MULTIPLE INTEGRATION FORMULAS FOR THE NUMERICAL INTEGRATION OF THE SATELLITE FACILITIES (Texas Univ. at Arlington) 146 PAC 11/4/81 CS61 22A
MULTISTEP INTEGRATION FORMULAS FOR THE NUMERICAL
INTEGRATION OF THE SATELLITE PROBLEM

John B. Lundberg

The University of Texas at Austin
Austin, Texas 78712

IASOM 81-1
April, 1981

INSTITUTE FOR ADVANCED STUDY
IN ORBITAL MECHANICS
The University of Texas at Austin
This report was prepared under
Contract NCR44-012-283

for the
National Aeronautics and Space Administration
Goddard Space Flight Center
Greenbelt, Maryland

by the
Institute for Advanced Study in Orbital Mechanics
The University of Texas at Austin
Austin, Texas 78712

under the direction of
Dr. Byron D. Tapley
Principal Investigator
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgements</td>
<td>iii</td>
</tr>
<tr>
<td>Chapter I Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter II PECE Algorithms for the General and Second-Sum Formulations</td>
<td>6</td>
</tr>
<tr>
<td>II.1 General Formulation</td>
<td>7</td>
</tr>
<tr>
<td>II.2 Second-Sum Formulation</td>
<td>11</td>
</tr>
<tr>
<td>II.3 Relationships Between the General and Second-Sum Coefficients</td>
<td>15</td>
</tr>
<tr>
<td>II.4 Development of the PECE Algorithm Equations</td>
<td>17</td>
</tr>
<tr>
<td>II.5 A Modification to the PECE Algorithm</td>
<td>21</td>
</tr>
<tr>
<td>II.6 Notes and Comments</td>
<td>24</td>
</tr>
<tr>
<td>Chapter III Starting Methods</td>
<td>26</td>
</tr>
<tr>
<td>III.1 Summary of Starting Algorithms</td>
<td>27</td>
</tr>
<tr>
<td>III.2 The Central Iterative Starting Algorithm</td>
<td>28</td>
</tr>
<tr>
<td>Chapter IV Comparisons and Results</td>
<td>33</td>
</tr>
<tr>
<td>IV.1 Integrator Comparisons for Differential Equations with Periodic Solutions</td>
<td>34</td>
</tr>
<tr>
<td>IV.1.1 Harmonic Oscillator Problem</td>
<td>37</td>
</tr>
<tr>
<td>IV.1.2 Circular Problem of Two Bodies</td>
<td>38</td>
</tr>
<tr>
<td>IV.1.3 Elliptic Problem of Two Bodies</td>
<td>57</td>
</tr>
<tr>
<td>IV.1.4 Euler Rigid-Body Problem</td>
<td>67</td>
</tr>
<tr>
<td>IV.1.5 Restricted Three-Body Problem</td>
<td>87</td>
</tr>
<tr>
<td>IV.1.6 Comments on the COMPAR Results</td>
<td>92</td>
</tr>
</tbody>
</table>
IV.2 Comparison of Integration Errors for a Typical Satellite Problem .......... 93
IV.2.1 Two-Body Force Model .............. 101
IV.2.2 Spherical Eleventh Degree and Order Force Model ................. 102
Chapter V Conclusions ............................ 121
Appendix A Coefficients for the General Formulation ...... 126
Appendix B Coefficients for the Second-Sum Formulation .. 128
Appendix C Modified, General Formulation, PECE Algorithm 130
Appendix D Modified, Second-Sum Formulation, PECE Algorithm 132
Appendix E Starting Algorithm for the General Formulation .......... 134
Appendix F Starting Algorithm for the Second-Sum Formulation ........ 137
Bibliography ............................................. 140
Vita
CHAPTER I

Introduction

The general initial value problem in the theory of ordinary differential equations may be defined as the problem of determining the values of the \( m \) dependent variables \( y^{(j)} \), \( j=0,\ldots,m-1 \) corresponding to some desired value of the independent variable, \( t \), subject to 1) the ordinary differential equation

\[
\frac{d^m y}{dt^m} = y^{(m)} = f(t,y^{(1)},\ldots,y^{(m-1)})
\]  

(1.1)

and 2) a specified set of initial conditions \( t_0, y^{(0)}(t_0), y^{(1)}(t_0), \ldots, y^{(m-1)}(t_0) \). The \( m \) dependent variables may be \( p \)-dimensional vectors, and hence, (1.1) may be a \( p \)-dimensional vector differential equation. The state vector of the initial value problem is defined to consist of the \( m \) dependent variables; whereas, (1.1) is referred to as a differential equation of order \( m \). The right-hand side of (1.1), \( f(t,y^{(1)},\ldots,y^{(m-1)}) \), is referred to as the function. The initial value problem is solved when the state corresponding to some desired value of the independent variable is determined. A numerical integration method which solves (1.1) directly, i.e., without reducing (1.1) to a set of first order differential equations, is referred to as a Class \( m \) integration method.
The initial value problem frequently occurs in science and engineering. One particular application of the initial value problem which is of interest is the satellite orbit problem. The ordinary differential equations which describe the motion of a satellite acted upon by gravitational and nongravitational forces are given by Newton's Second Law

$$\frac{d^2\mathbf{r}}{dt^2} = -\frac{\mu}{r^3} + \mathbf{r} - \mathbf{g}(\mathbf{r}, t) + \mathbf{n}(\mathbf{r}, \mathbf{r}^{(1)}, t)$$

$$= \mathbf{r} - \mathbf{r}^{(1)}, t \quad (1.2)$$

where $\mu \equiv$ gravitational constant (defined as the product of the universal constant of gravitation, $G$, and the mass, $M$, of the primary body)

$\mathbf{r} \equiv$ the position vector of the satellite

$r \equiv ||\mathbf{r}||$

$\mathbf{r}^{(1)} \equiv$ the velocity vector of the satellite

$\mathbf{g} \equiv$ vector of gravitational forces

$\mathbf{n} \equiv$ vector of nongravitational forces

Equation (1.2) is a second-order, nonlinear, ordinary differential equation where the function $\mathbf{g}(\mathbf{r}, t)$ is a smooth, periodic function representing the forces of gravity acting upon the satellite. The nongravitational contribution may be discontinuous, e.g., entry and exit of the satellite into a planet's shadow will suddenly affect the solar radiation pressure force.
There are many numerical techniques available for solving an initial value, ordinary differential equation, and each numerical method has a limiting degree of accuracy associated with it. The selection of a particular numerical method is subject to the accuracy and cost of using the numerical method, where the cost is usually measured in terms of computer time.

The numerical integration techniques may be divided into two categories: single-step methods and multistep methods. Single-step methods require only the value of the state at one value of the independent variable in order to advance the solution to another value of the independent variable. Multistep methods generally require values of the state or values of the function at more than one value of the independent variable in order to advance the solution. The values of the dependent variable or function, \( \bar{f} \), that are used to advance the solution are determined at values of the independent variable which are referred to as the nodes. The stepsizes used to determine the spacing between the nodes comprise the mesh sequence. Both techniques require the evaluation of the function, \( \bar{f} \), with values of the dependent variables which correspond to various values of the independent variable.

Equation (1.2) may be reduced from a set of second-order ordinary differential equations to a set of first-order equations, thus allowing the option of integrating (1.2) with a Class I or a Class II method. However, Krogh (1970) and Solis (1975) indicate that (1.2) may be integrated more efficiently with a Class II method than with a Class I method. Since \( \bar{f} \) is a well-behaved, periodic function when
gravitational forces are dominant, multistep methods are generally more efficient than single-step methods from an accuracy versus number of function evaluations point of view. Lambert (1973) discusses some of the advantages and disadvantages of using a multistep method instead of a single-step method.

Unlike fixed-order/fixed-mesh multistep integrators, variable-order/variable-mesh methods estimate the local truncation error at each node of the integration in order to satisfy specified tolerances. Thus, if a variable-step/variable-mesh integrator uses approximately a constant stepsize and order, then a fixed-order/fixed-mesh integrator should require fewer computations to integrate the same problem with comparable accuracy. For the satellite problem as described by (1.2), the fixed-order/fixed-mesh integrators may have a significant advantage over variable-mesh/variable-order integrators in the amount of computations required, particularly if the orbital eccentricity is small.

The most common algorithm for solving the satellite problem with a multistep integrator is the PECE algorithm. In the PECE algorithm, the solution at \( t_j \) is extrapolated, or predicted (P), forward to \( t_{j+1} \). The predicted solution at \( t_{j+1} \) is used to evaluate (E) the function, \( \bar{r} \). Using this evaluation, the extrapolated solution is corrected (C), and a second evaluation (E) of \( \bar{r} \) is made with the corrected solution at \( t_{j+1} \). Other multistep algorithms include the PE(CE)\( ^n \) and P(EC)\( ^n \) algorithms where \( n \) indicates the number of times the steps in parenthesis are applied. However, Krogh (1970) notes that
for satellite orbits with small eccentricities, the Class II/PECE methods are more efficient than Class II/PEC methods or Class I/PECE methods.

It is common in the literature to refer to all Class II, multistep formulations as Cowell methods and to refer to Class I methods as Adams methods. However, at least three distinct Class II methods and two distinct Class I methods are available. To avoid any ambiguities, the terminology of Mersman (1965) will be adopted in which the Class II methods are referred to as the general, second-sum, and Stormer-Cowell formulations, while the Class I methods are referred to as the Adams-Bashforth-Moulton and the first-sum formulations.

The purpose of this report is to examine the use of two Class II/fixed-mesh/fixed-order/multistep integration packages of the PECE type for the numerical integration of (1.2). These two methods are referred to as the general and the second-sum formulations. Chapter II discusses the derivation of the basic equations which characterize each formulation and discusses the role of the basic equations in the PECE algorithm. Chapter III discusses possible starting procedures which may be used to supply the initial set of values required by the fixed-mesh/multistep integrators. In Chapter IV, the results of the general and second-sum integrators are compared to the results of various fixed-step and variable-step integrators.
CHAPTER II

PECE Algorithms for the General and Second-Sum Formulations

In this chapter, two fixed-mesh/multistep PECE algorithms are developed for the numerical integration of second-order, initial value, ordinary differential equations represented by (1.1). These two formulations are the general and the second-sum formulations. The basic assumptions in developing these fixed-mesh/multistep integrators are:
1) the value of the state vector \( y(t_n), y^{(1)}(t_n) \) and the value of \( t_n \) are known;
2) the function values \( f_j = f(t_j, y_j, y_j^{(1)}) \), \( j = n, n-1, ..., n-i+1 \), are known at \( i \) distinct nodes where \( f \) represents the right-hand side of (1.1), and
3) the nodes, \( t_j \), associated with the values \( f_j \) are known and satisfy the condition \( t_j = t_{j-1} + h \), where \( j = n, n-1, ..., n-i+1 \) and \( h \) is a constant step-size. The first assumption is satisfied by definition of the initial value problem. The second and third are assumed to be satisfied by appropriate values supplied from some starting procedure (see Chapter III).

The development of the PECE algorithm begins by deriving the basic equations for each of the formulations. The basic equations developed in Sections II.1 and II.2 are used to extrapolate or interpolate the state at \( t \) from the state at \( t_n \) by the proper choice of coefficients. If the value of \( t \) is between \( t_n \) and \( t_{n-i+1} \), i.e.,
\[ t_n - t_{n-1} = \frac{t_n - t}{h}, \text{ then the basic equations interpolate the state at } t, \text{ and if the value of } t \text{ is not between } t_n \text{ and } t_{n-1}, \text{ then the basic equations extrapolate the state at } t. \text{ Using the extrapolation and interpolation capabilities of the basic equations, the PECE algorithms are determined. A discussion of some modifications made available by the use of back differences concludes the development of the PECE algorithms in this chapter.} \]

II.1 General Formulation

The solution of the second-order ordinary differential equation given by

\[ y^{(2)} = f(t, y, y^{(1)}) \]

is

\[ y = y_n + h y_n^{(1)} + \int_{t_n}^{1} f(x, y, y^{(1)}) dx \quad d_1 \]

(2.1)

\[ y^{(1)} = y_n^{(1)} + \int_{t_n}^{1} f(x, y, y^{(1)}) dx \quad d_2 \]

(2.2)

In most applications, the function \( f \) cannot be readily integrated by analytical means. Thus, \( f \) can be replaced by an approximating function which represents it to some specified degree of accuracy. The derivation of the general formulation algorithm uses the available nodes \( t_j \) and function \( f(t, y, y^{(1)}) \) to form a polynomial \( P(t) \) which is assumed to yield \( f_j \) when evaluated at \( t_j \). Hence,
The polynomial, $P(t)$, may be written in divided difference form as

$$P(t) = f_n + (t-t_n)g[f_n,f_{n-1}] + \cdots + (t-t_n)^{n-i+1}g[f_n,f_{n-1},\ldots,f_{n-i+1}]$$

where $g[ ]$ is the divided difference operator which is defined by

$$g[x_1,x_2,\ldots,x_n] = \frac{g[x_1,\ldots,x_{n-1}]-g[x_2,\ldots,x_n]}{x_1-x_n}$$

and the degree of the polynomial is $(i-1)$. Imposing the fixed-mesh criterion, the divided differences may be written as back differences, and the polynomial, $P(t)$, becomes

$$P(t) = f_n + \frac{(t-t_n)}{1!} h f_n + \cdots + \frac{(t-t_n)^{i-1}}{(i-1)! h^{i-1}} v^{i-1} f_n$$

where

$$y_1 = 1; \quad y_j = \frac{(t-t_n)^{j-1}}{(j-1)! h^{j-1}} \quad j=2,3,\ldots,i$$

and where the back difference operator, $\nabla^j$, is defined by

$$\nabla^0 r_k = r_k \quad \text{and} \quad \nabla^j r_k = \nabla^{j-1} r_k - \nabla^{j-1} r_{k-1}$$

$$= \sum_{m=0}^{j} (-1)^m r_{k-m} \binom{j}{m}$$
where \( \binom{j}{m} \) are binomial coefficients.

Using (2.4), it follows that (2.1) and (2.2) become

\[
y = y_n + h y_n^{(1)} + \int_{t_n}^{x_1} \sum_{j=1}^{i} y_j(x) v^{j-1} f_n \, dx \, dx_1
\]

\[
y^{(1)} = y_n + \int_{t_n}^{t} \sum_{j=1}^{i} y_j(x) v^{j-1} f_n \, dx.
\]

Since only the coefficients, \( y_j \), are a function of the independent variable, (2.5) and (2.6) may be rewritten as

\[
y(t_{n+r}) = y_{n+r} = y_n + h y_n^{(1)} + h^2 \sum_{j=1}^{i} \alpha_{j,r} v^{j-1} f_n
\]

\[
y^{(1)}(t_{n+r}) = y_{n+r}^{(1)} = y_n^{(1)} + h \sum_{j=1}^{i} \beta_{j,r} v^{j-1} f_n
\]

where

\[
\alpha_{j,r} = \int_{0}^{1} \int_{0}^{1} \frac{s_1 (s_1 h + \cdots + (j-2)h) \cdots (s_1 + (j-2)h)}{(j-1)! h^{j-1}} ds \, ds_1
\]

\[
\beta_{j,r} = \int_{0}^{1} \frac{(s_1 h + \cdots + (j-2)h)}{(j-1)! h^{j-1}} ds
\]

for \( j = 2, 3, \ldots, i \)

\[
\alpha_{1,r} = \frac{1}{2}, \beta_{1,r} = 1, s = \frac{t-t_n}{h_I}, h_I = t_{n+r} - t_n \text{ and } t_{n+r} = t_n + rh.
\]

As defined by Shampine and Gordon, basic equations which use \( i \) nodes are of order \( i \); thus, Equations (2.7) and (2.8) are defined to be order \( i \).
Equations (2.7) and (2.8) are the basic equations for solving a second-order ordinary differential equation using the general formulation. Equation (2.8) can be used for solving a first-order ordinary differential equation. Krogh (1970) and Shampine and Gordon (1975) discuss the idea of extending the set of basic equations to solve differential equations of order greater than two.

