-1}$ and $x = x_{n+1}$ (i.e., at $x_{n+1} = x_n + h$).

$$y_{n-1} = y_n - h y_n' + h^2\left(\frac{y_n''}{2} - \frac{\mu}{6}\delta y_n'' + \frac{\delta^2 y_n''}{24} - \ldots\right) \tag{3.44}$$

$$y_{n+1} = y_n + h y_n' + h^2\left(\frac{y_n''}{2} + \frac{\mu\delta y_n''}{6} + \frac{1}{24}\delta^2 y_n + \ldots\right) \tag{3.45}$$

Addition of equations (3.44) and (3.45) now gives

$$y_{n+1} + y_{n-1} - 2y_n = h^2\left(y_n'' + \frac{1}{12}\delta^2 y_n - \frac{1}{240}\delta^4 y_n'' + \frac{31}{60480}\delta^6 y'' + \ldots\right) \tag{3.46}$$

Since $\delta^2 x_n = y_{n+1} - 2y_n + y_{n-1}$, equation (3.46) can be written in its final form

$$\delta^2 x_n = h^2\left(y_n'' + \frac{1}{12}\delta^2 y_n'' - \frac{1}{240}\delta^4 y_n'' + \frac{31}{60480}\delta^6 y_n'' + \ldots\right) \tag{3.47}$$

2.3.4.6.3 Truncation Error. In order to obtain a solution for y_{n+1} by the use of the equations derived in the last section, it is necessary to know the values of higher order differences such as $\delta^2 \ddot{y}_n$, $\delta^4 \ddot{y}_n$, etc. Since these values cannot be calculated directly without $\ddot{y}_{n+1}$, it is necessary to approximate y and $\dot{y}$ at the next epoch by a truncated series and subsequently, estimate the desired derivatives. With these approximate values of $\delta\ddot{y}_n$ etc., a corrected valculation for y_{n+1} can be made. This corrected estimate can then be used to obtain a more accurate estimate of the higher differences needed in the equation for y_{n+1} and a new value for y_{n+1} can be found if desired. The difference between the preliminary values of y_{n+1} and the corrected value will serve as an estimate of the truncation error. It should be noted that the actual truncation error cannot be obtained by this technique; however, it provides a criterion for selecting the proper step size.

2.3.4.6.4 Step Size Modification. Any change in step size which increases the previous step size by an integer amount is easily handled by simple deleting the unwanted solution values provided sufficient preceding solution points have been retained. In general, unless all previous solution points have been retained, the number of previous solution points available will be less than

the minimum number required to increase the step size. In such a case, the solution simply proceeds with the current increment until the minimum number required to change the step is reached. If the step size is to be decreased, solution points must be generated within the intervals of the previous steps. Generation of these points is easily accomplished by the use of an interpolation formula. The accuracy of the interpolated points can be made the same as the original solution points by choosing a suitably high-order interpolation formula. A selection of interpolation formulas suitable for this purpose can be found in Hildebrand. (Reference 2.12).

2.3.4.6.5 Computational Procedure. The use of the δ^2 formula will be illustrated by performing one step in the solution of the equation

$$\ddot{x} = -x \tag{3.48}$$

This example is reproduced by consent of Dr. S. Herrick (Reference 2.25). It is assumed that a starting procedure is available which has generated the data in Table 1 which corresponds to the beginning of n = 9 step. The values in parentheses are part of the calculation of the step n = 9 and are described below. The last values in the several columns (except those in parentheses) are x_8, $\delta x_{7\frac{1}{2}}$, $\delta^2 x_7$, $\delta \ddot{x}_{7\frac{1}{2}}$, $\delta^2 \ddot{x}_7$, $\delta^3 \ddot{x}_{7\frac{1}{2}}$, $\delta^4 \ddot{x}_6$, and $\delta^5 \ddot{x}_{5\frac{1}{2}}$; those values are correct within the limits of rounding error. The first step in the calculation is to estimate $\delta^2 \ddot{x}_8$. This estimate can be obtained from the equation

$$\delta^2 x_n = \delta^2 x_{n-1} + \delta^3 x_{n-1\frac{1}{2}} + \delta^4 x_{n-2} + \delta^5 x_{n-2\frac{1}{2}} \tag{3.49}$$

This equation is obtained by assuming that $\delta^6 x_{n-2}$ is zero and working backward by substituting the definition of δ^6 etc. A similar expression for $\delta^4 \ddot{x}_n$ is

$$\delta^4 \ddot{x}_n = \delta^4 x_{n-2} + 2\delta^5 \ddot{x}_{n-2\frac{1}{2}} \tag{3.50}$$

Using equation (3.49) and equation (3.50) and the data in Table 3.4, there is obtained

$$\left.\begin{aligned} \delta^2 \ddot{x}_8 &= 64368 + 7950 - 566 - 89 \\ &= 71663 \\ \delta^4 \ddot{x}_8 &= -566 - 178 \\ &= -744 \end{aligned}\right\} \tag{3.51}$$

The values from equation (3.51) are now used in the δ^2 equation developed in Section 2.3.4.5.3, i.e.,

$$\delta^2 x_n = h^2 (y_n'' + \tfrac{1}{12}\delta^2 y_n'' - \tfrac{1}{240}\delta^4 y_n'')$$

