WIS-CS-74-208

COMPUTER SCIENCES DEPARTMENT
University of Wisconsin
1210 West Dayton Street
Madison, Wisconsin 53706

Received February 1, 1974

ENERGY AND MOMENTUM CONSERVING METHODS OF ARBITRARY ORDER FOR THE NUMERICAL INTEGRATION OF EQUATIONS OF MOTION

I. Motion of a Single Particle

by

Robert A. LaBudde
and
Donald Greenspan

Technical Report #208

March 1974

ENERGY AND MOMENTUM CONSERVING METHODS OF ARBITRARY ORDER FOR THE NUMERICAL INTEGRATION OF EQUATIONS OF MOTION

I. Motion of a Single Particle

by

Robert A. LaBudde*

and

Donald Greenspan+

ABSTRACT

Conventional numerical methods, when applied to the ordinary differential equations of motion of classical mechanics, conserve the total energy and angular momentum only to the order of the truncation error. Since these constants of the motion play a central role in mechanics, it is a great advantage to be able to conserve them exactly. A new numerical method is developed, which is a generalization to arbitrary order of the "discrete mechanics" described in earlier work, and which conserves the energy and angular momentum to all orders. This new method can be applied much like a "corrector" as a modification to conventional numerical approximations, such as those obtained via Taylor series, Runge-Kutta, or predictor-corrector formulae. The theory is extended to a system of particles in Part II of this work.

---

*Mathematics Research Center, University of Wisconsin.

+Computer Sciences Department and Academic Computing Center, University of Wisconsin.

TABLE OF CONTENTS

1. Introduction

The theory of the motion of a system of particles is dominated by the concepts of the conservation of the total energy and linear and angular momenta. These quantities are unique in the sense of being the only extensive constants of motion: i.e., simple additive functions of the individual particles [1]. Furthermore, they are directly related to the fundamental transformational invariances of the equations of motion [2].

Because of the importance of these additive constants of motion in mechanics, it is particularly disturbing when their values are changed because of truncation error in conventional numerical solutions. In addition, since the energy and momenta are easily ascertained from the initial conditions, it is reasonable to suppose this information could be used to improve the solution.

In previous work [3]-[4], an unconventional second-order numerical solution to the equations of motion of a system of particles was developed for the case of a separable, radial potential. The method was shown to be of comparable accuracy to the third-order Adams' method, and had the desirable property of exactly conserving the additive constants of motion [4]. The major disadvantage of this new method - denoted "discrete mechanics" - is its low order of approximation: third-order in the coordinates; second-order in the velocities.

In what follows, the "discrete mechanics" presented previously [4] is found to be the lowest-order member of a family of conservative solutions. In addition, the theory to be developed shows it to be actually third-order in the velocities (explaining the comparison in [4] with the third-order Adams' method). The general formula for an arbitrary degree of numerical approximation is developed in Sect. 4 and is compared with results obtained from conventional Adams' methods of orders 3 through 8 in Section 6. Attention will be directed in the present paper to the motion of a single particle: the theory is extended to a system of particles in Part II, to be presented in [5].

As mentioned previously [4], the only requirement for conservation of the total linear momentum (and consequently the center-of-mass motion) is consistency, which is a property of every numerical method of order 2 or higher. Since all the methods to be mentioned (including the conventional Adams' methods) satisfy this criterion, conservation of the total linear momentum will not receive special attention.

2. Definitions

For simplicity, the problem will be restricted initially to the case of the motion of a single particle in a central field. This case includes most common two-body interactions [6]: e.g., the motion of the earth about the run.

Suppose the particle has mass m and position vector

$$\vec{r} = \langle X, Y, Z \rangle \tag{2-1}$$

with respect to the origin of the central field. Let $\vec{v}$ denote the velocity of the particle

$$\vec{v} = \frac{d\vec{r}}{dt} = \langle \frac{dX}{dt}, \frac{dY}{dt}, \frac{dZ}{dt} \rangle \tag{2-2}$$

Newton's second law of motion specifies $\vec{r}(t)$ as the solution of the second-order ordinary differential equation

$$m \frac{d^2\vec{r}}{dt^2} = \vec{F}(\vec{r}) \tag{2-3}$$

where the force field $\vec{F}(\vec{r})$ is given as the negative gradient of a central potential $\phi(r)$:

$$\vec{F}(\vec{r}) = - \frac{\partial \phi(r)}{\partial \vec{r}} \tag{2-4a}$$

$$= - \left(\frac{d\phi}{dr}\right) \frac{\vec{r}}{r} \tag{2-4b}$$

where

$$\frac{\partial}{\partial \vec{r}} = \langle \frac{\partial}{\partial X}, \frac{\partial}{\partial Y}, \frac{\partial}{\partial Z} \rangle \tag{2-5}$$

is the gradient with respect to $\vec{r}$, and

$$r = ||\vec{r}|| = \sqrt{X^2+Y^2+Z^2} \tag{2-6}$$

Therefore the exact solution for the position $\vec{r}'$ and velocity $\vec{v}'$ at any later time $t' = t+\Delta t$ is given by the solution of the differential equations (2-3) and (2-4) with initial conditions $\vec{r}(t)$ and $\vec{v}(t)$.