Equations (2.7) and (2.8) show that the problem of evaluating the integrals of (2.1) and (2.2) has been reduced to a problem of evaluating the coefficients, $\alpha$ and $\beta$. Shampine and Gordon (1975) present the derivation of the algorithm that is used to evaluate the coefficients for a variable-mesh/multistep algorithm. The algorithm for calculating fixed-mesh coefficients is a simplification of the variable-mesh coefficient algorithm and is presented in Appendix A.

It should be noted that even though (2.7) and (2.8) are written in terms of the back differences, they may also be written explicitly in terms of the function values as

$$y_{n+r} = y_n + h \sum_{j=1}^{i} \alpha_{j,r} f_{n+1-j}$$

$$y^{(1)}_{n+r} = y^{(1)}_n + h \sum_{j=1}^{i} \beta_{j,r} f_{n+1-j}$$

where

$$\alpha_{j,r} = (-1)^{j-1} \sum_{q=j}^{(q)} \binom{j-1}{q-1} \alpha_{q,r}$$
\[ f_{j,r} = (-1)^{j-1} \sum_{q=j}^{q=j-1} \binom{q}{j-1} f_{q,r}. \]

Thus, the basic equations of the general formulation may be written in terms of back differences as in (2.7) and (2.8) or in terms of function values as in (2.9) and (2.10).

II.2 Second-Sum Formulation

In order to obtain the basic equations for the second-sum formulation, certain operators and relationships are required. The necessary operators are:

1. the back difference operator, \( V \), where \( V^{k+1} f_n = V^k f_n - V^k f_{n-1} \)

2. the shift operator, \( E \), where \( E^k f_n = f(t_n + kh) \), and

3. the differential operator, \( D \), where

\[
D^r = \frac{d^r}{dt^r} \quad \text{and} \quad x^{(1)} = \frac{dx}{dt} = D^{-1} \frac{d^2 x}{dt^2}. \tag{2.11}
\]

As shown by Hildebrand (1974), the above operators satisfy the following relationships:

\[ e^{-hD} = 1 - V = E^{-1} \tag{2.12} \]

and

\[ D^{-1} = \frac{-h}{\ln(1 - V)} \tag{2.13} \]

The use of the \( V \) operator implies that the stepsize \( h \) is a con-
stant.

The derivation of the basic equations begins by making use of the three operators above and the relationships from (2.11), (2.12) and (2.13). From (2.12), we may write

\[ y_{n+s} = E^s y_n = (1-V)^{-s} y_n \]  
(2.14)

\[ y_{n+s}^{(1)} = E^s y_n^{(1)} = (1-V)^{-s} y_n^{(1)} \]  
(2.15)

Using equations (2.11) and (2.13), we find that

\[ y_n = d^{-2} y_n = \left[ \frac{-h}{\ln(1-V)} \right] y_n \]  
(2.16)

\[ y_n^{(1)} = d^{-1} y_n = \left[ \frac{-h}{\ln(1-V)} \right] y_n \]  
(2.17)

By combining (2.14) and (2.15) with (2.16) and (2.17), we obtain the following relationships

\[ y(t_n + sh) = y_{n+s} = (1-V)^{-s} \left[ \frac{-h}{\ln(1-V)} \right] ^2 y_n \]

\[ = h^2 \left[ \frac{(1-V)^{-s}}{(\ln(1-V))^2} \right] f_n \]  
(2.18)

\[ y^{(1)}(t_n + sh) = y_{n+s}^{(1)} = (1-V)^{-s} \left[ \frac{-h}{\ln(1-V)} \right] y_n \]

\[ = h \left[ \frac{- (1-V)^{-s}}{\ln(1-V)} \right] f_n \]  
(2.19)

By expanding the terms in the brackets of (2.18) and (2.19) in an infinite series, i.e.,

\[ \frac{(1-V)^{-s}}{(\ln(1-V))^2} = \sum_{j=-1}^{\infty} a_{j,s} y^j f_n \]  
(2.20)
\[
\frac{(1-V)^{-s}}{\ln(1-V)} = \sum_{j=1}^{\infty} b_{j,s} \ V^j f_n
\]  
(2.21)

It follows that (2.18) and (2.19) may be rewritten as

\[
y_{n+s} = h^2 \sum_{j=2}^{\infty} a_{j,s} \ V^j f_n
\]  
(2.22)

\[
y_{n+s}^{(1)} = h \sum_{j=-1}^{\infty} b_{j,s} \ V^j f_n
\]  
(2.23)

The coefficients \(a_{j,s}\) and \(b_{j,s}\) are defined in terms of \(s\) by expanding the left-hand sides of (2.20) and (2.21) in an infinite series in \(V\) and comparing the coefficients of the different powers of the \(V\) operator. Derivation of the recursive algorithm used to calculate the coefficients is cumbersome and is discussed by Spier (1971) and Velez and Maury (1970). The algorithm used to generate the coefficients for the fixed-step second-sum formulation is summarized in Appendix B. It should be noted that the coefficients \(a_{-2,s}\), \(a_{-1,s}\) and \(b_{-1,s}\) are such that

\[
a_{-2,s} = b_{-1,s} = 1
\]

\[
a_{-1,s} = s-1
\]

for all \(s\). Thus, (2.22) and (2.23) may be rewritten as

\[
y_{n+s} = h^2 \left[ V^{-2} f_n - (1-s) V^{-1} f_n + \sum_{j=0}^{i-1} a_{j,s} V^j f_n \right]
\]  
(2.24)

\[
y_{n+s}^{(1)} = h \left[ V^{-1} f_n + \sum_{j=0}^{i-1} b_{j,s} V^j f_n \right]
\]  
(2.25)
where the summation terms have been truncated to include the first \( i \) terms. Equations (2.24) and (2.25) are said to be of order \( i \). In (2.24) and (2.25), the terms \( V^{-1}f_n \) and \( V^{-2}f_n \) are referred to as the first and second terms, respectively. The first and second terms satisfy the following relationships,

\[
V^{-1}f_n = V^{-1}f_{n-1} + f_n \tag{2.26}
\]

\[
V^{-2}f_n = V^{-2}f_{n-1} + V^{-1}f_n \tag{2.27}
\]

from the definition of the back difference operator. It should be noted that only (2.25) is required to solve a first-order ordinary differential equation.

By using function values instead of back difference, (2.24) and (2.25) may be rewritten as

\[
y_{n+s} = h^2 \left[ V^{-2}f_n - (1-s)V^{-1}f_n + \sum_{j=0}^{1-1} a_{j,s} f_{n-j} \right] \tag{2.28}
\]

\[
y_{n+s}^{(1)} = h \left[ V^{-1}f_n + \sum_{j=0}^{1-1} b_{j,s} f_{n-j} \right] \tag{2.29}
\]

where

\[
a_{j,s} = (-1)^{j-1} \sum_{q=j}^{1-1} \binom{q}{j-1} a_{j,s}
\]

\[
b_{j,s} = (-1)^{j-1} \sum_{q=j}^{1-1} \binom{q}{j-1} b_{j,s}
\]
II.3 Relationships Between the General and Second-Sum Coefficients

To illustrate the relationship between the coefficients of the general formulation and second-sum formulation coefficients, each set of basic equations for extrapolating the solution one step is presented and compared. To predict the solution at \( t_{n+1} \) from \( t_n \), the general formulation basic equations are

\[
y_{n+1} = y_n + h \gamma_{n}^{(1)} + h^2 \sum_{j=1}^{i} \alpha_{j,1} \psi_{j-1} f_n \quad (2.30)
\]

\[
y_{n+1}^{(1)} = y_{n}^{(1)} + h \sum_{j=1}^{i} \beta_{j,1} \psi_{j-1} f_n . \quad (2.31)
\]

The form of the basic equations for the second-sum formulation that predict the solution at \( t_{n+1} \) from \( t_n \) are

\[
y_{n+1} = h^2 \left[ \nabla^{-2} f_n + \frac{1}{\sum_{j=1}^{i} \alpha_{j-1,1} \psi_{j-1} f_n } \right] \quad (2.32)
\]

\[
y_{n+1}^{(1)} = h \left[ \nabla^{-1} f_n + \frac{1}{\sum_{j=1}^{i} \beta_{j-1,1} \psi_{j-1} f_n } \right] \quad (2.33)
\]

and that interpolate the solution at \( t_n \) from \( t_n \) are

\[
y_{n} = h^2 \left[ \nabla^{-2} f_n - \nabla^{-1} f_n + \frac{1}{\sum_{j=1}^{i} \alpha_{j-1,0} \psi_{j-1} f_n } \right] \quad (2.34)
\]

\[
y_{n}^{(1)} = h \left[ \nabla^{-1} f_n + \frac{1}{\sum_{j=1}^{i} \beta_{j-1,0} \psi_{j-1} f_n } \right] . \quad (2.35)
\]

From (2.31), it follows that
\[ V_{n+1}^{(1)} = h \sum_{j=1}^{i} \beta_{j,1} V_{j-1}^{n} \]  

(2.36)

Using (2.33) and (2.35), we find

\[ V_{n+1}^{(1)} = h \sum_{j=1}^{i} (b_{j-1,1} - b_{j-1,0}) V_{j-1}^{n} \]

and comparing this equation to (2.36), we also find

\[ \beta_{j-1} = b_{j-1,1} - b_{j-1,0} = \nabla b_{j-1,1} \]  

(2.37)

Mersman (1965) has shown that, in general, \( \nabla b_{j-1,1} = b_{j,1} \).

Thus, (2.37) becomes

\[ \beta_{j,1} = b_{j,1} \]

or

\[ b_{j,k} = \beta_{j,k} \]

Similarly, by forming the back difference \( V_{n+1} \) for each formulation and comparing coefficients, we also find that

\[ a_{j,1} = a_{j-1,1} - a_{j-1,0} - b_{j-1,0} \]

\[ = \nabla a_{j-1,1} - b_{j-1,0} \]  

(2.38)

Mersman (1965) has also shown that \( \nabla a_{k-1,1} = a_{k,1} \). Thus, (2.38) may be written as

\[ a_{j,1} = a_{j,1} - b_{j-1,0} \]

and
\[\alpha_{j,k} = a_{j,k} - b_{j-1,0}.\]

II.4 Development of the PECE Algorithm Equations

The fixed-mesh, PECE algorithm assumes that the values \( f_j \) and \( t_j \), for \( j=n,n-1,\ldots,n-1+1 \), are known, that \( t_j \) satisfies the condition \( t_{j+1} = t_j + h \), and that if \( t, y \) and \( y^{(1)} \) are known, then \( f = f(t,y,y^{(1)}) \) can be calculated. The PECE algorithm is based upon the extrapolation property of the basic equations. The basic equations for the general formulation are

\[
y_p = y_n + hy^{(1)} + h^2 \sum_{j=1}^{1} \alpha_{j,s} v^{d-1} f_n
\]

\[
y_p = y_n + hy^{(1)} + h^2 \sum_{j=1}^{1} \alpha^*_{j,s} f_{n-j+1}\]

(2.39)

\[
y^{(1)}_p = y^{(1)}_n + h \sum_{j=1}^{1} \beta_{j,s} v^{d-1} f_n
\]

\[
y^{(1)}_p = y^{(1)}_n + h \sum_{j=1}^{1} \beta^*_{j,s} f_{n-j+1}\]

(2.40)

and for the second-sum formulation are

\[
y_p = h^2 \left[ v^{-2} f_n + (s-1)v^{-1} f_n + \frac{1}{2} a_{j-1,s} v^{d-1} f_n \right]
\]

\[
y_p = h^2 \left[ v^{-2} f_n + (s-1)v^{-1} f_n + \frac{1}{2} a^*_{j-1,s} f_{n-j+1} \right]\]

(2.41)

\[
y^{(1)}_p = h \left[ v^{-1} f_n + \sum_{j=1}^{1} b_{j-1,s} v^{d-1} f_n \right]
\]
where $s = p-n$.

The prediction step of the PECE algorithm is an extrapolation of the solution from $t_n$ to $t_{n+1}$. By setting $s=1$ in (2.39) through (2.42), the predicted solution at $t_{n+1}$, i.e., $p_{n+1}$ and $p^{(1)}_{n+1}$, is given by the general form of the basic equations as

$$p_{n+1} = y_n + h^2\sum_{j=1}^{i} a_{j,1} y_{j-1} f_n$$

$$= y_n + h^2\sum_{j=1}^{i} a_{j,1} f_{n-j+1}$$

$$p^{(1)}_{n+1} = y^{(1)}_n + h \sum_{j=1}^{i} \beta_{j,1} y_{j-1} f_n$$

$$= y^{(1)}_n + h \sum_{j=1}^{i} \beta_{j,1} f_{n-j+1}$$

and by the second-sum form of the basic equations as

$$p_{n+1} = h^2 \left[ v^{-2} f_n + \sum_{j=1}^{i} a_{j-1,1} y_{j-1} f_n \right]$$

$$= h^2 \left[ v^{-2} f_n + \sum_{j=1}^{i} a_{j-1,1} f_{n-j+1} \right]$$

$$p^{(1)}_{n+1} = h \left[ v^{-1} f_n + \sum_{j=1}^{i} b_{j-1,1} y_{j-1} f_n \right]$$

$$= h \left[ v^{-1} f_n + \sum_{j=1}^{i} b_{j-1,1} f_{n-j+1} \right].$$
Using the values for \( t_{n+1} \), \( p_{n+1} \) and \( p_{n+1}^{(1)} \), the first evaluation of the PECE algorithm is given by \( r_{n+1}^p = f(t_{n+1}, p_{n+1}, p_{n+1}^{(1)}) \).

A new set of back differences, \( \nabla^j f_{n+1}^* \), or function evaluations, \( f_{n+1-j}^* \) are formed at \( t_{n+1} \). The new set of back differences are defined by

\[
\nabla^0 f_{n+1}^* = r_{n+1}^p \\
\nabla^j f_{n+1}^* = \nabla^{j-1} f_{n+1}^* - \nabla^{j-1} f_n^*, \quad j=1, \ldots, 1
\]

and the new set of function evaluations are defined

\[
f_{n+1}^* = r_{n+1}^p \\
\quad f_j^* = f_j, \quad j=n, n-1, \ldots, n-i+1.
\]

With the inclusion of \( r_{n+1}^p \), there are \( i+1 \) back differences or function evaluations available to approximate the function \( f \). The new polynomial approximation of \( f \) now spans from \( t_{n-i+1} \) to \( t_{n+1} \) instead of \( t_{n-i+1} \) to \( t_n \).

The predicted solution, \( p_{n+1} \) and \( p_{n+1}^{(1)} \), is corrected by extrapolating the solution at \( t_n \), i.e., \( y_n \) and \( y_n^{(1)} \), to \( t_{n+1} \) by use of the new polynomial order \( i \). The basic equations give the corrected solution as

\[
y_{n+1} = y_n + h y_n^{(1)} + h^2 \sum_{j=1}^{i+1} a_j \nabla^{j-1} f_{n+1}^* \\
\quad = y_n + h y_n^{(1)} + h^2 \sum_{j=1}^{i+1} a_j f_{n+1-j}^{(*)}
\]

\[ (2.47) \]
\( y_{n+1}^{(1)} = y_n^{(1)} + h \sum_{j=1}^{i+1} \beta_{j,0} f_n^{i+1-j} \)

\[ y_{n+1}^{(1)} = y_n^{(1)} + h \sum_{j=1}^{i+1} \beta_{j,0} f_n^{i+1-j} \tag{2.48} \]

for the general formulation and

\[ y_{n+1}^{(1)} = h^2 \left[ \nabla^{-2} f_n - \nabla^{-1} f_n + \sum_{j=1}^{i+1} a_{j-1,0} \nabla^{j-1} f_{n+1}^* \right] \]

\[ y_{n+1}^{(1)} = h^2 \left[ \nabla^{-2} f_n - \nabla^{-1} f_n + \sum_{j=1}^{i+1} a_{j-1,0} \nabla^{j-1} f_{n+1}^* \right] \tag{2.49} \]

\[ y_{n+1}^{(1)} = \left[ \nabla^{-1} f_n + \sum_{j=1}^{i+1} b_{j-1,0} \nabla^{j-1} f_{n+1}^* \right] \]

\[ y_{n+1}^{(1)} = \left[ \nabla^{-1} f_n + \sum_{j=1}^{i+1} b_{j-1,0} \nabla^{j-1} f_{n+1}^* \right] \tag{2.50} \]

for the second-sum formulation.