So

Table 3.4. Difference Table for Sample Problem $\ddot{x} = -x$

τ	x	δx	$\delta^2 x$	$\ddot{x} = -x$	$\delta\ddot{x}$	$\delta^2\ddot{x}$	$\delta^3\ddot{x}$	$\delta^4\ddot{x}$	$\delta^5\ddot{x}$	$\varepsilon = x - \sin\tau$
0.0	0.000,000,000		0,000,000	0.000,0000		0,0000		000		0 0
		+99,833,417			-99,8334		+9975		-100	0
0.1	+0.099,833,417		-0,997,502	-0.099,8334		+0,9975		-100		0 +1
		+98,835,915			-98,8359		+9875		-097	+1
0.2	+0.198,669,332		-1,985,038	-0.198,6693		+1,9850		-197		+1 0
		+96,850,877			-96,8509		+9678		-100	+1
0.3	+0.295,520,209		-2,952,740	-0.295,5202		+2,9528		-297		+2 +1
		+93,898,137			-93,8981		+9381		-091	+2
0.4	+0.389,418,346		-3,890,939	-0.389,4183		+3.8909		-388		+4 -1
		+90,007,198			-90,0072		+8993		-089	+1
0.5	+0.479,425,544		-4,790,261	-0.479,4255		+4,7902		-477		+5 +2
		+85,216,937			-85,2170		+8516		-089	+3
0.6	+0.564,642,481		-5,641,721	-0.564,6425		+5,6418		-566		+8 -1
		+79,575,216			-79,5752		+7950			+2
0.7	+0.644,217,697		-6,436,810	-0.644,2177		+6,4368		-744		+10 0
		+73,138,406			-73,1384					+2
0.8	+0.717,356,103		(-7,167,586)	-0.717,3561		7,1663				+12
		(+65,970,820)								
0.9	(+0.783,326,923)									

$$\left.\begin{aligned} \delta^2 X_8 &= .01(-.7173561 + .005971\,|\,9 + .0000003\,|\,1) \\ &= -.007167586 \end{aligned}\right\} \qquad (3.52)$$

The vertical lines in equation (3.51) are drawn to set off the digits that are carried only to reduce the accumulation of round-off error. This value, from equation (3.51) is entered in Table 1 in parenthesis and is used to calculate X_9 from the relation $\delta^2 x_n = x_{n+1} - 2x_n + x_{n-1}$.

$$X_9 = .78332\,6923 \qquad (3.53)$$

If it is desired to correct this value of X_9, the procedure would be to obtain $\ddot{X}_9$ from the equation (3.48) using equation (3.53) as the value of X_9. The differences $\delta^2 \ddot{x}_9$ and $\delta^4 \ddot{x}_9$ would then be determined by differencing. The estimated differences are then used in equation (3.47) to predict a value of X_9 which may be compared with the value obtained in equation (3.53) to determine if any significant change has been made.

2.3.4.6.6 Advantages and Disadvantages. One advantage of the δ^2 is its relative speed in solving a second-order differential equation as compared to methods which reduce the second-order differential equations to a system of first-order equations and perform two single integrations. Further, it will be more accurate than such methods because there will be truncation error for <u>each</u> of the variables in the system of first-order equations. One disadvantage is that, like other multi-step methods, a separate starting procedure is required. The method is not as desirable as the Gauss-Jackson method, which also directly integrates a second-order differential equation. (Gauss-Jackson is discussed in the next section.

2.3.4.7 Σ^2 (Gauss-Jackson) Method

2.3.4.7.1 Discussion. The Σ^2 or Gauss-Jackson method is, like the δ^2 method from which it is derived, another method which directly integrates a second-order differential equation. (Σ is the inverse operation for δ). As was the case with the δ^2 method, double integration by the Σ^2 formula is simpler and faster than two single integrations. The Σ^2 method provides a degree of smoothing to the δ^2 equations by effectively shifting the coefficients of the higher order differences so that errors in these quantities have less effect on the value of the solution point. [Compare equation (3.47) with equation (3.55)]. Thus, the Σ^2 method converges better than the δ^2 method and for this reason is generally preferrable. The remark made for the δ procedure that the solution for position should be made independently of the solution for velocity (i.e., a single calculation for position by the Σ^2 formula rather than repetive single integrations) is equally valid for the Σ^2 method.

2.3.4.7.2 Derivation of Equations. In the derivation presented here, the Σ^2 formula is obtained from the δ^2 formula of Section 2.2.4.5.3 by considering δ to be an operator whose inverse is Σ. (The Σ operator bears the same

relation to the integral operator as the δ operator bears to the differential operator.) Consider the operator δ transforms X_n according to the equation

$$\delta X_n = X_{n+\frac{1}{2}} - X_{n-\frac{1}{2}} \tag{3.54}$$

Thus, if Σ is the inverse of δ, the following equation must hold

$$\Sigma(\delta X_n) = X_n$$

Substituting equation (3.54) for δX_n now gives

$$\Sigma X_{n+\frac{1}{2}} - \Sigma X_{n-\frac{1}{2}} = X_n$$

or

$$\Sigma X_n = X_{n-\frac{1}{2}} + \Sigma X_{n-1} \tag{3.55}$$

If the set of numbers under consideration begins with X_o, then it will be necessary to arbitrarily define ΣX_o. Such an arbitrary definition should not seem strange in view of the relation between the Σ operator and the integral operator which it is recalled that an arbitrary constant, dependent upon initial conditions, is required for integration.

Now consider the δ^2 formula from Section 2.3.4.5.2

$$\delta^2 y_i = h^2\left(y''_i + \tfrac{1}{12}\delta^2 y''_i - \tfrac{1}{240}\delta^4 y''_i + \tfrac{31}{60480}\delta^6 y''_i - \frac{289}{3628800}\delta^8 y'' + \ldots\right)$$

Applying the Σ operator twice to this equation yields the Σ^2 integration formula

$$y_i = h^2\left(\Sigma^2 y''_i + \frac{1}{12}y''_i - \frac{1}{240}\delta^2 y_i + \frac{31}{60480}\delta^4 y'' - \frac{289}{36288}\delta^6 y'' + \ldots\right) \tag{3.56}$$

An equivalent backward-difference predictor formula, the "Σ^2_b" formula, can be similarly derived and is written below

$$y_{i+1} = h^2\left(\Sigma^2 \ddot{y}_{i+1} + \frac{1}{12}\ddot{y}_i + \frac{1}{12}\delta\ddot{y}_{i-1/2} + \frac{19}{240}\delta^2\ddot{y}_{i-1} + \frac{18}{240}\delta^3\ddot{y}_{i-1\frac{1}{2}} + \frac{1726}{24192}\delta^4\ddot{y}_{i-2} + \frac{1650}{24192}\delta^5\ddot{y}_{i-2\frac{1}{2}} + \ldots\right) \tag{3.57}$$

It is significant to note in both of these equations that the coefficients of the various powers have been shifted, thereby permitting larger estimation errors to occur without penalty in the "Σ^2" method as opposed to the "δ^2" method.