Since ϕ is differentiable and a function of r alone, the energy E and angular momentum $\vec{L}$ are conserved [2], i.e., are independent of t. For the present problem,

$$E = \frac{1}{2} m(\vec{v}\cdot\vec{v}) + \phi(r) \tag{2-7a}$$

$$\vec{L} = m(\vec{r}\times\vec{v}) \tag{2-7b}$$

where $\cdot$ and $\times$ denote the scalar and vector products, respectively.

Central potentials are the most common form of interactions. Typical examples of $\phi(r)$ that occur in practice are: (a) gravitational,

$$\phi(r) = -\frac{GMm}{r} \tag{2-8}$$

where GM is the gravitational constant for, e.g., the sun; and (b) intermolecular interaction in a dilute gas characterized by, e.g., the Lennard-Jones potential [7],

$$\phi(r) = 4\epsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^{6}] \tag{2-9}$$

where ϵ and σ are constants of the gas involved.

3. Conventional Numerical Solutions

Among the most popular techniques for solving the system of differential eqs. (2-3) and (2-4) numerically are truncated Taylor series, predictor-corrector, and Runge-Kutta formulae. The Taylor series solution is constructed by noting that, when no explicit time dependence occurs in the equations of motion,

$$\frac{d}{dt} = \frac{d\vec{r}}{dt} \cdot \frac{\partial}{\partial \vec{r}} + \frac{d\vec{v}}{dt} \cdot \frac{\partial}{\partial \vec{v}} \tag{3-1a}$$

$$= \vec{v} \cdot \frac{\partial}{\partial \vec{r}} + \frac{\vec{F}}{m} \cdot \frac{\partial}{\partial \vec{v}} \tag{3-1b}$$

Eqs. (3-1) may be used to obtain all higher time-derivatives of $\vec{r}$ in terms of the initial conditions $\vec{r}$ and $\vec{v}$: e.g.,

$$m \frac{d^3\vec{r}}{dt^3} = \frac{d\vec{F}}{dt} \tag{3-2a}$$

$$= -\left(\frac{1}{r}\frac{d\phi}{dr}\right)\vec{v} + \frac{1}{r^2}\left[\frac{1}{r}\frac{d\phi}{dr} - \frac{d^2\phi}{dr^2}\right](\vec{r} \cdot \vec{v})\vec{r} \tag{3-2b}$$

Further derivatives may be developed recursively using eqs. (3-1).

By writing $\vec{r}\,'$ and $\vec{v}\,'$ in the form of Taylor series, one has

$$\vec{r}\,' = \vec{r} + \vec{v}\,\Delta t + \frac{\vec{F}}{m}\frac{(\Delta t)^2}{2} + \frac{d^3\vec{r}}{dt^3}\frac{(\Delta t)^3}{6} + \ldots \tag{3-3a}$$

$$\vec{v}\,' = \vec{v} + \frac{\vec{F}}{m}\Delta t + \frac{d^3\vec{r}}{dt^3}\frac{(\Delta t)^2}{2} + \ldots \tag{3-3b}$$

Then, because vector dependences only on $\vec{r}$ and $\vec{v}$ occur (see (2-4b) and, e.g., eqs. (3-2)), both $\vec{r}\,'$ and $\vec{v}\,'$ can be written in the form

$$\vec{r}\,' = f_1 \vec{r} + f_2 \vec{v} \tag{3-4a}$$

$$\vec{v}\,' = f_3 \vec{r} + f_4 \vec{v} \tag{3-4b}$$

where the f_i are scalar functions. This observation will be used later in Sect. 4.