The second evaluation of the PECE algorithm is made by using the corrected solution, i.e.,

\[ f_{n+1} = f(t_{n+1}, y_{n+1}, y_{n+1}^{(2)}) . \]

The back differences are updated by the formulas

\[ \nabla^j f_{n+1} = \nabla^{j-1} f_{n+1} - \nabla^{j-1} f_n , \quad j=1, \ldots, i \]

and the first and second sums are updated by

\[ \nabla^{-1} f_{n+1} = \nabla^{-1} f_n + f_{n+1} \]
\[ V^{-2}f_{n+1} = V^{-2}f_n + V^{-1}f_{n+1}. \]

II.5 A Modification to the PECE Algorithm

The use of back differences instead of function evaluations in the PECE algorithm allows for a simplification of the correction formulas. The simplification will be demonstrated here as it applies to the general formulation, but it may also be applied to the second-sum formulation. The process of propagating the back differences is also discussed in this section in order to complement the use of the back difference form of the PECE algorithm.

Equation (2.3) states the approximation of the function \( f \) as a polynomial in \( t \) and using \( i \) nodes and \( i \) function evaluations. In divided difference notation, the polynomial approximation of \( f \) was given as

\[
P(t) = f_n + (t-t_n)g[f_n, f_{n-1}] + \ldots + (t-t_n) \ldots (t-t_{n-i+2})g[f_n, f_{n-1}, \ldots, f_{n-i+1}] \tag{2.51}
\]

where \( g[\ ] \) is the divided difference operator. The general formulation basic equation was derived by writing (2.51) with back differences as

\[
P(t) = \sum_{j=1}^{i} y_j(t) V^{j-1}f_n \tag{2.52}
\]

and integrating the coefficients, \( y_j(t) \), from \( t_n \) to \( t \).
Equation (2.51) represents the polynomial used in the prediction step of the PECE algorithm. The polynomial used in the correction step of the algorithm uses the \( n \) nodes and \( n \) function evaluations used in the predicting step and an additional function evaluation and node, \( f_{n+1}^p \) and \( t_{n+1} \). The polynomial that incorporates the \( n+1 \) nodes and \( n+1 \) function evaluations may be written as

\[
P^p(t) = f_n + (t-t_n)g[f_n,f_{n-1}] + \ldots \\
+ (t-t_n)(t-t_{n-1})g[f_n,f_{n-1},f_{n-2}] \\
+ (t-t_n)(t-t_{n-1})(t-t_{n-2})g[f_n,f_{n-1},f_{n-2},f_{n-3}] \\
= P(t) + (t-t_n)(t-t_{n+1})g[f_n,f_{n-1},f_{n-2},f_{n+1}] 
\tag{2.53}
\]

Equation (2.53) represents the polynomial used in the correction step of the PECE algorithm. The corrected solution is given by the calculus solution

\[
y_{n+1} = y_n + h^*(1) + \int_{t_n}^{t_{n+1}} \int_0^1 p^*(x)dx \, dx_1 
\tag{2.54}
\]

\[
y_{n+1} = y_n + \int_{t_n}^{t_{n+1}} p^*(x)dx 
\tag{2.55}
\]

Using (2.53) with (2.54) and (2.55), we find
\[ y_{n+1} = y_n + h^{(1)}_n + \int_{t_n}^{t_{n+1}} \int_0^1 p(x) dx \, dx_1 \]
\[ + \int_{t_n}^{t_{n+1}} \int_0^1 y_{i+1}(x) \gamma^1 f_{n+1} dx \, dx_1 \]
\[ = p_{n+1} + h^2 y_{i+1,1} V_i^P \]
\[ (2.56) \]

\[ y_n^{(1)} = y_n + \int_{t_n}^{t_{n+1}} p(x) dx + \int_{t_n}^{t_{n+1}} y_{i+1}(x) V_i^P dx \]
\[ = p_{n+1} + h \beta_{i+1,1} V_i^P \]
\[ (2.57) \]

where
\[ V_i^P = f_{n+1} - \sum_{j=1}^i V_j f_n. \]
\[ (2.58) \]

Equations (2.56) and (2.57) represent the correction formulas that would be used in place of (2.47) and (2.48) for the general formulation. Similarly, for the second-sum formulation, (2.49) and (2.50) could be replaced by

\[ y_{n+1} = p_{n+1} + h^2 a_{i+1,1} V_i^P \]
\[ (2.59) \]

\[ y_{n+1}^{(1)} = p_{n+1} + h b_{i+1,1} V_i^P \]
\[ (2.60) \]

One advantage of (2.56) and (2.57) is that, instead of 2i, only \( i+1 \) coefficients are needed to predict and correct the solution. Thus, (i-1) fewer multiplications per step are required. The additional computations required to form \( V_i^P \) and for propagating the back differences from \( t_n \) to \( t_{n+1} \) are a possible disadvantage to
(2.56) and (2.57) However, the propagation of the back differences may take advantage of the intermediate calculations resulting from the calculation of $\nabla^i f_{n+1}$. By retaining the intermediate sums,

$$S_k = \sum_{j=k}^{1} \nabla^{j-1} f_n, \quad k=1, \ldots, i$$

resulting from the evaluation of (2.58), the back differences at $t_{n+1}$ are then calculated from

$$\nabla^i f_{n+1} = S_o = f_{n+1} - S_1$$

$$\nabla^{k-1} f_{n+1} = S_k + S_o, \quad k=1, \ldots, i$$

The modified PECE algorithm is given in Appendix C for the general formulation and in Appendix D for the second-sum formulation.

II.6 Notes and Comments

The algorithm discussed in this chapter is for a PECE multistep integrator with a predictor of order $i$ and a corrector of order $i+1$. The PECE algorithm is one of a family of algorithms in which corrector formulas and the function evaluation using the corrected solution may be applied any number of times. This family of algorithms is represented by PE(CE)$^n$ and P(EC)$^n$. The choice of algorithm to be used to solve a given problem is discussed by Shampine and Gordon (1975) and Krogh (1970).

The order of the correcting formula is generally chosen to be
of the same order or one higher order than the predicting formula. The
decision of which order to use for the corrector is dependent on the
differential equation to be solved and on the order of the predictor.
Considerations for selection of the corrector order are discussed by

The fixed-mesh/multistep algorithm advances the solution in
intervals of the stepsize, h, of the independent variable t. However, the solution may be desired for some value of t, e.g., t_j, such that t_j is between two nodes. The solution at t_j is found by advancing the solution until t_j is bounded by two of the nodes. If the nodes range from t_n to t_{n-1+1} such that
|t_n - t_{n-1+1} | > |t_j - t_n|, the solution at t is interpolated by using r = s = j-n in the basic equations.

The problem of obtaining the initial set of function evaluations for the PECE algorithm is discussed in Chapter III. The general and second-sum algorithms discussed in this chapter are compared to the other numerical integration packages in Chapter IV.
CHAPTER III
Starting Methods

As discussed in Section II.4, a multistep numerical integration algorithm of order i assumes that i function evaluations and i nodes are known. The initial values of the nodes and function evaluations are found by application of an appropriate starting procedure. A starting procedure has only the initial conditions, \( t_o, y_o \) and \( y_o^{(1)} \), and the differential equation, \( y^{(2)} = f(t,y,y^{(1)}) \), available to calculate the values of the nodes and to evaluate the function. Some of the various starting algorithms, which can be used with multistep integrators, include 1) the bootstrap method, 2) the iterative method and 3) the use of a single-step integrator. This chapter begins by discussing the starting procedures for the Class II fixed-mesh/fixed-order multistep integration packages using the general and second-sum algorithms described in the previous chapter. The chapter concludes by detailing a proposed starting procedure for the Class II/fixed-mesh/fixed-order/multistep integration packages which are the emphasis of this report.
III.1 Summary of Starting Algorithms

The bootstrap starter begins with a first-order basic equation and increases the order as the solution is advanced one node at a time. By using the initial conditions, the first function evaluation is made, \( f_0 = f(t, y, y^{(1)}) \). The solution is extended to \( t_1 = t_n + h \) by using a first-order basic equation, e.g., \( i=1 \) in (2.39) and (2.40) or in (2.41) and (2.42). With the predicted solution at \( t_1 \), another function evaluation is made, \( f_1 = f(t_1, y_1, y_1^{(1)}) \). The correction form of the basic equation is employed and an additional function evaluation is made with the corrected solution at \( t_1 \) to complete the PECE algorithm at \( t_1 \). Now, with two function evaluations available, the solution is advanced to \( t_2 \) by application of the PECE algorithm for \( i=2 \). In this manner of bootstrapping, the solution is advanced until the required number of function evaluations are known. This starting procedure is generally performed with variable-step-variable-order multistep integrators in which the local truncation error is estimated after each step in order to determine if the step is acceptable and to determine the next stepsize or order to use.

A single-step numerical integrator, e.g., a Runge-Kutta method, may be used to generate the required function evaluations. By using the single-step integrator to advance the solution from one node to the next, the solution at the nodes is obtained and the function evaluations are calculated and stored.
An iterative starter assumes that a first approximation of the required function evaluations has been made by means of a bootstrap method, by use of a single-step integrator, by a Taylor series expansion or by any other suitable procedure. Once the first approximations have been provided, the basic equation is used to interpolate the solution at each of the nodes, and a second set of function evaluations is calculated. The second set of function evaluations is used with the basic equation to interpolate the solutions at the nodes again and a third set of function evaluations is calculated. The iterative method proceeds in this manner until some termination criteria is satisfied. The criteria may be a certain number of iterations, or it may depend upon the difference between the solutions of two consecutive iterations. Once a criterion is satisfied, the required function evaluations are known.

III.2 The Central Iterative Starting Algorithm

Two important criteria were considered in selecting a starting procedure for the general and second-sum packages of this report. The first criteria for the starting procedure was to generate a set of function evaluations and nodes with respect to the initial conditions which can be used to interpolate the initial conditions exactly. Secondly, a starting procedure should be consistent with the interpolation scheme of the integrator. The interpolation scheme chosen for use in the fixed-step/fixed-order integrators of this report requires that the solution be advanced far enough that the point at which the
interpolation is to be performed is approximately midway between the extreme nodes. This scheme helps to reduce discontinuities in the interpolation when the solution is advanced from one node to the next node. The starting procedure used in the general and second-sum integration packages described in this report is a central, iterative starting procedure.

Backward or forward iterative starting procedures are such that the initial set of nodes span the interval \( t_0 \) to \( t_{1-i} \) or the interval \( t_0 \) to \( t_{1-1} \). However, a central starting procedure selects the end nodes such that the initial conditions lie at a node approximately midway between the end nodes. With the initial conditions at \( t_0 \), the end nodes will be at \( t_j \) and \( t_q \), where \( q = \text{integer} \left( \frac{i}{2} \right) \) and \( j = q - i + 1 \). Thus, the central starting method will have advanced the solution to \( t_q \) and will return the function evaluations at the nodes \( t_k; k=q,q-1,...,q-i+1 \).

As mentioned above, one important requirement for the starting procedure is that the solution at \( t_q \) and the function evaluations must be consistent with the initial conditions. For the general formulation using back differences, this imposes the condition that

\[
y_0 = y_q - qh^\left(1\right) + q^2 h^2 \sum_{j=1}^{i} \alpha_{j,-q} v^{j-1}_{q} \tag{3.1}
\]

\[
y_0^{\left(1\right)} = y_q^{\left(1\right)} - qh \sum_{j=1}^{i} \beta_{j,-q} v^{j-1}_{q} \tag{3.2}
\]

The solution at any of the nodes used in the starting procedure is
given by

\[ y_k = y_q + (k-q)h \ y_q^{(1)} + (k-q)^2h^2 \ \sum_{j=1}^{i} \ \alpha_{j,k-q} \ y_{j-1}^{(1)} \]  
\[ (3.3) \]

\[ y_k^{(1)} = y_q^{(1)} + (k-q)h \ \sum_{j=1}^{i} \ \beta_{j,k-q} \ y_{j-1}^{(1)} \]  
\[ (3.4) \]

for \( k=q,q-1,\ldots,q+1 \).

The interpolation formulas for the central iterative starter may be found by using (3.1), (3.2), (3.3) and (3.4) to obtain

\[ y_k = y_o + kh y_o^{(1)} \]

\[ + \ h^2 \ \sum_{j=1}^{i} \left( (q-k)^2 \alpha_{j,k-q} - q^2 \alpha_{j,-q} + kq \beta_{j,-q} \right) \ y_{j-1}^{(1)} \]  
\[ (3.5) \]

\[ y_j^{(1)} = y_o^{(1)} + h \ \sum_{j=1}^{i} \left( \beta_{j,-q} + (k-q) \beta_{j,k-q} \right) \ y_{j-1}^{(1)} \]  
\[ (3.6) \]

\( k=q,q-1,\ldots,q+1 \).

Equations (3.5) and (3.6) use the initial conditions at \( t_o \) and the back differences at \( t_q \) to interpolate the solution at the nodes. Equations (3.5) and (3.6) will always satisfy the initial conditions.

The initial values of the function evaluations are generated by using the Taylor series expansions

\[ t_j = t_o + kh \]
\[
y_j = y_0 + k y^{(1)}_o + k^2 h^2 f_o
\]
\[
y^{(1)}_k = y^{(1)}_o + k h f_o
\]
\[
f_o = f(t_o, y_o, y^{(1)}_o)
\]
\[
f_k = f(t_k, y_k, y^{(1)}_k)
\]

To approximate the solution at the nodes in order to evaluate the function \( f(t, y, y^{(1)}) \). Using the initial values of the function evaluations, the iterative procedure uses (3.5) and (3.6) to find the final, required set of function evaluations.

Since the function \( f \) is assumed to be smooth and continuous, the solutions at the nodes during the iterative procedure are assumed to converge upon a trajectory. The final trajectory, the accuracy of the trajectory and the number of iterations required to converge on the trajectory are a function of the stepsize \( h \), the order \( i \) and the function \( f \). The starting algorithm employed by the general and second-sum packages measures the convergence of the solution by calculating the relative norm of the difference between two successive solutions at \( t_q \). If \( z_j \) is defined to be a vector composed of the dependent variables \( y \) and \( y^{(1)} \) at \( t = t_q \) on the \( j \)th iteration, then the relative norm, \( u \), of the difference between two successive iterations at \( t \) is defined to be

\[
||z||
\]

where, for \( k = 1, \ldots, 2n \).
\[ z(k) = \begin{cases} \frac{z_j(k) - z_{j-1}(k)}{z_j(k)} & \text{if } z_j(k) \neq 0 \\ z_{j-1}(k) & \text{if } z_j(k) = 0 \end{cases} \]

and \( n \) is defined to be the number of elements in each of the vectors \( y \) and \( y^{(1)} \). The solution at \( t_q \) is assumed to contain the greatest error and to be the slowest converging solution at any of the nodes because \( t_q \) is generally the furtherest node from \( t_0 \).