Since equation (3.56) is effectively a double integration, there are two arbitrarily constants to be evaluated. A special starting procedure is

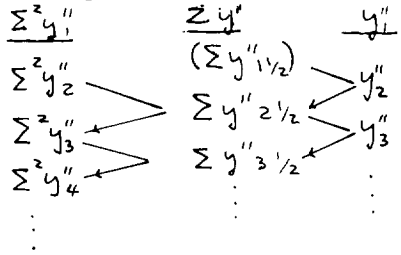
assumed to be available to generate sufficient points so that the constants can be evaluated. An equation for these constants (Σ^2_{y1} and Σ^2_{y2}) is obtained by rearranging equation (3.56)

$$\left.\begin{aligned}\Sigma^2 y_1'' &= \frac{y_1}{h_2} - \frac{1}{12} y_1'' + \frac{1}{240}\delta^2 y_1'' - \frac{31}{60480}\delta^4 y_1'' + \frac{289}{3628800}\delta^6 y'' + \dots \\ \Sigma^2 y_2'' &= \frac{y_2}{h_2} - \frac{1}{12} y_2'' + \frac{1}{240}\delta^2 y_2'' - \frac{31}{60480}\delta^4 y_2'' + \frac{289}{3628800}\delta^6 y'' + \dots\end{aligned}\right\} \quad (3.58)$$

With these initial conditions, a sum table can be constructed in the same manner that a difference table is constructed. The sum table will have three columns: a $\Sigma^2 y_n''$ column; a $\Sigma y_n''$ column; and a y_n'' column. The y_n'' column can be filled in by the use of a suitable starting procedure and the first two terms in the Σ^2 column are calculated by equation (3.58). If the Σ^2 operator is applied to equation (3.54) the result is

$$\Sigma x_n = \Sigma^2 x_{n+1/2} - \Sigma^2 x_{n-1/2} \quad (3.59)$$

This equation allows the calculation of the first member in the Σ column using $\Sigma^2 x_1$ and $\Sigma^2 x_2$. The rest of the table is then completed by repetitive application of equations (3.59) and (3.55).



Thus, for those problems in which the force field is conservative (i.e., the acceleration is a function only of the coordinates not the rates), the complete ephemeris can be generated without determining the velocity. However, while this capability is very useful in certain problems, it is not compatible with most of the problems of common interest. Thus, it is necessary to develop the first integral as well. This process is performed in exactly the same manner as was outlined for the double integral. The result is:

$$x_{i+1} = h\left[\tfrac{1}{2}\left(\Sigma \ddot{x}_{i+1/2} + \Sigma \ddot{x}_{i+1/2}\right) - \tfrac{1}{24}\left(\delta \ddot{x}_{i+1/2} + \delta \ddot{x}_{i+1/2}\right) + \frac{11}{1440}\left(\delta^3 \ddot{x}_{i+1/2} + \delta^3 \ddot{x}_{i+1/2}\right) - \dots\right] \quad (3.60)$$

Before continuing, it is noted that this equation provides a means of evaluating the second "constant" of integration based on the initial conditions uncontaminated directly by the noise in the integral at a point along the trajectory as in equation (3.59). This feature is derived by solving equation (3.60) for the averaged first sum at the epoch of the initial condition

$$\tfrac{1}{2}(\Sigma \ddot{x}_{1/2} + \Sigma \ddot{x}_{-\frac{1}{2}}) = \frac{\dot{x}_0}{h} + \frac{1}{24}(\delta \ddot{x}_{1/2} + \delta \ddot{x}_{-1/2})$$

$$- \frac{11}{1440}(\delta^3 \ddot{x}_{1/2} + \delta^3 \ddot{x}_{-1/2})$$

$$+ \frac{191}{120960}(\delta^5 \ddot{x}_{1/2} + \delta^5 \ddot{x}_{-1/2})$$

$$\equiv \Sigma \ddot{x}_0$$

and then making use of the fact that

$$\ddot{x}_0 = \Sigma \ddot{x}_{1/2} - \Sigma \ddot{x}_{-1/2}$$

$$\ddot{x}_0 + 2\Sigma \ddot{x}_{-1/2} = \Sigma \ddot{x}_{1/2} + \Sigma \ddot{x}_{-1/2}$$

$$\equiv 2\Sigma \ddot{x}_0$$

To yield

$$\Sigma \ddot{x}_{1/2} = \Sigma \ddot{x}_0 + 1/2\, \ddot{x}_0$$

Therefore, it is possible to construct the sum table without differencing two computed second sums.

2.3.4.7.3 Truncation Error. As with the δ^2 method, a measure of the truncation error may be obtained by comparing the resulting value for x_{n+1} calculated from equation (3.59) with that calculated from equation (3.60). Again, while the difference is not a pure function of truncation of the series, it does serve as an effective means of estimating the error involved in the mechanization of the solution process.

2.3.4.7.4 Step Size Modification. The discussion on step size modification made for the δ^2 procedure applies equally to the Σ^2 method that discussion is found in section 2.2.4.6.4.