Predictor-corrector approximations $\vec{r}\,'_c$ and $\vec{v}\,'_c$ for $\vec{r}\,'$ and $\vec{v}\,'$, respectively, can be written in the form

$$\vec{r}\,'_c = \vec{r}\,'_p + \delta\vec{r}_c \tag{3-5a}$$

$$\vec{v}\,'_c = \vec{v}\,'_p + \delta\vec{v}_c \tag{3-5b}$$

where

$$\delta\vec{r}_c = \gamma_0 (\Delta t)^2 (\vec{F}(\vec{r}\,'_c) - \vec{F}'_p)/m \tag{3-6a}$$

$$\delta\vec{v}_c = \gamma_1 \, \Delta t (\vec{F}(\vec{r}\,'_c) - \vec{F}'_p)/m \tag{3-6b}$$

In eqs. (3-5) and (3-6), $\vec{r}\,'_p$, $\vec{v}\,'_p$, and $\vec{F}'_p$ are "predicted" approximate values for $\vec{r}\,'$, $\vec{v}\,'$, and $\vec{F}' = \vec{F}(\vec{r}\,')$, respectively, obtained via interpolation over previous steps. (In the case of an Adams' method, $\vec{r}\,'_p$, $\vec{v}\,'_p$, and $\vec{F}'_p$ are obtained using only $\vec{r}$, $\vec{v}$, and previous values of the forces.) The "corrector" coefficients γ_0 and γ_1 are usually chosen to eliminate the leading terms in the truncation errors of $\vec{r}\,'_p$ and $\vec{v}\,'_p$.

Since $\delta\vec{r}_c$ and $\delta\vec{v}_c$ are implicit in $\vec{r}\,'_c$ via $\vec{F}(\vec{r}\,'_c)$, eqs. (3-5) are implicit in $\vec{r}\,'_c$ and must be solved by an iterative process. Normally this is done by simple successive substitutions, starting with $\vec{r}\,'_c = \vec{r}\,'_p$. If this process is to converge rapidly, $\vec{r}\,'_p$ must be a good approximation to $\vec{r}\,'$. Assuming the method is n-th order, i.e.,

$$\vec{r}\,' = \vec{r}\,'_c + O[(\Delta t)^{n+1}] \qquad (3\text{-}7a)$$

$$\vec{v}\,' = \vec{v}\,'_c + O[(\Delta t)^{n}] \qquad (3\text{-}7b)$$

the "predictors" $\vec{r}\,'_p$ and $\vec{v}\,'_p$ are usually at least (n-1)-st order:

$$\vec{r}\,' = \vec{r}\,'_p + O[(\Delta t)^{n}] \qquad (3\text{-}8a)$$

$$\vec{v}\,' = \vec{v}\,'_p + O[(\Delta t)^{n-1}] \qquad (3\text{-}8b)$$

It should be noted that, if the predictor-corrector system is constructed using only the initial conditions and information (such as the forces) from the differential eqs. (2-3), then $\vec{r}\,'_p$, $\vec{v}\,'_p$, $\vec{r}\,'_c$, and $\vec{v}\,'_c$ are linearly dependent on $\vec{r}$ and $\vec{v}$ in a way similar to eqs. (3-4) for $\vec{r}\,'$ and $\vec{v}\,'$.

As an illustration of the predictor-corrector method, the lowest-order Adams' formulae result for the case $n = 3$, where

$$\vec{r}\,'_p = \vec{r} + \vec{v}\,\Delta t + \frac{\vec{F}}{m}\frac{(\Delta t)^2}{2} \qquad (3\text{-}9a)$$

$$\vec{v}\,'_p = \vec{v} + \frac{\vec{F}}{m}\,\Delta t \qquad (3\text{-}9b)$$

$$\vec{F}\,'_p = \vec{F}(\vec{r}) \qquad (3\text{-}9c)$$

and $\gamma_0 = 1/6$, $\gamma_1 = 1/2$, i.e.,

$$\delta\vec{r}_c = \frac{(\Delta t)^2}{6}(\vec{F}(\vec{r}\,'_c) - \vec{F})/m \qquad (3\text{-}10a)$$

$$\delta\vec{v}_c = \frac{\Delta t}{2}(\vec{F}(\vec{r}\,'_c) - \vec{F})/m \tag{3-10b}$$

Substitution of eqs. (3-9) and (3-10) into eqs. (3-5) gives the following implicit equations for $\vec{r}\,'_c$ and $\vec{v}\,'_c$:

$$\vec{r}\,'_c = \vec{r} + \vec{v}\Delta t + \frac{\vec{F}}{m}\frac{(\Delta t)^2}{2} + \frac{(\Delta t)^2}{6}(\vec{F}(\vec{r}\,'_c) - \vec{F})/m \tag{3-11a}$$

$$\vec{v}\,'_c = \vec{v} + \frac{\vec{F}}{m}\Delta t + \frac{(\Delta t)^2}{2}(\vec{F}(\vec{r}\,'_c) - \vec{F})/m \tag{3-11b}$$