Equations (3.5) and (3.6) are the equations used in a starting procedure with the initial conditions at \( t_0 \). This notation does not lend itself to simple application in a computer program. Appendices E and F give the general and second-sum equations equivalent to (3.5) and (3.6), but where the initial conditions are at \( t_m \) and the end nodes are at \( t_n \) and \( t_{n-1+1} \) where \( t_m \) is between \( t_n \) and \( t_{n-1+1} \). It should be noted that the use of (3.5) and (3.6) depends upon the calculation of a special set of coefficients; however, the calculation of the coefficients require little effort and are only required for the starting procedures.
CHAPTER IV

Comparisons and Results

The evaluation of the general and second-sum fixed-mesh/fixed-order Class II multistep integrators was carried out in two phases. The first phase compared the performances of the general and the second-sum integrators to the performances of four Class I integrators and two Class II integrators in solving a selected group of differential equations with periodic solutions. The second phase compared the transverse errors of these integrators for two typical satellite problems with the following gravitational force models. 1) a spherical earth modeled as a point mass and 2) a non-spherical earth modeled by an eleventh degree and order spherical harmonic geopotential. The results of the first and second phases are discussed in Sections IV.1 and IV.2, respectively.

Three forms of the fixed-mesh/fixed-order Class II multistep integrators were adopted in this investigation. The general formulation of the Class II integrator package, referred to as KSGFS, used the back difference form of the PECE algorithm. Two forms of the second-sum formulation were used. The first second-sum package, referred to as SSFSBD, used back differences while the other second-sum package, referred to as SSFSFE, used the function evaluation form of the PECE algorithm.
The results of the general and the second-sum Class II integrators are compared with three documented integration packages and two undocumented integration packages. The documented software packages were ODE by Shampine and Gordon (1975), KROGH by F. T. Krogh (1970) and RK(7)8 as derived by Fehlberg (1968) and applied by McKenzie and Sepehnoori (1978). It should be noted that KROGH may be used as a Class I, Class II or a combination of both Class I and Class II integrators. To distinguish between the first two of these modes, the Class I form and the Class II form of KROGH will be referred to as KROGH1 and KROGH2, respectively. The undocumented packages were ABFS and RKN7(8). ABFS is the Class I equivalent of KSGFS. RKN7(8) is a variable-step Runge-Kutta-Nystrom integrator used for the solution of a general ordinary differential equation of Class II. The coefficients for RKN7(8) were derived and published by Fehlberg (1974).

All computer work was carried out on a CDC Cyber 170/750 computer using the FTN compiler at the University of Texas at Austin under the UT2D operating system. The Cyber 170 computer utilizes a 60-bit word; 12 bits are used for the sign and exponent, and 48 bits are used for the mantissa which results in about 14 decimal digits of accuracy.

IV.1 Integrator Comparisons for Differential Equations
with Periodic Solutions

The first phase of the evaluation of KSGFS, SSFSBD and SSFSFE consisted of comparing their results with those from ODE, KROGH1,
KROGH2, ABFS, RK(7)8 and RKN7(8) for the solution of five sets of differential equations that have periodic solutions. The differential equations and results are discussed in Sections II.1.1 through II.1.5.

The comparisons of the integrators for this phase was accomplished by using the software package COMPAR.* COMPAR allows the user to plot the efficiency curves for each integrator or combinations of integrators. The efficiency curves are defined as the endpoint error versus the number of function evaluations required, endpoint error versus the amount of central processor time required, the maximum global error versus the number of function evaluations and the maximum global error versus the amount of central processor time. The global error at each step is defined as

\[ GE = \text{Maximum} \left[ \frac{x_c(i)-x_i(i)}{b} \right], \text{ } i=1, \ldots, n \]

where \( n \) is the number of elements in the state vector \( \vec{x} \), \( x_c \) is the state vector calculated by the integrator, \( x_r \) is the reference solution and \( b \) is the maximum \( (x_r(i),1) \). The endpoint error is the global error at the final point of the integration interval. By varying the tolerances, the efficiency curves of the variable-step integrators could be determined. To obtain the efficiency curves for ABFS, KSGFS, SSFSBD and SSPSFE, several computer runs were made to determine the optimum order for each integrator on each problem. By using the

* COMPAR was developed at the Department of Aerospace Engineering and Engineering Mechanics, The University of Texas at Austin, by Richard McKenzie. COMPAR is used to compare one or more integrators, each at one or more tolerances, for solving a set of differential equations.
optimum orders and varying the stepsize, the efficiency curves for the fixed-mesh/fixed-order integrators can be found. It should be noted that the starting procedures for ABFS, KSGFS, SSFSBD and SSFSFE were allowed to converge the relative error norm (see Chapter III) as far as possible. Generally, the relative error norm was reduced to zero.

Table IV.1.1 through IV.1.5 are summaries of the results from COMPAR for each of the five problems discussed in Sections IV.1.1 through IV.1.5. The first two columns of these tables are labelled either ABSERR and RELERR or ITER and NORDER. For the variable-step integrators, ABSERR and RELERR are the absolute and relative error tolerances to be used by each integrator. For the fixed-mesh multistep integrators, ITER represents the maximum number of iterations to be used in the starting procedure, and NORDER represents the order of the integration formula to be used. The next three columns in the tables are labelled NFE, NSTPA and NSTPR and are the total number of function evaluations made, the total number of accepted steps by each integrator and the total number of steps rejected by RK(7)8 or RKN7(8). The two columns labelled CP-TIME and OVHD indicate the amount of central processor time used by each integrator for each set of ABSERR and RELERR tolerances. CP-TIME is the total amount of central processor time used by the integrator and the derivative evaluation routines, while OVHD is the overhead or the CP-TIME excluding the time spent in performing derivative evaluations. The remaining columns are self-explanatory.

The problems discussed in Sections IV.1.2 through IV.1.5 were
suggested by Shampine and Gordon (1975). These four problems have non-linear differential equations and "give a good indication of how the codes perform on problems of celestial mechanics."** The harmonic oscillator problem of Section IV.1.1 has the same analytic solution as the circular two-body problem of Section IV.1.2 but has a set of linear differential equations. The elliptic two-body problem of Section IV.1.3 has the same set of differential equations as the circular two-body problem but has a different analytical solution. The Euler rigid-body problem of Section IV.1.4 has a Jacobian elliptic function as the solution. The restricted three-body problem of Section IV.1.5 does not have an analytical solution and presents a case of rapidly changing derivatives near the close approaches.

**IV.1.1 Harmonic Oscillator Problem**

The harmonic oscillator problem may be modeled by the set of first-order differential equations

\[
\frac{dx_1}{dt} = x_2
\]

\[
\frac{dx_2}{dt} = -x_1
\]

or by the second-order differential equation

\[
\frac{d^2x_1}{dt^2} = -x_1
\]

**Shampine and Gordon (1975), page 242.**
With the initial conditions $x_1(0)=0.0$ and $x_2(0)=1.0$, the analytical solution is given by

$$x_1(t) = \sin(t)$$

$$x_2(t) = \cos(t)$$

which has a period of $2\pi$.

Several computer runs were made with the integration interval ranging from 0.0 to $2\pi$, 0.0 to 12$\pi$ and 0.0 to 20$\pi$. By varying the stepsize from 0.5 to 0.3 for these three intervals, the optimum orders for the fixed-mesh/fixed-order multistep integrators were determined to be 8 for ABFS, 15 for KSGFS and 11 for SSFSBD and SSFSFE.

The efficiency curves for this problem were determined for the integration interval of 0.0 to 20$\pi$. The efficiency curves for the variable-step integrators RK(7)8, RKN7(8), ODE, KROGH1 and KROGH2 are shown in Figures IV.1.1a through IV.1.1d. Figures IV.1.1e through IV.1.1h show the efficiency curves for ABFS, KSGFS, SSFSBD and SSFSFE relative to those of ODE, KROGH1 and KROGH2. Figures IV.1.1i through IV.1.1l show the efficiency curves of ABFS, KSGFS, SSFSBD and SSFSFE relative to those for RK(7)8 and RKN7(8). A summary of these results are given in Table IV.1.1.

IV.1.2 Circular Problem of Two Bodies

The circular problem of two bodies may be modeled in two dimensions by the set of first-order differential equations
Figures IV.1.1 c and d

Efficiency Curves for the Harmonic Oscillator Problem
Figures IV.1.1 e and f
Efficiency Curves for the Harmonic Oscillator Problem
Figures IV.1. g and h

Efficiency Curves for the Harmonic Oscillator Problem
Figures IV.1.1 and IV.1.2 show Efficiency Curves for the Harmonic Oscillator Problem.
Figures IV.1.1  k and l
Efficiency Curves for the Harmonic Oscillator Problem
<table>
<thead>
<tr>
<th>METHOD</th>
<th>NSTPR/ AVERAGE</th>
<th>NSTPR/ STEP SIZE</th>
<th>END POINT</th>
<th>GLOBAL MINIMUM</th>
<th>GLOBAL MAXIMUM</th>
<th>START TIME</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERGATION METHOD: SPM</td>
<td>1.00E-01 1.00E-12</td>
<td>257 136</td>
<td>0.00000 3.515E-01</td>
<td>0.07300</td>
<td>0.05616 2.374E-03</td>
<td>3.280E-03</td>
</tr>
<tr>
<td>INTERGATION METHOD: RKS(7)</td>
<td>1.00E-01 1.00E-12</td>
<td>547 42</td>
<td>0.00000 1.496E+00</td>
<td>0.02775</td>
<td>0.01396 7.330E-04</td>
<td>7.300E-04</td>
</tr>
<tr>
<td>INTERGATION METHOD: ODE</td>
<td>1.00E-01 1.00E-12</td>
<td>157 250</td>
<td>0.00000 2.435E-01</td>
<td>0.09800</td>
<td>0.06786 1.956E+05</td>
<td>1.960E+05</td>
</tr>
<tr>
<td>INTERGATION METHOD: ERODE1</td>
<td>1.00E-01 1.00E-12</td>
<td>250 136</td>
<td>0.00000 3.095E-01</td>
<td>0.09550</td>
<td>0.08709 2.400E-05</td>
<td>2.480E-05</td>
</tr>
<tr>
<td>INTERGATION METHOD: ERODE1</td>
<td>1.00E-01 1.00E-12</td>
<td>627 246</td>
<td>0.00000 2.554E-01</td>
<td>0.11300</td>
<td>0.10390 5.776E-08</td>
<td>5.716E-08</td>
</tr>
<tr>
<td>Table IV.1.1</td>
<td>COMPAR Summary of Statistics for the Harmonic Oscillator Problem</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iteration Method</td>
<td>ARFS</td>
<td>KSGFS</td>
<td>SSFSBD</td>
<td>SSFSFE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>------</td>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iter</td>
<td>Order</td>
<td>NFE</td>
<td>Time</td>
<td>NSE</td>
<td>AVERAGE</td>
<td>END POINT</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>8.00E+00</td>
<td>2699</td>
<td>1256</td>
<td>0.00</td>
<td>5.00</td>
<td>0.02</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>8.00E+00</td>
<td>1769</td>
<td>628</td>
<td>0.00</td>
<td>1.00</td>
<td>0.01</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>8.00E+00</td>
<td>1117</td>
<td>502</td>
<td>0.00</td>
<td>1.00</td>
<td>0.01</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>8.00E+00</td>
<td>963</td>
<td>418</td>
<td>0.00</td>
<td>1.00</td>
<td>0.01</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>8.00E+00</td>
<td>773</td>
<td>314</td>
<td>0.00</td>
<td>2.00</td>
<td>0.01</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>8.00E+00</td>
<td>635</td>
<td>251</td>
<td>0.00</td>
<td>2.50</td>
<td>0.01</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>8.00E+00</td>
<td>505</td>
<td>209</td>
<td>0.00</td>
<td>3.00</td>
<td>0.01</td>
</tr>
</tbody>
</table>

### Table IV.1.1

**COMPAR Summary of Statistics for the Harmonic Oscillator Problem**
\[
\frac{dx_1}{dt} = x_3 \\
\frac{dx_2}{dt} = x_4 \\
\frac{dx_3}{dt} = -\frac{x_1}{r^3} \\
\frac{dx_4}{dt} = -\frac{x_2}{r^3}
\]

or by the set of second-order differential equations

\[
\frac{dx_1}{dt^2} = -\frac{x_1}{r^3} \\
\frac{dx_2}{dt^2} = -\frac{x_2}{r^3}
\]

where \( r^2 = x_1^2 + x_2^2 \). For the initial conditions \( x_1(0) = 1.0 \), \( x_2(0) = 0.0 \), \( x_3(0) = 0.0 \), \( x_4(0) = 1.0 \), the analytic solution is given by:

\[
\begin{align*}
    x_1(t) &= \cos(t) \\
    x_2(t) &= \sin(t) \\
    x_3(t) &= -\sin(t) \\
    x_4(t) &= \cos(t)
\end{align*}
\]

The circular two-body orbit is shown in Figure IV.1.2a along with an elliptic two-body orbit of eccentricity 0.6.

By varying the final time of integration to \( 2\pi, 12\pi \) and \( 20\pi \), the optimum orders for ABFS, KSGFS, SSFSBD and SSFSFE were found to be 8, 13, 11, 11, respectively, for a range of step sizes of 0.1 to 0.35.

The efficiency curves for the variable-step integrators ODE, RK(7)8, RKN7(8), KRGH1 and KROGH2 are shown in Figures IV.1.2b through IV.1.2e. The efficiency curves for ABFS, KSGFS, SSFSBD and SSFSFE relative to those for ODE, KROGH1 and KROGH2 are shown in Figures IV.1.2f through
Figure IV.1.2a

Circular ($e=0$) and Elliptic ($e=.6$)
Two Body Orbits
Figures IV.1.2 b and c

Efficiency Curves for the Circular Two Body Problem
Figures IV.1.2 f and g

Efficiency Curves for the Circular Two Body Problem
Figures IV.1.2 j and k

Efficiency Curves for the Circular Two Body Problem
Figures IV.1.2.1 and m
Efficiency curves for the Eccentric Two Body Problem
<table>
<thead>
<tr>
<th>ABERR</th>
<th>RELERR</th>
<th>NF</th>
<th>NSPTA</th>
<th>NSPTR</th>
<th>NSPT</th>
<th>AVERAGE</th>
<th>END POINT</th>
<th>EXPERIMENTAL</th>
<th>GLOBAL</th>
<th>ERROR</th>
<th>EXPERIMENTAL</th>
<th>GLOBAL</th>
<th>ERROR</th>
<th>EXPERIMENTAL</th>
<th>GLOBAL</th>
<th>ERROR</th>
<th>EXPERIMENTAL</th>
<th>GLOBAL</th>
<th>ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00E-04 1.00E-12</td>
<td>590 53</td>
<td>0.00000 1.186±0.0</td>
<td>0.0700</td>
<td>0.05735</td>
<td>2.352E-02</td>
<td>2.372E-02</td>
<td>2.00E-02</td>
<td>1.334±00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00E-06 1.00E-12</td>
<td>598 89</td>
<td>0.00000 7.060±01</td>
<td>0.111000</td>
<td>0.08641</td>
<td>9.599E-04</td>
<td>9.599E-04</td>
<td>2.00E-02</td>
<td>7.502±01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00E-08 1.00E-12</td>
<td>2061 155</td>
<td>0.00000 4.054±01</td>
<td>2.000100</td>
<td>1.4819</td>
<td>2.169E-05</td>
<td>2.169E-05</td>
<td>2.00E-02</td>
<td>2.81±01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00E-10 1.00E-12</td>
<td>3524 271</td>
<td>0.00000 1.339±01</td>
<td>2.250±010</td>
<td>2.2516</td>
<td>4.380E-03</td>
<td>4.380E-03</td>
<td>2.00E-02</td>
<td>2.40±01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00E-12 1.00E-13</td>
<td>4694 361</td>
<td>0.00000 7.401±01</td>
<td>4.8300</td>
<td>3.3831</td>
<td>5.668E-08</td>
<td>5.668E-08</td>
<td>2.00E-02</td>
<td>1.80±01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00E-14 1.00E-13</td>
<td>6241 480</td>
<td>0.00000 1.304±01</td>
<td>0.58300</td>
<td>0.45006</td>
<td>7.646E-09</td>
<td>7.646E-09</td>
<td>2.00E-02</td>
<td>1.35±01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.00E-13 1.00E-14</td>
<td>8308 639</td>
<td>0.00000 9.833±02</td>
<td>0.77300</td>
<td>0.59656</td>
<td>1.051E-09</td>
<td>1.051E-09</td>
<td>2.00E-02</td>
<td>1.01±01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table IV.1.2**

**COMPAR Summary of Statistics for the Circular Two Body Problem**
### Table IV.1.2

COMPAR Summary of Statistics for the Circular Two Body Problem
IV.1.2i. Figures IV.1.2j through IV.1.2m show the efficiency curves for ABFS, KSGFS, SSFSFD and SSFSFE relative to those of RK(7)8 and RKN7(8). To illustrate the efficiency curves, the integration interval was taken to be from 0.0 to $20\pi$. A summary of these results are given in Table IV.1.2.