2.3.4.7.5 Computational Procedure. The following discussion is reproduced, with a few modifications, from Herrick's book on astrodynamics (Reference 2.2.5) by consent of the author. The equation for solving a second order differential equation is

$$x_{i+1} = h^2\left(\Sigma^2 \ddot{x}_{i+1} + \frac{1}{12}\ddot{x}_{i+1} - \frac{1}{240}\delta^2 \ddot{x}_{i+1} + \ldots\right) \tag{3.61}$$

where $h = \tau_{i+1} - \tau_i$ is the uniform interval of the argument. The problem to be illustrated is the solution of the equation

$$\ddot{x}(\tau) = -x(\tau), \qquad x(0) = 0$$

The analytic solution is recognized to be

$$x = \sin\tau$$

Table 3.5 shows the Σ^2 integration table for the example above. It is noted that the table omits the decimal place in the differences of x since the last digit may be recognized as being in the 7th decimal place.

In the several columns of Table 1, the last values above the dashed line are known at the beginning of the integration step to be illustrated. These are $\Sigma^2\ddot{x}_{i+1}, \Sigma\ddot{x}_{i+1/2}, \ddot{x}_i, \delta\ddot{x}_{i-1/2}, \delta^2\ddot{x}_{i-1}, \delta^3\ddot{x}_{i-1/2}, \delta^4\ddot{x}_{i-2}, \delta^5\ddot{x}_{i-2\frac{1}{2}}$ where i=8, i+1=9. The first operation in the integration is the prediction of $\ddot{x}_{i+1}$ x_{i+1}. This objective can be accomplished in two distinctly different modes. The first and most accurate mode, in general, would be to step the function and its first derivative using the predictor formula; then, the equations of motion would be solved to yield the predicted acceleration. In contrast, the second mode employs the summation of the last acceleration and the differences along the trailing diagonal containing the last known acceleration (this process is equivalent to assuming that the last difference being carried is invariant to a sufficient degree) to eliminate one of the required functional evaluations. For the purpose of the sample problem, the simple expansion technique is assumed completely adequate

Table 3.5. Sum Table for Sample Problem $\ddot{x} = -x$

t	$\sum^2 \ddot{x}$	$\sum \ddot{x}$	$\ddot{x}$	$\delta\ddot{x}$	$\delta^2\ddot{x}$	$\delta^3\ddot{x}$	$\delta^4\ddot{x}$	$\delta^5\ddot{x}$
0.0	0.000,0000		-0.000,0000		0,0000		000	
		+9.991,6653		-99,8334		+9975		-100
0.1	+ 9.991,6653		-0.099,8334		+0,9975		-100	
		+9.891,8319		-98,8354		+9875		-097
0.2	+19.883,4972		-0.198,6693		+1,9850		-197	
		+9.693,1626		-96,8509		+9678		-100
0.3	+29.576,6598		-0.295,5202		+2,9528		-297	
		+9.397,6426		-93,8981		+9381		-091
0.4	+38.974,3022		-0.389,4183		+3,8909		-388	
		+9.008,2241		-90,0072		+8993		-089
0.5	+47.982,5263		-0.479,4255		+4,7902		-477	
		+8.528,7986		-85,2170		+8516		-089
0.6	+56.511,3249		-0.564,6425		+5,6418		-566	-----
		+7.964,1561		-79,5752		+7950	--------/	-076
0.7	+64.475,4810		-0.644,2177		+6,4368	--------/	-642	
		+7.319,9384		-73,1384	----------/	+7308		
0.8	+71.795,4194		-0.717,3561	----------/	+7,1676			
		+6.602,5823	----------/	-65,9708				
0.9	+78.398,0017	----------/	-0.783,3269					
-----	----------/	+5.819,2554						
1.0	+84,217,2571							

$$\ddot{x}_{i+1} = \ddot{x}_i + \delta\ddot{x}_{i-1/2} + \delta^2\ddot{x}_{i-1} + \delta^3\ddot{x}_{i-1\frac{1}{2}} + \delta^4\ddot{x}_{i-2} + \delta^5\ddot{x}_{i-2\frac{1}{2}} + \cdots$$

$$= -0.717,3561 - 73,1384 + 6,4368 + 7950 - 566 - 0^* + \cdots$$

$$= -0.783,3193^* + \cdots -.783,3282$$

and

$$\delta^2\ddot{x}_{i+1} = \delta^2\ddot{x}_{i-1} + 2\delta^3\ddot{x}_{i-1\frac{1}{2}} + 3\delta^4\ddot{x}_{i-2} + 4\delta^5\ddot{x}_{i-2\frac{1}{2}} + \cdots$$

$$= +6,4368 + 1,5900 - 1698 - 0^* + \cdots$$

$$= +7,8570^* + \cdots + 7,8214$$

*Note that the 5th difference is replaced by 0 in order to show that a large error (+76) can be introduced into the estimation process in the Σ^2 procedure without requiring a corrector cycle such as that made necessary by a smaller error in the δ^2 procedure.

By equation (3.61), using $h = \tau_{i+1} - \tau_i = 0.1$,

$$x_{i+1} = h^2\left(\Sigma^2\ddot{x}_{i+1} + \frac{1}{12}\ddot{x}_{i+1} - \frac{1}{240}\delta^2\ddot{x}_{i+1} + \cdots\right)$$

$$= 0.01\,(+78.398,00\,|\,2$$
$$-65,27\,|\,7$$
$$-\quad 3\,|\,3 + \cdots)$$
$$= 0.01\,(+78.332,69\,|\,2 + \cdots)$$
$$= +0.783,3269$$

(The vertical line is drawn to set off digits that are carried only to reduce the accumulation of rounding error). Then the value

$$\ddot{x}_{i+1} = -x_{i+1} = -0.783,3269 \qquad \text{(for this sample)}$$

is entered into Table 1 and summed and differenced to obtain the numbers in the diagonal below the dashed line. These values are correct, and the cycle is complete, as the reader may verify by comparing them with the analytic solution.