The truncation errors in $\vec{r}\,'_c$ and $\vec{v}\,'_c$ are given by (see, e.g., [8]):

$$\vec{r}\,' = \vec{r}\,'_c - \frac{(\Delta t)^4}{24}\frac{d^4\vec{r}}{dt^4} + O[(\Delta t)^5] \tag{3-12a}$$

$$\vec{v}\,' = \vec{v}\,'_c - \frac{(\Delta t)^3}{12}\frac{d^4\vec{r}}{dt^4} + O[(\Delta t)^4] \tag{3-12b}$$

In general with methods of this type, special starting and step-size changing formulae are necessary, although of course this is not the case for the simple $n = 3$ formulae above.

Runge-Kutta formulae for second-order differential equations, although easily applied, are more complicated in form, and will be of little heuristic value in the discussion that follows. For this reason these methods will not be discussed in any detail in the present work.

4. Conservative Solutions

Since conventional numerical methods of order n conserve the values of E and $\vec{L}$ only to terms of $O[(\Delta t)^n]$, it is of interest to find a method which conserves E and $\vec{L}$ to all orders, i.e., exactly. For these quantities to be maintained at their initial values for all steps, ΔE and $\Delta\vec{L}$ must vanish, where

$$\Delta E = E' - E \tag{4-1a}$$

$$\Delta\vec{L} = \vec{L}' - \vec{L} \tag{4-1b}$$

and $E' = E(t')$ and $\vec{L}' = \vec{L}(t')$. But

$$\Delta E = \frac{1}{2} m(\vec{v}'\cdot\vec{v}' - \vec{v}\cdot\vec{v}) + \Delta\phi \tag{4-2}$$

where

$$\Delta\phi = \phi(r') - \phi(r) \tag{4-3}$$

and

$$\Delta\vec{L} = m(\vec{r}'\times\vec{v}' - \vec{r}\times\vec{v}) \tag{4-4}$$

For conservation of energy and angular momentum, eqs. (4-2) and (4-4) must vanish for all values of $\vec{r}$, $\vec{v}$, and Δt.

The objective in this section is to derive formulae for numerical estimates $\vec{r}'_e$ and $\vec{v}'_e$ for $\vec{r}'$ and $\vec{v}'$ which exactly satisfy the equations $\Delta E = 0$ and $\Delta\vec{L} = \vec{0}$. Suppose that $\vec{r}'_a$ and $\vec{v}'_a$ are approximations to $\vec{r}'$ and $\vec{v}'$, respectively, obtained, e.g., by any of the methods of Sect. 3 above. In analogy with the predictor-corrector methods, define $\delta\vec{r}_e$ and $\delta\vec{v}_e$ by

$$\vec{r}'_e = \vec{r}'_a + \delta\vec{r}_e \tag{4-5a}$$

$$\vec{v}'_e = \vec{v}'_a + \delta\vec{v}_e \tag{4-5b}$$

In the above equations, as with the "correctors" of Sect. 3, $\delta\vec{r}_e$ and $\delta\vec{v}_e$ are of the same orders in Δt as the truncation errors in $\vec{r}'_a$ and $\vec{v}'_a$, respectively.

Two formulations of conservative numerical methods will be developed. In Sect. 4A below, $\vec{r}'_e = \vec{r}'_a$ (i.e., $\delta\vec{r}_e = \vec{0}$) and $\delta\vec{v}_e$ is determined so that conservation of energy and angular momentum results. In Sect. 4B, $\delta\vec{r}_e$ is expressed in terms of $\delta\vec{v}_e$ by assuming a truncated Taylor series for $\vec{r}'$, and $\delta\vec{v}_e$ is determined via conservation of energy and angular momentum.

For the purpose of illustration, the Adams' predictors of eqs. (3-9) will be carried through as examples for $\vec{r}'_a$ and $\vec{v}'_a$, i.e.,

$$\vec{r}'_a = \vec{r} + \vec{v}\,\Delta t + \frac{\vec{F}}{m}\frac{(\Delta t)^2}{2} \tag{4-6a}$$

$$\vec{v}'_a = \vec{v} + \frac{\vec{F}}{m}\,\Delta t \tag{4-6b}$$

Of course, the Adams' correctors of eqs. (3-11) would result for $\vec{r}'_e$ and $\vec{v}'_e$ from the choices for $\delta\vec{r}_e$ and $\delta\vec{v}_e$ from eqs. (3-10), but these values would not satisfy the conservation principles exactly. Note that there is no requirement that $\vec{r}'_a$ and $\vec{v}'_a$ be obtained from explicit formulae: only their values are needed.