IV.1.3 Elliptic Problem of Two Bodies

The elliptic problem of two bodies is modeled in two dimensions by the set of first-order differential equations

$$\frac{dx_1}{dt} = x_3$$
$$\frac{dx_2}{dt} = x_4$$
$$\frac{dx_3}{dt} = -\frac{x_1}{r^3}$$
$$\frac{dx_4}{dt} = -\frac{x_2}{r^3}$$

or by the set of second-order differential equations

$$\frac{d^2x_1}{dt^2} = -\frac{x_1}{r^3}$$
$$\frac{d^2x_2}{dt^2} = -\frac{x_2}{r^3}$$

where $r^2 = x_1^2 + x_2^2$. The analytic solution of this problem may be written in terms of the eccentricity, $e$, of the orbit as

$$x_1(t) = \cos u$$
$$x_2(t) = (1-e^2)^{1/2} \sin u$$
$$x_3(t) = -\sin u/(1-e \cos u)$$
$$x_4(t) = (1-e^2)^{1/2} \cos u/(1-e \cos u)$$

where $u$ is found by solving Kepler's equation
\[ M = u - e \sin u \]

and \( M = t \). The eccentricity was chosen to be equal to 0.6, and the initial conditions were

\[ x_1(0) = 0.4, \quad x_2(0) = 0.0, \quad x_3(0) = 0.0 \quad \text{and} \quad x_4(0) = 2.0. \]

Figure IV.1.2a shows the eccentric two-body orbit relative to the circular two-body orbit.

In order to determine the optimum order for ABFS, KSGFS, SSFSBD and SSFSFE, the final time of the integration intervals examined were the same as those used in the circular two-body problem, i.e., 2\( \pi \), 12\( \pi \) and 20\( \pi \). However, due to the relatively large eccentricity of the orbit, the range of stepsize examined was from 0.001 to 0.025, about one-tenth of the values used for the circular two-body problem. The optimum orders were determined to be 12 for ABFS, 12 for KSGFS and 14 for SSFSBD and SSFSFE.

The efficiency curves for the variable-step integrators are shown in Figures IV.1.3a through IV.1.3d. The efficiency curves for ABFS, KSGFS, SSFSBD and SSFSFE relative to those for ODE, KROGH1 and KROGH2 are shown in Figures IV.1.3e through IV.1.3h. Figures IV.2.3i through IV.2.3l show the efficiency curves for ABFS, KSGFS, SSFSBD and SSFSFE relative to those of RK(7)8 and RKN7(8). The integration interval used to illustrate the efficiency curves was taken to be from 0.0 to 20\( \pi \). The data for these figures are summarized in Table IV.1.3.
Figures IV.1.3 a and b

Efficiency Curves for the Eccentric Two Body Problem
Figures IV.1.3 c and d

Efficiency Curves for the Eccentric Two Body Problem
Figures IV.1.3 e and f

Efficiency Curves for the Eccentric Two Body Problem
Figures IV.1.3 8 and h
Efficiency Curves for the Eccentric Two Body Problem
Figures IV.1.3 i and j

Efficiency Curves for the Eccentric Two Body Problem
Figures IV.1.3 k and 1

Efficiency Curves for the Eccentric Two Body Problem
Table IV.1.3

Comparative Summary of Statistics for the Eccentric Two Body Problem
### Table IV.1.3

**COMPAR Summary of Statistics for the Eccentric Two Body Problem**
IV.1.4 Euler Rigid-Body Problem

The Euler equations of motion for a rigid body without external forces may be written as

\[
\begin{align*}
\frac{dx_1}{dt} &= x_2 x_3 \\
\frac{dx_2}{dt} &= -x_1 x_3 \\
\frac{dx_3}{dt} &= -0.51 x_1 x_2 \\
\end{align*}
\]

By defining \( x_4 = \frac{dx_1}{dt} \), \( x_5 = \frac{dx_2}{dt} \) and \( x_6 = \frac{dx_3}{dt} \), the original set of three first-order ordinary differential equations may be extended to six equations. The new set of six differential equations may be written in two forms. The first form is given by the set of first-order differential equations

\[
\begin{align*}
\frac{dx_1}{dt} &= x_2 x_3 \\
\frac{dx_2}{dt} &= -x_1 x_3 \\
\frac{dx_3}{dt} &= -0.51 x_1 x_2 \\
\frac{dx_4}{dt} &= x_1 (x_3^2 + 0.51 x_2^2) \\
\frac{dx_5}{dt} &= x_2 (x_3^2 - 0.51 x_1^2) \\
\frac{dx_6}{dt} &= -0.51 x_3 (x_2^2 - x_1^2) \\
\end{align*}
\]

or the equivalent set of second-order differential equations

\[
\begin{align*}
\frac{d^2 x_1}{dt^2} &= -x_1 (x_3^2 + 0.51 x_2^2) \\
\frac{d^2 x_2}{dt^2} &= -x_2 (x_3^2 - 0.51 x_1^2) \\
\frac{d^2 x_3}{dt^2} &= -0.51 x_3 (x_2^2 - x_1^2) \\
\end{align*}
\]
The second form is given by the set of first-order differential equations

\[
\begin{align*}
\frac{dx_1}{dt} &= x_4 \\
\frac{dx_2}{dt} &= x_5 \\
\frac{dx_3}{dt} &= x_6 \\
\frac{dx_4}{dt} &= x_5 x_3 + x_6 x_2 \\
\frac{dx_5}{dt} &= -(x_4 x_3 + x_1 x_6) \\
\frac{dx_6}{dt} &= -.51(x_4 x_2 + x_5 x_1)
\end{align*}
\]

or the equivalent set of second-order differential equations

\[
\begin{align*}
\frac{d^2x_1}{dt^2} &= x_5 x_3 + x_6 x_2 \\
\frac{d^2x_2}{dt^2} &= -(x_4 x_3 + x_1 x_6) \\
\frac{d^2x_3}{dt^2} &= -.51(x_4 x_2 + x_5 x_1)
\end{align*}
\]

By expanding the state from three to six elements, the right-hand side of the differential equations may be written as a function of \(x_1, x_2\) and \(x_3\) or of \(x_1, x_2, x_3, x_4, x_5\) or \(x_6\). The analytic solutions are given by the Jacobian elliptic functions

\[
\begin{align*}
x_1(t) &= \text{sn}(t; .51) & x_4(t) &= \text{cn}(t; .51)\text{dn}(t; .51) \\
x_2(t) &= \text{cn}(t; .51) & x_5(t) &= -\text{sn}(t; .51)\text{dn}(t; .51) \\
x_3(t) &= \text{dn}(t; .51) & x_6(t) &= -.51 \text{ sn}(t; .51)\text{cn}(t; .51)
\end{align*}
\]

with the initial conditions \(x_1(0) = 0.0, x_2(0) = 1.0, x_3(0) = 1.0,\)
\( x_4(0) = 1.0, \ x_5(0) = 0.0 \) and \( x_6(0) = 0.0 \). The Jacobian elliptic functions are periodic. The functions \( \text{sn}(t|\frac{\pi}{2}) \) and \( \text{cn}(t|\frac{\pi}{2}) \) have a quarter-period of \( K \) and \( \text{dn}(t|\frac{\pi}{2}) \) has a half-period of \( k \) where \( K = 1.86264080233273855203... \). The functions \( \text{sn}(t|\frac{\pi}{2}), \text{cn}(t|\frac{\pi}{2}) \) and \( \text{dn}(t|\frac{\pi}{2}) \) are shown in Figure IV.1.4a. The integration interval was taken to be from \( 0.0 \) to \( 8K \).

To obtain the efficiency curves for the fixed-step/fixed-order integrators, the stepsize was varied from \( 0.03 \) to \( 0.10 \). The optimum orders for ABFS, KSGFS, SSFSBD and SSFSFE were 11, 16, 14 and 14 for the first form of the differential equations and 11, 11, 10 and 10 for the second form of the differential equations.

The efficiency curves presented in Figures IV.1.4b through y are for an integration interval from \( 0.0 \) to \( 8K \). The efficiency curves for solving the first form of the differential equations are shown in Figures IV.1.4b through IV.1.4m. Figures IV.1.4b through IV.1.4e show the efficiency curves for the variable-step integrators. Figures IV.1.4f through IV.1.4i show the efficiency curves of ABFS, KSGFS, SSFSBD and SSFSFE relative to those for ODE, KROGH1 and KROGH2. Figures IV.1.4j through IV.1.4m show the efficiency curves of ABFS, KSGFS, SSFSBD and SSFSFE relative to those for RK7(8) and RKN7(8). Similar comparisons for the solution of the second form of the differential equations are shown in Figures IV.1.4n through IV.1.4y. The results of the comparisons are summarized in Tables IV.1.4A and IV.1.4B.
Jacobian Elliptic Functions:
$\text{sn}(t|\frac{1}{2})$, $\text{cn}(t|\frac{1}{2})$, and $\text{dn}(t|\frac{1}{2})$
Figures IV.1, 4, b and c

Efficiency Curves for the Euler Rigid Body Problem,
Figures IV.1.4 d and e

Efficiency Curves for the Euler Rigid Body Problem.
Figures IV.1.4 f and g

Efficiency Curves for the Euler Rigid Body Problem,
$F=F(T,Y)$
Figures IV.1.4 j and k

Efficiency Curves for the Euler Rigid Body Problem,
F\(=F(T,Y)\)
Figures IV.1.4 1 and m

Efficiency Curves for the Euler Rigid Body Problem.

Pm(T,Y)
Figures IV.1.4 n and o
Efficiency Curves for the Euler Rigid Body Problem,
\( F = F(T, Y, Y^{(1)}) \)
Figures IV.14 p and q
Efficiency Curves for the Euler Rigid Body Problem,
P(F, Y, Y(t))
Figures IV.1.4 r and s
Efficiency Curves for the Euler Rigid Body Problem,
P=F(T,Y,Y(1))
Figures IV.1, 4 t and u
Efficiency Curves for the Euler Rigid Body Problem,
\[ F = F(T, Y, Y^{(1)}) \]
Figures IV.1.4 v and w
Efficiency Curves for the Euler Rigid Body Problem,
\[ F = F(I, Y, Y^{(1)}) \]
Figures IV.1.4 and y:

Efficiency Curves for the Euler Rigid Body Problem,

\[ F(T, Y, Y') \]

(1)
<table>
<thead>
<tr>
<th>METHOD</th>
<th>NSTP</th>
<th>NFF</th>
<th>NSTPA</th>
<th>NSTPT</th>
<th>AVERAGE</th>
<th>END POINT</th>
<th>MINIMUM</th>
<th>GLOBAL</th>
<th>STEP SIZE</th>
<th>STEP SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table IV.1.4A**

**Comparative Summary of Statistics for the Euler Rigid Body Problem F=F(T,Y)**
<table>
<thead>
<tr>
<th>Iter</th>
<th>Order</th>
<th>NFE</th>
<th>NSTPA</th>
<th>NSTPR</th>
<th>Average</th>
<th>Step Size</th>
<th>CP</th>
<th>Time</th>
<th>OvhD</th>
<th>Global Error</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.50E+01</td>
<td>1.00E+01</td>
<td>1114</td>
<td>496</td>
<td>0.00000</td>
<td>3.010E-02</td>
<td>.16400</td>
<td>.3.010E-13</td>
<td>2.560E+13</td>
<td>3.000E-02</td>
<td>1.800E-01</td>
<td>.14377</td>
<td>1.670E-13</td>
<td>2.560E-13</td>
<td>3.000E-02</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>1.00E+01</td>
<td>888</td>
<td>372</td>
<td>0.00000</td>
<td>5.000E-02</td>
<td>.12800</td>
<td>.10869</td>
<td>6.615E-13</td>
<td>7.190E-13</td>
<td>4.000E-02</td>
<td>2.400E-01</td>
<td>.08390</td>
<td>5.403E-11</td>
<td>5.527E-11</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>1.00E+01</td>
<td>740</td>
<td>298</td>
<td>0.00000</td>
<td>5.000E-02</td>
<td>.12500</td>
<td>.10891</td>
<td>5.350E-12</td>
<td>6.121E-12</td>
<td>5.000E-02</td>
<td>3.000E-01</td>
<td>.08329</td>
<td>5.037E-11</td>
<td>5.350E-11</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>1.00E+01</td>
<td>660</td>
<td>248</td>
<td>0.00000</td>
<td>6.000E-02</td>
<td>.09600</td>
<td>.08291</td>
<td>5.202E-10</td>
<td>5.202E-10</td>
<td>5.000E-02</td>
<td>3.000E-01</td>
<td>.08291</td>
<td>5.202E-10</td>
<td>5.202E-10</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>1.00E+01</td>
<td>579</td>
<td>212</td>
<td>0.00000</td>
<td>7.029E-02</td>
<td>.09350</td>
<td>.08291</td>
<td>5.202E-10</td>
<td>5.202E-10</td>
<td>5.000E-02</td>
<td>3.000E-01</td>
<td>.08291</td>
<td>5.202E-10</td>
<td>5.202E-10</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>1.00E+01</td>
<td>527</td>
<td>186</td>
<td>0.00000</td>
<td>8.011E-02</td>
<td>.07750</td>
<td>.06064</td>
<td>1.385E-09</td>
<td>1.417E-09</td>
<td>8.000E-02</td>
<td>4.800E-01</td>
<td>.06821</td>
<td>5.724E-09</td>
<td>5.624E-09</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>1.00E+01</td>
<td>496</td>
<td>165</td>
<td>0.00000</td>
<td>9.031E-02</td>
<td>.07900</td>
<td>.06891</td>
<td>3.185E-08</td>
<td>3.095E-08</td>
<td>6.000E-02</td>
<td>3.000E-01</td>
<td>.06891</td>
<td>3.185E-08</td>
<td>3.095E-08</td>
</tr>
<tr>
<td>2.50E+01</td>
<td>1.00E+01</td>
<td>464</td>
<td>149</td>
<td>0.00000</td>
<td>1.000E-01</td>
<td>.07900</td>
<td>.06891</td>
<td>3.185E-08</td>
<td>3.095E-08</td>
<td>6.000E-02</td>
<td>3.000E-01</td>
<td>.06891</td>
<td>3.185E-08</td>
<td>3.095E-08</td>
</tr>
</tbody>
</table>