2.3.4.7.6 Advantages and Disadvantages. The main advantage of the Σ^2 method, as with the δ^2 method, is the speed with which second order differential equations involving only conservative forces, are handled. The Σ^2 method is, however, less subject to round off error problems than the δ^2 method because of the smaller coefficients for the higher order differences. Thus, the Σ^2 method generally accomplishes with a "predictor" alone what the δ^2 procedure (and other related procedures such as Adams-Bashforth) require a "predictor" and "corrector" to accomplish. However, it is generally a good

idea to provide a "corrector" for the Σ^2 method, even though it isn't always required, to assure consistency and provide a higher level of stability. The method suffers from the usual multi-step starting problems (i.e., it is not self starting).

2.3.5 STARTING PROCEDURES

2.3.5.1 Starting by Taylor Series

2.3.5.1.1 Discussion. One of the most straightforward methods of finding the matrix of starting values for those routines which require them is to determine the coefficients of a Taylor expansion

$$y_s = y(x_o + hs) = y_o + hy_o's + \frac{h^2 y_o'' s^2}{2!} + \ldots + h^n y_o^{(n)} \frac{s^n}{n!} + h^{n+1} y_o^{(n+1)}(\xi) \frac{s^{n+1}}{(n+1)!} \tag{3.62}$$

$s = \pm$ integer

$h = x - x_o$

where the last term compensates for the error and $x_o \leq \xi \leq x_o + h$. Different integer values of S give distinct values of y, and the series is evaluated for as many integer values of S as are required for starting. Generally, $y_o = y(x_o)$ will be known, and S will range from I to N, or from $-\frac{N}{2}$ to $+\frac{N}{2}$.

The problem lies in finding the derivatives of y_o ; however, by recalling that $\frac{d}{dx} = \frac{\partial}{\partial x} + y' \frac{\partial}{\partial y}$, the relations

$$y' = f(x, y)$$

$$y'' = f_x(x, y) + y' f_y(x, y)$$

$$y''' = f_{xx}(x, y) + 2y' f_{xy}(x, y) + y'^2 f_{yy}(x, y) + y'' f_y(x, y)$$

$$\vdots$$

may be obtained and thus, it follows that

$$y_o' = f(x_o, y_o)$$

$$y_o'' = f_x(x_o, y_o) + y_o' f_y(x_o, y_o)$$

and so forth for higher derivatives.

2.3.5.2.2 Remarks. Since y(x) is not generally known, it is usually difficult to obtain a reliable estimate of the truncation-error term given in equation (3.62) even though it be known that the series itself is convergent for the values of S under consideration. Further, in many cases, the equations for finding higher order derivatives will either rapidly increase in complexity or be undefinable. Thus, while Taylor series is the basis for most starting procedures, it is seldom used in its simplest form.

2.3.5.2 Starting by Modified Euler Method

2.3.5.2.1 Discussion. The first order differential equation may be integrated

$$y' = f(x, y) \tag{3.63}$$

to give y as a function of x, y=F(x). (3.64)

The graph of equation (3.64) is a curve in the x-y plane which may be defined in an approximate sense by setting $\Delta x = x - x_0$, $\Delta y = y(x) - y(x_0)$ and writing the following approximate relations

$$\Delta y \cong \left(\frac{dy}{dx}\right)_0 \Delta x$$

$$y_1 \cong y_0 + \left(\frac{dy}{dx}\right)_0 \Delta x$$

Then the values of y corresponding to $x_2 = x_1 + h$, $x_3 = x_2 + h$, ...are

$$y_2 \simeq y_1 + \left(\frac{dy}{dx}\right)_1 h$$

$$y_{n+1} = y_n + \left(\frac{dy}{dx}\right)_n h$$

Thus, if h is small enough, proceeding in this manner allows the integral of equation (3.63) to be tabulated as a set of corresponding values of x and y. However, the method is generally either too slow for h small or too inaccurate for a larger h to be of practical use. These considerations have led to a modification of Euler's method.

Starting with the initial value, y_0, and approximate value for y_1 is computed from the relation

$$y_1^{(1)} \simeq y_0 + \left(\frac{dy}{dx}\right)_0 h$$

where the superscript, (1), indicates the first approximation. This approximate value of y′ is substituted into the differential equation (3.63) to get an approximate value of dy/dx

$$\left(\frac{dy}{dx}\right)_1^{(1)} = f\left[x_1, y_1^{(1)}\right]$$

An improved value of Δy is then found by multiplying h by the mean of the values of dy/dx at the ends of the interval x_0 and x_1 , or

$$\Delta y = \left[\frac{(dy/dx)_0 + (dy/dx)_1^{(1)}}{2}\right] h$$

The second approximation for y_1 is now

$$y_1^{(2)} = y_0 + \frac{(dy/dx)_0 + (dy/dx)_1^{(1)}}{2} h$$

This improved value of $y_1^{(2)}$ is substituted into equation (3.63) to get a second approximation for $(dy/dx)_1$, or $(dy/dx)_1^{(2)} = f[x, y_1^{(2)}]$. Likewise, the third approximation for y_1 is found to be

$$y_1^{(3)} = y_0 + \frac{(dy/dx)_0 + (dy/dx)_1^{(2)}}{2} h$$

and this process is repeated until no change is produced in y_1 for the number of digits retained.

The computation for the next interval x_1 to $x_2 = x_1 + h$ is similarly carried out, by first finding an approximate value of Δy and then applying the averaging process until no improvement is made in y_2.

Although the first approximations to y_2, y_3, ... could be found using

$$y_{n+1} \simeq y_n + h\left(\frac{dy}{dx}\right)_n$$

the first approximations to succeeding y's can be more accurately found from the formula

$$y_{n+1} = y_{n-1} + 2hy_n'$$

as soon as two consecutive values of y are known. This fact may be readily proven by subtracting two Taylor expansions (one for +h one for -h and observing that the term involving the second derivative cancels. Thus, this approximation is correct to the third order).

2.2.5.2.2 Remarks. The modified Euler method is slow and has limited accuracy. However, it can be recommended for applications where simplicity is more important than accuracy or for those cases where corrector cycles (employing closed integration formula) will be applied to the results once the process has converged to a satisfactory degree.

2.3.5.3 Starting by Milne's Formulas

2.3.5.3.1 Discussion. When higher derivatives of y cannot easily be found but the first and second derivatives of y can be determined without difficulty, the starting values can be found by Milne's formulas. The formulas will be derived here for the case when five starting values of y are required.

First, representing the function y' in the neighborhood of $x = x_0$ by a fourth-order Taylor series yields

$$y_1' = y_0' + y_0'' h + y_0''' \frac{h^2}{2!} + y_0^{IV} \frac{h^3}{3!} + y_0^{V} \frac{h^4}{4!} + \dots \tag{3.65}$$

$$y_{-1}' = y_0' - y_0'' h + y_0''' \frac{h^2}{2!} - y_0^{IV} \frac{h^3}{3!} + y_0^{V} \frac{h^4}{4!} + \dots \tag{3.66}$$

(The fourth order terms will not be used in the derivation and will be carried only to provide an error estimate). Now, adding (3.65) and (3.66) yield

$$y_1' + y_{-1}' = 2y_0' + y_0''' h^2 + \frac{y_0^{V} h^4}{12} + \dots \tag{3.67}$$

Similarly subtracting (3.64) from (3.65) yields

$$y_1' - y_{-1}' = 2y_0'' h + y_0^{IV} \frac{h^3}{3} + \dots \tag{3.68}$$

Now solving (3.67) for y_0''' and (3.68) for y_0^{IV}, and substituting these results into (3.65) and (3.66) provides

$$y_1 = y_0 + \frac{h}{24}\left[y_{-1}' + 16y_0' + 7y_1'\right] + y_0'' \frac{h^2}{4} - y^{V} \frac{h^5}{180} \tag{3.69}$$

$$y_{-1} = y_0 - \frac{h}{24}\left[7y_{-1}' + 16y_0' + y_1'\right] + y_0'' \frac{h^2}{4} + y_0^{V} \frac{h^5}{180} \tag{3.70}$$

Similarly, y_2 and y_{-2} are found to be

$$y_2 = y_0 + \frac{2h}{3}\left[5y_1' - y_0' - y_{-1}'\right] - 2y_0'' h^2 + \frac{7}{45} y_0^{V} h^5 \tag{3.71}$$

$$y_{-2} = y_0 - \frac{2h}{3}\left[5y_{-1}' - y_0' - y_1'\right] - 2y_0'' h^2 - \frac{7}{45} y_0^{V} h^5 \tag{3.72}$$

Milne uses additional formula in checking these values of y_2 and y_{-2}. Subtracting (3.70) from (3.69) yields

$$y_1 - y_{-1} = \frac{h}{3}\left[y_{-1}' + 4y_0' + y_1'\right] - y_0^{V} \frac{h^5}{90}$$

Since this formula holds for any interval of width 2h, a general formula may be written

$$y_{n+1} = y_{n-1} + \frac{h}{3}\left[y_{n-1}' + 4y_n' + y_{n+1}'\right] - y_n^{V} \frac{h^5}{90} \tag{3.73}$$

It can be noted that the quantity $(h/3)[y'_{n-1} + 4y'_n + y'_{n+1}]$ is evidently Simpson's Rule and is an approximation to the definite integral $\int_{-h}^{h} y'\,dx$.

In the application of these formulas, the terms in y_0^V are omitted; and hence, the formulas as used are accurate up to and including fourth differences. The formulas (3.69) - (3.73) are used in two iterations to obtain the starting values. The first iteration obtains y_1 and y_{-1} and proceeds in four steps.

(a) Trial values of y_1' and y_{-1}' are computed from Euler's method (sect.)

$$y_1' = y_0' + hy_0''$$

$$y_{-1}' = y_0' - hy_0''$$

(b) These values are substituted into equations (3.69) and (3.70) to get first approximations to y_1 and y_{-1}.

$$y' = f(x, y)$$

(c) These approximate values, with the corresponding x_1 and x_{-1} are substituted into the given differential equation

$$y' = f(x, y)$$

and improved values for y_1^1 and y_{-1}^1 are obtained.

(d) The improved values for y_1 and y_{-1} are then substituted back into equations (3.69) and (3.70) to get the improved values for y_1 and y_{-1}.

Steps (c) and (d) are repeated until no change is produced in y_1 and y_{-1}. At this point, the iteration is said to have converged and the values of y_1 and y_{-1}, and the corresponding y_1' and y_{-1}', are accepted as final values.

The second iteration uses (3.71), (3.72), and (3.73) to compute values of y_2 and y_{-2}; and the iteration proceed in three steps.

(a) The three consecutive values y_{-1}', y_0', and y_1' of the desired degree of accuracy and substituted into (3.71) and (3.72) to get approximate values of y_2 and y_{-2}.

(b) These values together with x_2 and x_{-2} are substituted into differential equation to obtain approximate values of y_2' and y_{-2}' .

(c) The approximate values of y_2' and y_{-2}' are substituted into (3.73) to get improved values of y_2 and y_{-2}.

Steps (b) and (c) are repeated until no change occurs in y_2 and y_{-2}. At this point, five values of y and y' have been obtained, and the starting procedure is complete.

2.3.5.3.2 Remarks. The requirement that the derivatives y' and y'' exist is less restrictive than the requirements for starting by Taylor series, and more restrictive than the requirements for Euler's method or Runge-Kutta starting. Milne's method is more complicated than the previous methods in that two iterations are required; however, the additional complication can be justified by a higher accuracy potential.

2.3.5.4 Starting by Runge-Kutta

2.3.5.4.1 Discussion. The Runge-Kutta method of numerical integration is discussed in section 2.2.3 of this monograph; a starting procedure using the standard fourth order method, will be illustrated here.