**Table IV.1.4A**

**COMPAR Summary of Statistics for the Euler Rigid Body Problem F=F(T,Y)**
### Table IV.1.48

**COMPAR Summary of Statistics for the Euler Rigid Body Problem $F=F(T,Y,Y')$**

<table>
<thead>
<tr>
<th>Integration Method</th>
<th>NFE</th>
<th>NSPAP</th>
<th>NSTPR/ Average</th>
<th>NSTPR/ Step Size</th>
<th>Time</th>
<th>END POINT ERROR</th>
<th>GLOBAL ERROR</th>
<th>MINIMUM</th>
<th>MAXIMUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>RK(7,8)</td>
<td>1.00E-04</td>
<td>1.00E-12</td>
<td>282</td>
<td>17</td>
<td>5</td>
<td>0.27277</td>
<td>7.865E-01</td>
<td>0.02240</td>
<td>0.01706</td>
</tr>
<tr>
<td>ODDE</td>
<td>1.00E-04</td>
<td>1.00E-12</td>
<td>175</td>
<td>87</td>
<td>0</td>
<td>0.00000</td>
<td>1.713E-01</td>
<td>0.05650</td>
<td>0.05119</td>
</tr>
<tr>
<td>FRODDE</td>
<td>1.00E-04</td>
<td>1.00E-12</td>
<td>278</td>
<td>14</td>
<td>0</td>
<td>0.00000</td>
<td>1.057E-01</td>
<td>0.10800</td>
<td>0.10747</td>
</tr>
<tr>
<td>RK(6,8)</td>
<td>1.00E-04</td>
<td>1.00E-12</td>
<td>269</td>
<td>138</td>
<td>0</td>
<td>0.00000</td>
<td>1.080E-01</td>
<td>0.12500</td>
<td>0.11991</td>
</tr>
<tr>
<td>RK(5,6)</td>
<td>1.00E-04</td>
<td>1.00E-12</td>
<td>297</td>
<td>164</td>
<td>0</td>
<td>0.00000</td>
<td>1.185E-01</td>
<td>0.13900</td>
<td>0.13523</td>
</tr>
</tbody>
</table>

Note: The table above presents comparative summary statistics for various integration methods applied to the Euler Rigid Body Problem, indicating the number of function evaluations (NFE), step size (NSTPR), average step size, time, and errors at end points and globally. The minimum and maximum values are also provided for each metric.
<table>
<thead>
<tr>
<th>Iter</th>
<th>Order</th>
<th>NFE</th>
<th>NSTPA</th>
<th>NSTPR/</th>
<th>Average</th>
<th>END POINT</th>
<th>GLOBAL</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>TIME</td>
<td>ERROR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CPE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table IV.1.4B**

COMPAR Summary of Statistics for the Euler Rigid Body Problem $F=F(T,Y,Y')$
IV.1.5 Restricted Three-Body Problem

The restricted three-body problem in two dimensions may be described by the set of first-order differential equations

\[
\begin{align*}
\frac{dx_1}{dt} &= x_3 \\
\frac{dx_2}{dt} &= x_4 \\
\frac{dx_3}{dt} &= x_4 + x_1 - \frac{(1-\mu)(x_1+\mu)}{r_1^3} - \frac{\mu(x_1+\mu-1)}{r_2^3} \\
\frac{dx_4}{dt} &= -2x_3 + x_2 - \frac{x_2(1-\mu)}{r_1^2} - \frac{\mu x_2}{r_2^3}
\end{align*}
\]

where \( r_1^2 = (x_1+\mu)^2 + x_2^2 \), \( r_2^2 = (x_1+\mu-1)^2 + x_2^2 \), and \( \mu \) is the mass ratio. By selecting the initial conditions \( x_1(0)=1.2, x_2(0)=0.0, x_3(0)=0.0 \) and \( x_4(0) = -1.04935750983031990726 \), the mass ratio as \( \mu = 1.0/82.45 \), the period of the orbit becomes \( T = 6.19216933131963970674 \). Although the state at \( T \) is equal to the initial conditions, i.e., \( \bar{x}(T) = \bar{x}(0) \), there is no analytic solution in the interval \( 0 < t < T \). The reference solution for this problem was generated by using ODE with an absolute error tolerance of 10^{-12} and a relative error tolerance of 10^{-12}. The reference orbit is shown in Figure IV.2.5a.

Several computer runs were made in an attempt to determine the optimum order for the fixed-mesh/fixed-order integrators. However, to obtain maximum global errors of less than 1.0, the range of possible stepsizes had to be restricted to less than 0.001. For a stepsize of 0.001, the multistep integrators require approximately 12,500 function evaluations, while the most function evaluations required by any of the
Figure IV.1.5a

Orbit for the Restricted Problem of Three Bodies
Figures IV.1, 5 b and c

Efficiency Curves for the Restricted Problem of Three Bodies
Figures IV.1.5 d and e

Efficiency Curves for the Restricted Problem of Three Bodies
<table>
<thead>
<tr>
<th>Table IV.1.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMPAR Summary of Statistics for the Restricted Problem of Three Bodies</td>
</tr>
</tbody>
</table>
variable-step integrators was 302 and, in general, less than 2000. Since the performances of the fixed-mesh/fixed-order integrators are inferior to the variable-step integrators for this problem, the efficiency curves of the fixed-mesh/fixed-order integrators are not presented. Figures IV.1.5b through IV.1.5e show the efficiency curve for ODE, RK(7)8, RKN7(8), KROGH1 and KROGH2 for an integration of one period. The results of the variable-step integrators are summarized in Table IV.1.5.

IV.1.6 Comments on the COMPAR Results

In general, the efficiency curves indicating the maximum global errors for ODE, KROGH1 and KROGH2 increase dramatically for the more strict absolute and relative error tolerances. The maximum global error for high tolerances occurs within the first few steps in the integration by ODE, KROGH1 or KROGH2, i.e., during the bootstrapping starting procedure of these integrators when the order of the integration formulas is being increased.

The maximum step sizes listed in Tables IV.1.1 through IV.1.4 for the fixed-mesh multistep integrators result from the procedure COMPAR used to calculate the maximum step size and the starting algorithm used by the fixed-mesh integrator. As a result, the maximum step size listed for the fixed-mesh integrators is the difference in the value of the independent variable at the initial conditions and at the last node point in the starting procedure.
IV.2 Comparison of Integration Errors for a Typical Satellite Problem

The major focus of this report is the evaluation of the fixed-mesh/fixed-order multistep integrators for solving the satellite problem. The evaluation is carried out by comparing the results of the fixed-step integrators with the results of the variable-step integrators for two different force models: a two-body model and an eleventh degree and order geopotential model. The comparisons are carried out for an integration interval of 30 days. In order to reduce the number of parameters that could be varied during the evaluation, only one set of initial conditions was examined. This set of initial conditions corresponds to the Lageos satellite. The initial conditions in rectangular coordinates are

\[
\begin{align*}
x &= -9123000.0 \text{ m} & x &= 2708.0 \text{ m/sec} \\
y &= -1894000.0 \text{ m} & y &= 3145.0 \text{ m/sec} \\
z &= -7822000.0 \text{ m} & z &= -3930.0 \text{ m/sec}
\end{align*}
\]

and the osculating set of Keplerian orbital elements are

\[
\begin{align*}
a &= 1.226538 \times 10^7 \text{ m} & \Omega &= .5092370 \text{ radians} \\
e &= .003530571 & \omega &= -1.659311 \text{ radians} \\
i &= 1.918821 \text{ radians} & M &= 5.553645 \text{ radians}
\end{align*}
\]

The comparisons for each force model are carried out in rectangular coordinates, orbital elements and radial-transverse-normal
(RTN) coordinates relative to the reference orbits. The errors are computed by subtracting the integrated solution from the reference solution.

The RTN coordinates are defined relative to the position and velocity vectors of the reference orbit. The radial unit vector is defined to be in the direction of the position vector. The normal unit vector is both parallel to and in the direction of the angular momentum vector. The transverse unit vector is defined to be orthogonal to the normal and radial unit vectors in the approximate direction of the velocity vector.

An example of a set of comparison plots are shown in Figures IV.2.0a through IV.2.0r. The differences illustrated in Figures IV.2.0a through 2.0r are the differences of integrating a spherical eleventh degree and order model with KSGFS using a stepsize of 300 seconds and an order of 14 and integrating the same model with a double-precision RK(7)8 using absolute and relative error tolerances of $10^{-18}$. To compare the six components of each of the three coordinate systems would be cumbersome and probably ambiguous. Since it has been noted that a major portion of the integration error is reflected in the transverse error, the transverse error results were chosen as the basis for the comparisons.
Figures IV.2.0 a, b, c

Example of Integration Comparisons in Rectangular Position Coordinates
Figures IV.2.0 d, e, f
Example of Integration Comparisons in Rectangular Velocity Coordinates
Figures IV.2.0 & 1.4
Example of Integration
Comparison in Semimajor Axis, Eccentricity, and Inclination
Figures IV.2.0 j, k, l
Example of Integration Comparisons in Ascending Node (CAPOMEGA), Argument of Perigee (SMALL OMEGA), and Mean Anomaly
Figures IV.2.0 m, n, o
Example of Integration Comparisons in Radial, Transverse, and Normal Position Coordinates
Figures IV.2.0 p, q, r
Example of Integration Comparisons in Radial, Transverse, and Normal Coordinates
IV.2.1 Two-Body Force Model

The reference orbit for the two-body model was generated by using the analytical solution of the two-body problem using orbital elements. The Keplerian elements $a$, $e$, $i$, $\Omega$ and $\omega$ are constant for the two-body, point-mass problem. The mean anomaly, $M$, was calculated from the equation $M = M_0 + nt$ where $n$ is the mean motion and $M_0$ is the value of the mean anomaly at $t = 0$.

The first-order ordinary differential equations for this problem were

\[
\begin{align*}
\frac{dx_1}{dt} &= x = x_4 \\
\frac{dx_2}{dt} &= y = x_5 \\
\frac{dx_3}{dt} &= z = x_6 \\
\frac{dx_4}{dt} &= -\frac{\mu}{r^3} x_1 = -\frac{\mu}{r^3} x \\
\frac{dx_5}{dt} &= -\frac{\mu}{r^3} x_2 = -\frac{\mu}{r^3} y \\
\frac{dx_6}{dt} &= -\frac{\mu}{r^3} x_3 = -\frac{\mu}{r^3} z
\end{align*}
\]

and the second-order differential equations are

\[
\begin{align*}
\dot{x}_1 &= -\frac{\mu x_1}{r^3} = -\frac{\mu x}{r^3} \\
\dot{x}_2 &= -\frac{\mu x_2}{r^3} = -\frac{\mu y}{r^3} \\
\dot{x}_3 &= -\frac{\mu x_3}{r^3} = -\frac{\mu z}{r^3}
\end{align*}
\]

where $r^2 = x_1^2 + x_2^2 + x_3^2$. 

A plot of the transverse errors of the variable-step integrators ODE, RK(7)8, RKN7(8), KROGH1 and KROGH2 are shown in Figure IV.2.1a. An absolute error tolerance of $10^{-10}$ and a relative error tolerance of $10^{-12}$ was used for ODE, RK(7)8 and RKN7(8), and for KROGH2, an absolute error tolerance of $10^{-9}$ and relative error tolerance of $10^{-11}$ was used. It should be noted that, using the most stringent tolerances allowed, the transverse error for KROGH1 is approximately two orders of magnitude larger than the results from any of the other variable-step integrators. using the most stringent tolerances allowed by KROGH1.

To determine the behavior of the transverse error from the fixed-step integrators, they were applied with a variety of stepsizes and orders to integrate this problem. Table IV.2.1 is a summary of the results and lists the maximum transverse error encountered for each integrator for each combination of order and stepsize. Figure IV.2.1b shows the transverse error curves of ABFS for an integration order of 9 and stepsizes of 50.0, 75.0, 100.0 and 150.0 seconds, while Figure IV.2.1c shows the transverse error curves of KSGFS for an integration order of 13 and stepsizes of 100.0, 200.0 and 250.0 seconds.

IV.2.2 Spherical Eleventh Degree and Order Force Model

This force model does not have an analytical solution and is characterized by the first-order ordinary differential equations
Figure IV.2.1 a

Transverse position error of the variable step and variable mesh integrators for the two body force model with an analytic two body solution.
Table IV.2.1

Maximum Transverse Errors (m)

Two Body Force Model
Thirty Day Integration Interval

<table>
<thead>
<tr>
<th>AbFS</th>
<th>STEPSIZE/ORDER</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50 sec</td>
<td></td>
<td>- .221</td>
<td>-.097</td>
<td>-.034</td>
<td>-.048</td>
</tr>
<tr>
<td>100 sec</td>
<td></td>
<td>-16.90</td>
<td>4.20</td>
<td>-.005</td>
<td>-.048</td>
</tr>
<tr>
<td>150 sec</td>
<td></td>
<td>- 642.</td>
<td>159.</td>
<td>3.65</td>
<td>-.767</td>
</tr>
<tr>
<td>200 sec</td>
<td></td>
<td>-8449.</td>
<td>2035.</td>
<td>85.91</td>
<td>-16.08</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>KSG</th>
<th>STEPSIZE/ORDER</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 sec</td>
<td></td>
<td>- .038</td>
<td>-.041</td>
<td>-.037</td>
<td>-.059</td>
</tr>
<tr>
<td>200 sec</td>
<td></td>
<td>- .004</td>
<td>-.057</td>
<td>-.055</td>
<td>-.043</td>
</tr>
<tr>
<td>300 sec</td>
<td></td>
<td>7.79</td>
<td>-.307</td>
<td>-.224</td>
<td>-.052</td>
</tr>
<tr>
<td>400 sec</td>
<td></td>
<td>260.</td>
<td>11.59</td>
<td>-7.64</td>
<td>-1.55</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SSFSFE</th>
<th>STEPSIZE/ORDER</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 sec</td>
<td></td>
<td>- .053</td>
<td>-.035</td>
<td>-.041</td>
<td>-.024</td>
</tr>
<tr>
<td>200 sec</td>
<td></td>
<td>- .062</td>
<td>-.045</td>
<td>-.050</td>
<td>-.050</td>
</tr>
<tr>
<td>300 sec</td>
<td></td>
<td>- .319</td>
<td>-.077</td>
<td>-.030</td>
<td>-.051</td>
</tr>
<tr>
<td>400 sec</td>
<td></td>
<td>- 8.83</td>
<td>-2.80</td>
<td>.094</td>
<td>.104</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SSFSBD</th>
<th>STEPSIZE/ORDER</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100 sec</td>
<td></td>
<td>- .064</td>
<td>-.046</td>
<td>-.040</td>
<td>-.047</td>
</tr>
<tr>
<td>200 sec</td>
<td></td>
<td>- .042</td>
<td>-.046</td>
<td>-.043</td>
<td>-.042</td>
</tr>
<tr>
<td>300 sec</td>
<td></td>
<td>- .088</td>
<td>-.036</td>
<td>-.042</td>
<td>-.046</td>
</tr>
<tr>
<td>400 sec</td>
<td></td>
<td>- 3.01</td>
<td>-.206</td>
<td>+.108</td>
<td></td>
</tr>
</tbody>
</table>
Figure IV.2.1 b

Comparison of transverse errors from solving the Two Body force model with ABFS (order = 9) using various stepsizes.
Figure IV.2.1 c

Comparison of transverse errors from solving the Two Body model with KSGFS (order = 13) using various stepsizes.
or the second-order ordinary differential equations

\[
\frac{d^2 x_1}{dt^2} = -\frac{\mu x_1}{r^3} + g_1
\]

\[
\frac{d^2 x_2}{dt^2} = -\frac{\mu x_2}{r^3} + g_2
\]

\[
\frac{d^2 x_3}{dt^2} = -\frac{\mu x_3}{r^3} + g_3
\]

where \( r^2 = x_1^2 + x_2^2 + x_3^2 \) and \( g_1, g_2 \) and \( g_3 \) are components of which is the acceleration vector, \( \ddot{g} \), due to a nonspherical earth. The reference orbit for this model was generated by using a double-precision version of RK(7)8 with an absolute error tolerance of \( 10^{-18} \) and a relative error tolerance of \( 10^{-18} \). There were 927,317 function evaluations made during the generation of the reference orbit for an orbital arc of 2,700,000 seconds. To verify that the tolerances used in generating the reference orbit were not too strict, the reference orbit was compared to another integration by the double-precision RK(7)8 that used an absolute error tolerance of \( 10^{-15} \) and a relative tolerance of \( 10^{-15} \). This second integration required 376,897 function
evaluations to integrate an orbital arc of 2,500,000 seconds. The comparison of these two trajectories was done in orbital elements and RTN components which are shown in Figures IV.2.2a through IV.2.21. The differencing of these two comparison trajectories was done in single-precision, and Figures 4.2.2a through 4.2.2f indicate differences in the two orbits at the level of the roundoff error of the computer, except for a small secular difference in the mean anomaly. The differences in the two orbits are magnified when the radial, transverse and normal differences are examined. Despite the growth of the transverse position difference between the orbits, the difference is still relatively small. Thus, from an accuracy and roundoff error point of view, restricting the tolerances to $10^{-18}$ is not inappropriate.