For a first order equation, $y' = f(x,y)$, assuming that the functional value y_i is known at the i^{th} step, the next value y_{i+1} is obtained by the formula

$$y_{i+1} = y_i + \frac{h}{6}\left[k_1 + 2k_2 + 2k_3 + k_4\right] + O(h^5)$$

$$= y_i + \Delta y_i + O(h^5) \qquad (3.74)$$

where

$$k_1 = h f(x_i, y_i)$$

$$k_2 = h f\left(x_i + \frac{h}{2}, y_i + \frac{k_1}{2}\right)$$

$$k_3 = h f\left(x_i + \frac{h}{2}, y_i + \frac{k_2}{2}\right)$$

$$k_4 = h f(x_i + h, y_i + k_3)$$

It can be shown that (3.74) will reduce to Simpson's one-third rule whenever $f(x,y)$ is a function of x alone. Hence, having obtained three successive equally spaced values y_{i-1}, y_i , y_{i+1} , Simpson's rule may be used as a corrector over the two intervals to obtain a new estimate of y_{i+1}

$$y_{i+1}^{s} = y_i + \frac{h}{3}\left[y'_{i-1} + 2y'_i + y'_{i+1}\right] .$$

This has the same local truncation error as the Runge-Kutta method, and one can use the difference $y_{i+1}^{s} - y_{i+1}$ as an error estimate. If the estimated error is greater than that expected in the integration procedure, the step size can be reduced.

2.3.5.4.2 Remarks. The main disadvantage of the Runge-Kutta method is the necessity for four evaluations of the function per step. However, this drawback is more than offset by the simplicity, and potential accuracy of the method. Thus, this procedure is well suited to the roll of starter. However, the use of Runge-Kutta for starting predictor correction integration formula requires one of several precautions to assure accuracy in the results compatible with those obtainable from the continuation routine since the order of accuracy of those formulas which were derived was somewhat less than most of the predictor-corrector formulas. The first and simplest alternative is the integration of the desired function at one-half or one-fourth of the expected integration step size. Then if the integration step must be halved, no additional starting values need be computed. The second alternative would be to employ a corrector cycle to revise the tabulated arrays of position and velocity. This mechanization could be accomplished using the differences as provided by Runge-Kutta and a central difference formula such as δ^2 or $\sum^2$. The last alternative is generally the most difficult and inefficient (the other technique's generally required no additional mechanization logic). This alternative is the inclusion of higher order terms in the Runge-Kutta formula.

3.0 RECOMMENDED PROCEDURES

This monograph considers the analytical techniques available for the solution of the n-body problem. The methods available are discussed in two sections. The first section (section 2.2) discusses the three analytical formulations most commonly used in handling the equations of the n-body problem. The second section (section 2.3) discusses numerical integration and the specific methods that have been most successful in dealing with the n-body problem. Recommendations in these two areas will be made independently.

The three most common mathematical formulations of the equations of motion in current use are: Cowell's method, Encke's method, and the variation of parameters. None of these methods is superior for all types of trajectories; however, for most practical problems a choice can be made based on the available information. If large pertabative accelerations which are changing rapidly are being experienced as, for example, during boost or re-entry, then small integration steps are required with the variation of parameters method and Encke's method as well as with Cowell's method. In this case, the accuracies of the three methods are comparable; however, since the Cowell method requires less time per integration step than the other two and since it avoids problems pertaining to the definition of the reference motion, it is preferred. For phases of the flight where the force field is nearly conservative and/or nearly central with respect to the body about which the motion is occuring, advantage can be taken of the more sophisticated approaches such as Encke's method or the method of variation of parameters. The choice between these two methods is seldom clear cut and is more often dictated by the analysts engrained preferences than by any demonstrable superiority. However, in the past, Encke's method has been used mainly for those lunar or interplanetary orbits where the number of reference trajectory rectifications is small (generally no thrust capability is considered). In contrast, the variation of parameters method has been used where small perturbations acting throughout the orbit, e.g. microthrust transfer produce relatively large changes in the trajectory relative to a conic reference.

Of the available numerical integration techniques, multi-step predictor-corrector methods have been proven to be significantly more efficient (i.e., faster for the same accuracy) than any other method. Two of the methods discussed, Lanczos and Nordsieck can be classified as new and promising. Thus, any contemplated use of these methods should be tempered by this observation. This comment should not be interpreted as disparagement since it is felt that more experience is needed with these methods; rather, their use is encouraged whenever circumstances allow an exploratory approach. However, the Gauss-Jackson method of integration is recommended for general application. This recommendation is substantiated by a long history of use in astronomy where its accuracy has been demonstrated and by the fact that the first and second sums carried in the solution afford much better predictions than possible with the other approaches considered. (Thus, fewer predictor-corrector cycles are generally required for a given accuracy). As a close alternate, the Adams-Moulton predictor-corrector method is recommended. Both of these methods have been in use for a considerable length of time and have emerged to a position of eminence in post digital computer re-evaluation

of integration techniques. It should be pointed out that while the Gauss-Jackson method has historically been used for integrating conservative forces there is nothing in the formulation which precludes its use with non-conservative forces provided the first integral is evaluated.

Even though the multi-step methods discussed above are more efficient than single-step methods, there is one outstanding attribute of the single-step Runge-Kutta method which makes its use desirable in some cases. This attribute is its stability, i.e., its tenacity in following the solution curve. Therefore, when the nature of the solution is unknown and a stability problem is suspected (e.g., a boost trajectory) the Runge-Kutta method should be used.

In section 2.2.5 several methods of starting multi-step methods were discussed. Of these methods, the most versatile as well as the simplest is Runge-Kutta. So straight forward, in fact, is Runge-Kutta that it is recommended without major reservation as a starter for the multi-step methods. In general, however, the tabular array of functional values generated by the Runge-Kutta method should be differentially corrected before initiating a predictor-corrector process to assure comparable accuracies in the two phases of the process. This corrector process can be accomplished by employing the corrector equations directly; thus, no additional logic is required.