The fixed-step integrators ABFS, KSGFS, SSFSBD and SSFSFE use the PECE algorithm as described in Chapters II and III. It has been noted that if the value of the state does not change significantly from the predicted value to the corrected value, then the value of the perturbing function, $\bar{g}$, may vary only slightly from the prediction step to the correction step. The calculation of $\bar{g}$ generally requires many more arithmetic operations than the two-body terms. The calculation of $\bar{g}$ for each calculation of the two-body terms is referred to as a "full evaluation." However, if $\bar{g}$ does not change significantly between the prediction step and correction step, then $\bar{g}$ need only be calculated during the prediction step. The calculation of $\bar{g}$ only once per integration step is referred to as a "partial evaluation."
Figures IV.2.2 a, b, c
Differences in Semimajor Axis, Eccentricity, and Inclination for Spherical Eleventh Degree and Order Model
Double Precision RK(7)8 using $10^{-15}$ tolerances vs Double Precision RK(7)8 using $10^{-18}$ tolerances
Figures IV.2.2 d, e, f
Differences in Ascending Node, Argument of Perigee, and Mean Anomaly for Spherical Eleventh Degree and Order Model

Double Precision RK(7)8 using $10^{-15}$ tolerances vs Double Precision RK(7)8 using $10^{-7}$ tolerances
Figures IV.2.2 g, h, i
Radial, Transverse, and Normal Position Differences for a Spherical Eleventh Degree and Order Model
Double Precision RK(7)8 using $10^{-15}$ tolerances vs Double Precision RK(7)8 using $10^{-16}$ tolerances
Figures IV.2.2 j,k,l
Radial, Transverse, and Normal Velocity Differences for a Spherical Eleventh Degree and Order Model
Double Precision RK(7)8 using 10^-15 tolerances vs Double Precision RK(7)8 using 10^-18 tolerances
Figure IV.2.2m is a plot of the transverse errors of the variable-step integrators, ODE, RK(7)8, RKN7(8), KROGH1 and KROGH2. The tolerances used for each integrator were

<table>
<thead>
<tr>
<th>INTEGRATOR</th>
<th>ABSOLUTE ERROR TOLERANCE</th>
<th>RELATIVE ERROR TOLERANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ODE</td>
<td>$10^{-10}$</td>
<td>$10^{-12}$</td>
</tr>
<tr>
<td>KROGH1</td>
<td>$10^{-7}$</td>
<td>$10^{-9}$</td>
</tr>
<tr>
<td>KROGH2</td>
<td>$10^{-9}$</td>
<td>$10^{-11}$</td>
</tr>
<tr>
<td>RK(7)8</td>
<td>$10^{-10}$</td>
<td>$10^{-12}$</td>
</tr>
<tr>
<td>RKN7(8)</td>
<td>$10^{-10}$</td>
<td>$10^{-12}$</td>
</tr>
</tbody>
</table>

As in Section IV.2.1, the fixed-step integrators were used with a variety of stepsizes and orders to integrate this problem. Tables IV.2.2A and IV.2.2B are the summaries of the results of each integrator for each pair of order and stepsize. Table IV.2.2A are the results when full evaluations are made, and IV.2.2B are the results when partial evaluations are made.

To illustrate the behavior of the transverse error for each of the fixed-step integrators, Figures IV.2.2n through IV.2.2q are the plots of the transverse errors for ABFS, KSGFS, SSFSBD and SSFSFE with integration orders of 9, 13, 13 and 12, respectively, with a variety of stepsizes and with full evaluations of $\bar{g}$. 
Figure IV.2.2 m

Transverse errors of the variable mesh and variable step integrators solving the spherical eleventh order and degree problem: 1) with KROGH1 and 2) without KROGH1.
Table IV.2.2A

Maximum Transverse Error (m)
Spherical Eleventh Degree and Order Geopotential
Thirty Day Integration Interval
Full Evaluation of Accelerations

<table>
<thead>
<tr>
<th>ABFS</th>
<th>STEPSIZE/ORDER</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50 sec</td>
<td>-0.032</td>
<td>-0.092</td>
<td>-0.026</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td>100 sec</td>
<td>-16.9</td>
<td>4.17</td>
<td>-0.019</td>
<td>-0.053</td>
</tr>
<tr>
<td></td>
<td>150 sec</td>
<td>-639.</td>
<td>158.</td>
<td>3.60</td>
<td>-0.750</td>
</tr>
<tr>
<td></td>
<td>200 sec</td>
<td>-8421.</td>
<td>2026.</td>
<td>85.0</td>
<td>-15.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>KSG</th>
<th>STEPSIZE/ORDER</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 sec</td>
<td>-0.025</td>
<td>-0.056</td>
<td>-0.044</td>
<td>-0.046</td>
</tr>
<tr>
<td></td>
<td>200 sec</td>
<td>-0.019</td>
<td>-0.019</td>
<td>-0.016</td>
<td>-0.031</td>
</tr>
<tr>
<td></td>
<td>300 sec</td>
<td>5.66</td>
<td>0.170</td>
<td>2.60</td>
<td>-0.683</td>
</tr>
<tr>
<td></td>
<td>400 sec</td>
<td>231.</td>
<td>-9.85</td>
<td>14.9</td>
<td>31.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SSFSFE</th>
<th>STEPSIZE/ORDER</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 sec</td>
<td>-0.052</td>
<td>-0.053</td>
<td>-0.046</td>
<td>-0.040</td>
</tr>
<tr>
<td></td>
<td>200 sec</td>
<td>-0.045</td>
<td>-0.039</td>
<td>-0.035</td>
<td>-0.054</td>
</tr>
<tr>
<td></td>
<td>300 sec</td>
<td>-0.191</td>
<td>0.104</td>
<td>0.032</td>
<td>-0.097</td>
</tr>
<tr>
<td></td>
<td>400 sec</td>
<td>-6.51</td>
<td>1.72</td>
<td>3.34</td>
<td>-2.77</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SSFSBD</th>
<th>STEPSIZE/ORDER</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100 sec</td>
<td>-0.044</td>
<td>-0.040</td>
<td>-0.055</td>
<td>-0.039</td>
</tr>
<tr>
<td></td>
<td>200 sec</td>
<td>-0.042</td>
<td>-0.040</td>
<td>-0.045</td>
<td>-0.036</td>
</tr>
<tr>
<td></td>
<td>300 sec</td>
<td>0.053</td>
<td>0.068</td>
<td>0.009</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>400 sec</td>
<td>-0.480</td>
<td>1.68</td>
<td>1.13</td>
<td>*</td>
</tr>
<tr>
<td>Table IV.2.2B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum Transverse Error (m)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spherical Eleventh Degree and Order Geopotential</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thirty Day Integration Interval</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partial Evaluation of Accelerations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STEPSIZE/ORDER</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 sec</td>
<td>-0.025</td>
<td>-0.090</td>
<td>-0.035</td>
<td>-0.045</td>
</tr>
<tr>
<td>100 sec</td>
<td>-16.9</td>
<td>4.17</td>
<td>-0.004</td>
<td>-0.066</td>
</tr>
<tr>
<td>150 sec</td>
<td>-641.</td>
<td>158.</td>
<td>3.65</td>
<td>-0.748</td>
</tr>
<tr>
<td>200 sec</td>
<td>-8431.</td>
<td>2027.</td>
<td>85.9</td>
<td>-15.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STEPSIZE/ORDER</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 sec</td>
<td>-0.040</td>
<td>-0.058</td>
<td>-0.025</td>
<td>-0.031</td>
</tr>
<tr>
<td>200 sec</td>
<td>-0.029</td>
<td>-0.038</td>
<td>-0.019</td>
<td>-0.043</td>
</tr>
<tr>
<td>300 sec</td>
<td>5.05</td>
<td>0.360</td>
<td>2.74</td>
<td>-0.707</td>
</tr>
<tr>
<td>400 sec</td>
<td>196.</td>
<td>-13.7</td>
<td>23.7</td>
<td>35.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STEPSIZE/ORDER</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 sec</td>
<td>-0.060</td>
<td>-0.046</td>
<td>-0.040</td>
<td>-0.045</td>
</tr>
<tr>
<td>200 sec</td>
<td>-0.042</td>
<td>-0.043</td>
<td>-0.041</td>
<td>-0.047</td>
</tr>
<tr>
<td>300 sec</td>
<td>-0.959</td>
<td>0.361</td>
<td>0.207</td>
<td>-0.140</td>
</tr>
<tr>
<td>400 sec</td>
<td>-49.9</td>
<td>-4.34</td>
<td>14.38</td>
<td>2.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>STEPSIZE/ORDER</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>100 sec</td>
<td>-0.054</td>
<td>-0.056</td>
<td>-0.050</td>
<td>-0.039</td>
</tr>
<tr>
<td>200 sec</td>
<td>-0.044</td>
<td>-0.043</td>
<td>-0.045</td>
<td>-0.038</td>
</tr>
<tr>
<td>300 sec</td>
<td>-0.706</td>
<td>0.320</td>
<td>0.165</td>
<td>0.014</td>
</tr>
<tr>
<td>400 sec</td>
<td>-42.7</td>
<td>-4.01</td>
<td>11.9</td>
<td>*</td>
</tr>
</tbody>
</table>
Figure IV.2.2 n
Comparison of transverse errors from solving the spherical eleventh degree and order model with ABFS (order = 9) using various stepsizes.
Figure IV.2.2 o
Comparison of transverse errors from solving the spherical eleventh degree and order model with KSGFS (order = 13) using various step sizes.
Figure IV.2.2 p

Comparison of transverse errors from solving the spherical eleventh degree and order model with SSFSBD (order = 13) using various stepsizes.
Comparison of transverse errors from solving the spherical eleventh degree and order model with SSFSFE (order = 12) using various step sizes.
CHAPTER V
Conclusions

Results from COMPAR give an indication of the relative advantages and disadvantages of the three groups of integrators: 1) variable-step Runge-Kutta integrators RK(7)8 and RKN7(8); 2) variable-mesh/variable-order multistep integrators ODE, KROGH1 and KROGH2; and 3) fixed-mesh/fixed-order multistep integrators ABFS, KSGFS, SSFSBD and SSFSFE.

The harmonic oscillator problem of Section IV.1.1 and the circular two-body problem of Section IV.1.2 have the same analytic solution but have linear and nonlinear differential equations, respectively. However, the relative performance of the three groups of integrators is the same in each problem. The fixed-mesh multistep integrators are the most efficient with respect to the number of function evaluations required and central processor time used. The variable-mesh multistep integrators are competitive with the Runge-Kutta integrators in central processor time used and are more efficient than the number of function evaluations required.

The elliptic two-body problem of Section IV.1.3 illustrates the major differences between the variable-mesh multistep integrators and the single-step Runge Kutta integrators. The variable-mesh multistep integrators require fewer function evaluations than the
variable-step Runge-Kutta integrators to achieve a certain accuracy; however, the variable-mesh multistep integrators require more central processor time. Thus, the variable-mesh multistep integrators require more overhead per function evaluation than the variable-step Runge-Kutta integrators. If the function evaluations of a differential equation are relatively inexpensive in central processor time, the Runge-Kutta integrators would have a distinct advantage; whereas, if the function evaluations are relatively expensive, the variable-mesh multistep integrators would have an advantage. The performance of the fixed-mesh multistep integrators lie between the performances of the other two groups of integrators. The fixed-mesh multistep integrators reduce the overhead associated with the variable-mesh integrators while still requiring fewer function evaluations than the variable-step Runge-Kutta integrators to achieve a certain accuracy.

The relative results of the three groups of integrators for the Euler rigid-body problem of Section IV.1.4 are similar to the results for the eccentric two-body problem. The variable-step Runge-Kutta integrators require the least amount of overhead per function evaluation, while the variable-mesh multistep integrators require the fewest number of function evaluations to achieve a certain accuracy. Again, the fixed-mesh multistep integrators require less overhead per step than the variable-mesh multistep integrators and fewer function evaluations than the variable-step Runge-Kutta integrators. A point of interest with this problem is the relative performance of the Class I/fixed-mesh multistep integrator ABFS with that of the Class
II/fixed-mesh multistep methods. When the derivatives are a function of the states \( x_1, x_2 \) and \( x_3 \) only, ABFS performs more efficiently. However, when the derivatives are a function of all six elements, the Class II methods perform as efficiently as ABFS.

The performance of numerical integration algorithms to solve the satellite problem depends on many parameters, e.g., the distance of the satellite from the primary, the eccentricity of the satellite orbit, the intricacy of the model used to represent the primary. However, by examining one scenario of the satellite problem, the characteristics of the fixed-mesh multistep integrators may be illustrated.

Figures IV.2.1a and IV.2.2m are comparisons of the transverse errors for the variable-step and variable-mesh integrators. Table V.1 is a summary of the approximate number of function evaluations required by each integrator. Figures IV.2.1a and IV.2.2m show that ODE and KROGH2 are more accurate than RK(7)8 and RKN7(8), while Table V.1 shows that ODE and KROGH2 also required significantly fewer function evaluations.

Tables IV.2.1, IV.2.2A and IV.2.2B give the maximum transverse errors of the fixed-mesh multistep integrators for a variety of stepsizes and orders. These tables point out the characteristics of the formulations. First, it is noted that for stepsize less than 300 seconds, the use of "partial evaluations" of the derivatives does not significantly alter the transverse errors. The second point is that the Class II formulations allow larger stepsizes than the Class I
Table V.1

Approximate Number of Function Evaluations
Required by Each Integrator for an Integration Interval of Thirty Days*

<table>
<thead>
<tr>
<th>INTEGRATOR:</th>
<th>TWO BODY MODEL</th>
<th>ELEVENTH DEGREE AND ORDER MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>RK(7)8</td>
<td>168,000</td>
<td>168,000</td>
</tr>
<tr>
<td>RKN7(8)</td>
<td>270,000</td>
<td>271,000</td>
</tr>
<tr>
<td>ODE</td>
<td>44,800</td>
<td>53,900</td>
</tr>
<tr>
<td>KROGH2</td>
<td>32,500</td>
<td>58,900</td>
</tr>
<tr>
<td>KROGH1</td>
<td>30,300</td>
<td>32,800</td>
</tr>
</tbody>
</table>

Fixed mesh multistep integrators with step sizes of:

<table>
<thead>
<tr>
<th>Step Size</th>
<th>Two Body Model</th>
<th>Eleventh Degree and Order Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>75 sec</td>
<td>69,000</td>
<td>69,000</td>
</tr>
<tr>
<td>100 sec</td>
<td>52,000</td>
<td>52,000</td>
</tr>
<tr>
<td>125 sec</td>
<td>41,600</td>
<td>41,600</td>
</tr>
<tr>
<td>200 sec</td>
<td>26,000</td>
<td>26,000</td>
</tr>
<tr>
<td>250 sec</td>
<td>20,900</td>
<td>20,900</td>
</tr>
<tr>
<td>300 sec</td>
<td>17,400</td>
<td>17,400</td>
</tr>
</tbody>
</table>

*The tolerances used by the variable step integrators are given in Sections IV.2.1 and IV.2.2.
formulation to obtain comparable accuracies. Finally, except for one combination of step size and order, there is no appreciable difference between the second-sum formulation using back differences and the second-sum formulation using function evaluations.

The results of Section IV.2 illustrate the advantages of fixed-mesh multistep integrators for solving the satellite problem. The fixed-mesh multistep integrators are capable of being as accurate as the variable-mesh multistep integrators while requiring fewer function evaluations than the variable-mesh integrators. As noted above, the variable-mesh multistep integrators are more efficient than the variable-step Runge-Kutta integrators for solving the satellite problem.

The results of Section IV.1 illustrate that the fixed-mesh multistep methods require less computer time overhead than the variable-mesh integrators. By coupling the results of Sections IV.1 and IV.2, the fixed-mesh multistep integrators are shown to be an attractive tool for use in the satellite problem.
Equations (2.7) and (2.8) are the general formulation basic equations and are used to approximate the state at \( t_{n+r} \) given the state and back differences at \( t_n \). The coefficients, \( \alpha \) and \( \beta \), to be used in (2.7) and (2.8) are determined by the following algorithm:* 

1. Define \( h_I = t_{n+r} - t_n \) 
   \[ \eta_s = \frac{h_I}{h} \] 
   where \( h \) is the integration stepsize.

2. Calculate the matrix of coefficients \( g_{k+q} \), where 
   
   for \( k = 1 \) \( g_{1,q} = \frac{1}{q} \) 
   
   for \( k > 1 \) \( g_{k,q} = \gamma_k g_{k-1,q} - \eta_s g_{k-1,q+1} \) 
   
   where \( k = 1, \ldots, i+1 \) 
   \( q = 1, \ldots, k \) 
   
   \( i \equiv \) number of coefficients required in (2.7) and (2.8) 
   
   \( \psi_k = h + \eta_s - 2 \)

*This fixed-mesh coefficient algorithm is determined from the variable-mesh coefficient algorithm given by Shampine and Gordon (1975).
\[ \gamma_k = \frac{1}{k} \psi_k \]
\[ \eta_k = \frac{1}{k} \eta_s \]

3. The coefficients in (II.7) and (II.8) are given by

\[ \theta_{j,r} = g_{1,j} \]
\[ \alpha_{j,r} = g_{2,j} \]
APPENDIX B

COEFFICIENTS FOR THE SECOND-SUM FORMULATION

Equations (2.24) and (2.25) are the second-sum formulation basic equations and are used to approximate the state at $t_{n+s}$ given the first and second sums and back differences at $t_n$. The coefficients, $a$ and $b$, to be used in (2.24) and (2.25) are determined from the following recursive relationships:

$$b_{j-1,s} = \sum_{k=0}^{j} y'_k y_{j-k,s} \quad j=1,\ldots,i+1$$
$$a_{j-2,s} = \sum_{k=0}^{j} y''_k y_{j-k,s} \quad j=2,\ldots,i+1$$

where $i$ is the number of coefficients required in (2.24) and (2.25) and

$$y'_{k} = -\sum_{q=0}^{k-1} \frac{1}{k-q+1} y'_{q} \quad ; \quad y'_0 = 1$$

$$y''_{k} = \sum_{q=0}^{k} y''_{q} y_{k-q}$$

$$y_{m,s} = \frac{s+m-1}{m} y_{m-1,s} \quad \text{for} \quad m > 0$$

*This is the algorithm for the second sum of coefficients as presented by Spier (1971).
\[ y_{0.5} = 1 \]

and \[ s = \frac{t_{n+s} - t_n}{h} \] where \( h \) is the integration stepsize.
APPENDIX C

MODIFIED, GENERAL FORMULATION, PECE ALGORITHM

Assuming the state \((y_n, y_n^{(1)})\) and the \(i\) back differences of the function \(f(t,y,y^{(1)})\) are known at \(t_n\), then the modified, general formulation, PECE algorithm that is used to advance the solution from \(t_n\) to \(t_{n+1}\) is given by:

1. **Predict** \((y, y^{(1)})\) at \(t_{n+1}\) using an \(i^{th}\) order formula

\[
P_{n+1} = y_n + hy_n^{(1)} + h^2 \sum_{j=1}^{i} \alpha_j y^{j-1} f_n
\]

\[
p^{(1)}_{n+1} = y_n^{(1)} + h \sum_{j=1}^{i} \beta_j y^{j-1} f_n
\]

2. **Evaluate** the function \(f\) with the predicted solution \((p_{n+1}, p^{(1)}_{n+1})\)

\[
r^p_{n+1} = f(t_{n+1}, p_{n+1}, p^{(1)}_{n+1})
\]

3. **Form** the modified back differences

\[
\nu^1 r^p_{n+1} = r^p_{n+1} - d_1
\]
4. Correct the solution using an \((i+1)\) order formula

\[ y_{n+1} = p_{n+1} + h^2 \alpha_{i+1,1} \frac{d_i}{2} f_{n+1} \]

\[ y^{(1)}_{n+1} = p_{n+1} + h \beta_{i+1,1} \frac{d_i}{2} f_{n+1} \]

5. Evaluate the function with the corrected solution

\[ f_{n+1} = f(t_{n+1}, y_{n+1}, y^{(1)}_{n+1}) \]

6. Advance the back differences from \(t_n\) to \(t_{n+1}\)

\[ d_0 = f_{n+1} - d_1 \]

\[ \forall q \leq 1 \quad f_{n+1} = d_q + d_0 : \quad q=1, \ldots, i \]
APPENDIX D

MODIFIED, SECOND-SUM FORMULATION, PECE ALGORITHM

Assuming the second and first sums \((-2f_n, -f_n)\) and the first back differences of the function \(f(t,y,y^{(1)})\) are known at \(t_n\), then the modified, second-sum formulation, PECE algorithm that is used to advance the solution from \(t_n\) to \(t_{n+1}\) is given by:

1. Predict \((y, y^{(1)})\) at \(t_{n+1}\) using an \(i^{th}\) order formula

   \[ p_{n+1} = h^2[V^{-2}f_n + \sum_{j=1}^{1} a_{j-1,1} V^{j-1}f_n] \]

   \[ p_{n+1}^{(1)} = h[V^{-1}f_n + \sum_{j=1}^{1} b_{j-1,1} V^{j-1}f_n] \]

2. Evaluate the function \(f\) with the predicted solution \((p_{n+1}, p_{n+1}^{(1)})\)

   \[ p_{n+1}^p = f(t_{n+1}, p_{n+1}, p_{n+1}^{(1)}) \]

3. Form the modified back differences

   \[ d_k = \sum_{j=k}^{1} V^{j-1}f_n \quad : \quad k=1, \ldots, 1 \]

   \[ V_{n+1}^{d} = r_{n+1}^{p} - d_1 \]
4. Correct the solution using an \((i+1)\) order formula

\[ y_{n+1} = p_{n+1} + h^2 a_{1,1} v^i f_{n+1} \]

\[ y_{n+1}^{(1)} = p_{n+1} + h b_{1,1} v^i f_{n+1} \]

5. Evaluate the function with the corrected solution

\[ f_{n+1} = f(t_{n+1}, y_{n+1}, y_{n+1}^{(1)}) \]

6. Advance the back differences from \(t_n\) to \(t_{n+1}\)

\[ d_0 = f_{n+1} - d_1 \]

\[ v^{q-1} f_{n+1} = d_q + d_0 \quad q = 1, \ldots, i \]

7. Advance the first and second sums from \(t_n\) to \(t_{n+1}\)

\[ v^{-1} f_{n+1} = v^{-1} f_n + f_{n+1} \]

\[ v^{-2} f_{n+1} = v^{-2} f_n + f_{n+1} \]
APPENDIX E

STARTING ALGORITHM FOR THE GENERAL FORMULATION

The purpose of a starting procedure for a multistep integrator is to use the initial conditions \((t_m, y_m, y_m^{(1)})\) and the function \(f(t, y, y^{(1)})\) to approximate the values of \(f\) at \(i\) nodes where \(i\) is the order of the integration formula. The Class II/fixed-mesh general formulation starting algorithm presented here is an iterative procedure where the \(i\) nodes are at \(t_k, k=n, n-1, \ldots, n-i+1\), \(t_m\) is between \(t_n\) and \(t_{n-i+1}\) and \(t_n\) is such that \(\frac{t_n - t_m}{h} > 0\), and \(n, n-i+1 > 0\). The algorithm can be summarized as follows:

1. Evaluate \(f\) at \(t_m\): \(f_m = f(t_m, y_m, y_m^{(1)})\)

2. Use a Taylor series expansion to obtain the first approximation of \(y, y^{(1)}\) and \(f\) at the nodes, e.g.,

\[
\begin{align*}
t_k &= t_m + (k-m)h \\
y_k &= y_m + (k-m)y_m^{(1)} + (k-m)^2 h y_m^{(2)} + (k-m)^2 h f_m \\
y_k^{(1)} &= y_m^{(1)} + (k-m)h f_m \quad k=n, n-1, \ldots, n-i+1 \\
f_k &= f(t_k, y_k, y_k^{(1)})
\end{align*}
\]

where \(h\) is the integration stepsize.
3. Form the back differences of \( f \) at \( t_n \) to obtain 
\[ \psi_{j-1} f_n ; \ j=1, \ldots, i \] and form the vector \( z_p = (y_n, y_n^{(1)}) \).

4. Approximate the solution and evaluate \( f \) at the nodes by using the following interpolation formulas:

\[ y_k = y_m + (k-m)h y_m^{(1)} + h^2 \sum_{j=1}^{i} \frac{1}{2} \alpha_{j,k-m} \psi_{j-1} f_n \]  
\[ y_k^{(1)} = y_m^{(1)} + h \sum_{j=1}^{i} \beta_{j,k-m} \psi_{j-1} f_n \]  
\[ f_k = f(t_k, y_k, y_k^{(1)}) \quad k=n, \ldots, n-i+1 \]

where
\[ \alpha_{j,k-m} = (k-n)^2 \alpha_{j,k-n} - (m-n)^2 \alpha_{j,m-n} + (m-n)(m-k) \beta_{j,m-n} \]
\[ \beta_{j,k-m} = (k-n) \beta_{j,k-n} + (n-m) \beta_{j,m-n} \]

and the coefficients \( \alpha_{j,k} \) and \( \beta_{j,k} \) are discussed in Appendix A.

5. Compute the norm of the relative difference between consecutive values of \((y_n, y_n^{(1)})\)

\[ u^2 = \sum_{j=1}^{q} \left( \frac{z_p(j) - z(j)}{D(j)} \right)^2 \]

where \( z \) is the vector of the most recent values of \((y_n, y_n^{(1)})\), 
\( z_p \) is the vector of the previous values of \((y_n, y_n^{(1)})\).
\[
D(j) = \begin{cases} 
0 & \text{if } z_p(j) = 0 \\
z_p(j) & \text{otherwise}
\end{cases}
\]

and \( q \) is the number of elements in \( z_p \).

6. Finally, if \( u \) is not less than some desired tolerance, then the iterative procedure is continued by repeating Steps 3, 4 and 5; if \( u \) is less than the desired tolerance, then the iterative procedure is complete and the solution at \( t_n \) is computed by setting \( k=n \) in (E.1) and (E.2).

It should be noted that this algorithm may be used for any set of nodes relative to \( t_m \). However, the interpolation formulas which are used to derive (E.1) and (E.2) are generally only valid for \((t_n - t_m)(t_n - t_{n-1+1}) \geq 0\). In the multistep integration packages KSGFS and ABFS, \( t_n \) is chosen so that \( t_m \) is approximately midway between \( t_n \) and \( t_{n-1} \). Also, the convergence criterion discussed in Step 5 may be replaced by any other appropriate criterion.
APPENDIX F
STARTING ALGORITHM FOR THE SECOND-SUM FORMULATION

The purpose of a starting procedure for a multistep integrator is to use the initial conditions \((t_m, y_m, y^{(1)}_m)\) and the function \(f(t, y, y^{(1)})\) to approximate the values of \(f\) at the \(i\) nodes where \(i\) is the order of the integration formula. The Class II/fixed-mesh/second-sum formulation starting algorithm presented here is an iterative procedure where the \(i\) nodes are at \(t_k, k=1, n-1, \ldots, n-i+1\), \(t_m\) is between \(t_n\) and \(t_{n-i+1} > 0\). The algorithm is summarized as follows:

1. Evaluate \(f\) at \(t_m\):
   \[ f_m = f(t_m, y_m, y^{(1)}_m) \]

2. Use a Taylor series expansion to obtain the first approximation of \(y, y^{(1)}\), and \(f\) at the nodes, e.g.,
   \[ t_k = t_m + (k-m)h \]
   \[ y_k = y_m + (k-m)h y^{(1)}_m + (k-m)^2 h^2 f_m \]
   \[ y^{(1)}_k = y^{(1)}_m + (k-m)h f_m \]
   \[ f_k = f(t_k, y_k, y^{(1)}_k) \quad k=n, n-1, \ldots, n-i+1 \]

where \(h\) is the integration stepsize.
3. Form the back differences of \( f \) at \( t_n \) to obtain 
\[ v^{j-1} f_n, \quad j=1, \ldots, i \]
and form the vector \( z_p = (y_n, y_n^{(1)}) \)

4. Approximate the solution and evaluate \( f \) at the nodes by using the following interpolation formulas:

\[
y_k = y_m + (k-m)h y_m^{(1)} + h^2 \sum_{j=1}^{i} \alpha_{j-1,k-m} v^{j-1} f_n \quad \text{(F.1)}
\]

\[
y_k^{(1)} = y_m^{(1)} + h \sum_{j=1}^{i} b_{j-1,k-m} v^{j-1} f_n \quad \text{(F.2)}
\]

\[
f_k = f(t_k, y_k, y_k^{(1)}) ; \quad k=n, \ldots, n-i+1
\]

where

\[
a_{j-1,k-m} = a_{j-1,k-n} - a_{j-1,m-n} - (k-m)b_{j-1,m-n}
\]

\[
b_{j-1,k-m} = b_{j-1,k-n} - b_{j-1,m-n}
\]

and the coefficients \( a_{j,k} \) and \( b_{j,k} \) are discussed in Appendix B.

5. Compute the norm of the relative difference between consecutive values of \( (y_n, y_n^{(1)}) \)

\[
u^2 = \sum_{j=1}^{q} \left( \frac{z_p(j) - z(j)}{D(j)} \right)^2
\]

where \( z \) is the vector of the most recent values of \( (y_n, y_n^{(1)}) \), \( z_p \) is the vector of the previous values of \( (y_n, y_n^{(1)}) \).
\[
D(j) = \begin{cases} 
1 & \text{if } z_p(j) = 0 \\
\sum_{j=1}^{q} z_p(j) & \text{otherwise}
\end{cases}
\]

and \( q \) is the number of elements in \( z_p \).

6. Finally, if \( u \) is not less than some desired tolerance, then the iterative procedure is continued by repeating Steps 3, 4 and 5; if \( u \) is less than the desired tolerance, the iterative procedure is complete and the first and second sums are computed at \( t_n \) by

\[
\begin{align*}
V^{-1}f_n &= \frac{y_m}{h} - \sum_{j=1}^{q} b_{j-1,m-n} V^{j-1}f_n \\
V^{-2}f_n &= \frac{y_m}{h^2} - (m-n-1) V^{-1}f_n + \sum_{j=1}^{q} a_{j-1,m-n} V^{j-1}f_n.
\end{align*}
\]

It should be noted that this algorithm may be used for any set of nodes relative to \( t_m \). However, the interpolation formulas which are used to derive (F.1) and (F.2) are generally only valid for \((t_n - t_m)(t_m - t_{n-i+1}) \geq 0\). In the multistep integration packages SSFSBD and SSFSFE, \( t_n \) is chosen so that \( t_m \) is about midway between \( t_n \) and \( t_{n-i+1} \). Also, the convergence criterion in Step 5 may be replaced by any other appropriate criterion.