4.0 REFERENCES

1.1 Baker, R.M.L. Jr., and Makemson, Maud, W. An Introduction to Astrodynamics. New York: Academic Press (1960).

1.2 Explanatory Supplement to Astronautical Ephemeris and the American Ephemeris and Nautical Almanac. London: Issued by H. M. Nautical Almanac Office by Order of the Lords Commissioners of the Admiralty, Her Majesty's Stationery Office (1961).

1.3 Townsend, G. E., et al. State Determination and/or Estimation. NAA S&ID, SID 65-1200-5 (February 1966).

1.4 Planetary Coordinates for the Years 1960-1980 Referred to the Equinox of 1950.0. London: Prepared by H. M. Nautical Almanac Office by Order of the Lords Commissioners of the Admiralty, Her Majesty's Stationery Office (1958).

1.5 The American Ephemeris and Nautical Almanac for the Year 1965. Washington, D. C. Issued by the Nautical Almanac Office, United States Naval Observatory by direction of the Secretary of the Navy and under the authority of Congress (1963).

1.6 Conte, Samuel D. The Computation of Satellite Orbit Trajectories. Space Technology Laboratories, Inc., Los Angeles, California (1961).

1.7 Scarborough, James B. Numerical Mathematical Analysis. Baltimore: The Johns Hopkins Press (1962).

1.8 Adams, J. C. and Bashforth, F. An Attempt to Test the Theories of Capillary Attraction. Cambridge, England: Cambridge University Press (1883).

1.9 Nordsieck, Arnold "On Numerical Integration of Ordinary Differential Equations," Mathematics of Computation. Volume 16, No. 77 (January 1962) pp. 22-49 .

1.10 Lanczos, Cornelius Linear Differential Operators. New York: D. Van Nostrand Company (1961).

2.0 Townsend, G. E. "Perturbations," The Orbital Flight Handbook. NASA SP-33 (1962).

2.1 Baker, R. "Three Dimensional Drag Perturbation Techniques," ARS Journal. Volume 30, No. 8 (August 1960).

2.2 Polyakhova, Y. N. "Solar Radiation Pressure and the Motion of Earth Satellites," AIAA Journal. Volume 1, No. 12 (December 1963) pp. 2893-2909.

2.3 Jastrow, R. "Atmospheric Drag on Satellites," J. Geophysical Research. Volume 62 (1957) pp. 413-423.

2.4 Danbe, J.M.A. Fundamentals of Celestial Mechanics. New York: The MacMillan Company (1962).

2.5 Garafalo, A. M. "A New Set of Variables for Astronomical Problems," Astronomical Journal. Volume 65 (April 1960) pp. 117-121.

2.6 Newton, R. R. "Variables That Are Determined for Any Orbit," ARS Journal. Volume 31 (March 1961) pp. 364-366.

2.7 Wong, P. "Nonsingular Variation of Parameter Equations for Computation of Space Trajectories," ARS Journal. (February 1962).

2.8 Townsend, G. E. and Tamburro, M. B. The Two Body Problem. NAA S&ID, SID 65-1200-3 (February 1966).

2.9 Boylan, Robert F. A Method for the Numerical Integration of a Special Class of Second-Order Differential Equations. R-463, Instrumentation Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts (1964).

2.10 Scarborough, J. P. Numerical Mathematical Analysis. Baltimore: The Johns Hopkins Press (1962).

2.11 Lanczos, C. Linear Differential Operators. New York: D. Van Nostrand Company (1961).

2.12 Hildebrand, F. B. Introduction to Numerical Analysis. New York: McGraw-Hill (1956).

2.13 Ralston, Anthony "Runge-Kutta Methods with Minimum Error Bound," Mathematics of Computations. Volume 16 (October 1962).

2.14 Richardson, L. F. and Gaunt, J. A. "The Deferred Approach to the Limit," Trans. Roy. Soc. Volume 226A (1927) pp. 300.

2.15 Carr III, J. W. "Error Bounds for the Runge-Kutta Single Step Integration Process," ACM Journal. Volume 5 (January 1958) pp. 39.

2.16 Ralston and Wilf Mathematical Method for Digital Computers. New York: John Whiley & Sons, Inc. (1960).

2.17 Galler, B. A. and Rozenbarg, D. P. "A Generalization of a Theorem of Carr on Error Bounds for Runge-Kutta Procedures," ACM Journal. Volume 7 (1960) pp. 57-60.

2.18 Gill, S. "A Process for the Step-by-Step Integration of Differential Equations in an Automatic Digital Computing Machine," Proc. Cambridge. Phil. Society. Volume 47 (1951).

2.19 Hute, A. "Une Amelioration de la Methode de Runge-Kutta-Nystron, Pour la Resolution Numerique des Equations Differentielles du Premir Ordre," Acts Facultatis Nerum Naturalium Universitatis Comeniarae, Mathematica. Bratislava: Tomus I, Fasc. IV-VI (1956).

2.20 Shanks, E. B. Higher-Order Approximations of Runge-Kutta Type. NASA Technical Note, TN D-2920 (September 1965).

2.21 Durham, H. L., et al. Study of Methods for the Numerical Solution of Ordinary Differential Equations. Engineering Experimentation Station, Georgia Institute of Technology, Document N65-20106 (November 1964).

2.22 Fehlberg, E. New High-Order Runge-Kutta Formulas with Step Size Control for Systems of First and Second-Order Differential Equations. Presented at the Meeting of the GAMM, Giessen, Germany (April 1964).

2.23 Bashforth, F. and Adams, J. C. An Attempt to Test the Theories of Capillary Attraction. Cambridge University Press (1883).

2.24 Herrick, S. "Astrodynamics," in publication.