A. First Formulation

The equations of motion (2-3) and (2-4) are a system of second-order ordinary differential equations with first-derivatives absent. It is therefore possible to obtain an estimate $\vec{r}'_e$ of $\vec{r}'$ independently of that for $\vec{v}'$. (For example, in eq. (3-11a), $\vec{v}'_c$ does not occur). Suppose

$\vec{r}'_a$ and $\vec{v}'_a$ are approximations to $\vec{r}'$ and $\vec{v}'$, respectively, and take $\vec{r}'_e = \vec{r}'_a$ (i.e., set $\delta\vec{r}_e = 0$); $\vec{v}'_e$ is then given by eq. (4-5b), and the problem is to determine δv_e so that conservation of energy and angular momentum results.

The calculated change $\Delta\vec{L}_e$ in $\vec{L}$ over the time step Δt is given by

$$\Delta\vec{L}_e/m = \vec{r}'_e \times \vec{v}'_e - \vec{r} \times \vec{v} \tag{4-7a}$$

$$= [\vec{r}'_e \times \vec{v}'_a - \vec{r} \times \vec{v}] + \vec{r}'_e \times \delta\vec{v}_e \tag{4-7b}$$

with the unknown part $\delta\vec{v}_e$ of $\vec{v}'_e$ isolated in the second term of eq. (4-7b). Taking the cross product with respect to $\vec{r}'_e$ from the left gives

$$\vec{r}'_e \times \Delta\vec{L}_e/m = [\vec{r}'_e \times (\vec{r}'_e \times \vec{v}'_a) - \vec{r}'_e \times (\vec{r} \times \vec{v})] + \vec{r}'_e \times (\vec{r}'_e \times \delta\vec{v}_e) \tag{4-8}$$

The vector triple products may be decomposed via the formula

$$\vec{a} \times (\vec{b} \times \vec{c}) = (\vec{a} \cdot \vec{c})\vec{b} - (\vec{a} \cdot \vec{b})\vec{c} \tag{4-9}$$

giving for eq. (4-8)

$$\vec{r}'_e \times \Delta\vec{L}_e/m = \vec{\beta} - (r'_e)^2\, \delta\vec{v}_e + \epsilon\vec{r}'_e \tag{4-10}$$

where

$$\vec{\beta} = \vec{r}_e \times (\vec{r}'_e \times \vec{v}'_a) - \vec{r}'_e \times (\vec{r} \times \vec{v}) \tag{4-11a}$$

$$= [(\vec{r}'_e \cdot \vec{v}'_a)\vec{r}'_e - (\vec{r}'_e \cdot \vec{v})\vec{r}\,] - [(r'_e)^2\, \vec{v}'_a - (\vec{r}'_e \cdot \vec{r})\vec{v}] \tag{4-11b}$$

and

$$\epsilon = \vec{r}'_e \cdot \delta\vec{v}_e \tag{4-12}$$

Eq. (4-10) may now be solved for the value of $\delta\vec{v}_e$ that makes

$$\vec{r}'_e \times \Delta\vec{L}_e/m = \vec{0} \tag{4-13}$$

namely

$$\delta\vec{v}_e = \frac{1}{(r'_e)^2} [\epsilon\vec{r}'_e + \vec{\beta}] \tag{4-14}$$

where ϵ is arbitrary and $\vec{\beta}$ is given by eq. (4-11).

For $\delta\vec{v}_e$ given by eq. (4-14), the components of $\Delta\vec{L}_e$ perpendicular to $\vec{r}'_e$ vanish by eq. (4-13), and only

$$\Delta\vec{L}_e/m = (\vec{r}'_e \cdot \Delta\vec{L}_e/m(r'_e)^2)\vec{r}'_e \tag{4-15}$$

remains to be considered.

Since $\delta\vec{v}_e$ enters eq. (4-7b) only via the cross product $\vec{r}'_e \times \delta\vec{v}_e$, this component of $\Delta\vec{L}_e$ in the $\vec{r}'_e$ direction cannot be affected by the choice of $\delta\vec{v}_e$. It will now be shown this component of $\Delta\vec{L}_e$ vanishes from other considerations.

Now,

$$\vec{r}'_e \cdot \Delta\vec{L}_e/m = \vec{r}'_e \cdot [\vec{r}'_e \times \vec{v}'_a - \vec{r} \times \vec{v}] \tag{4-16a}$$

$$= -\vec{r}'_e \cdot (\vec{r} \times \vec{v}) \tag{4-16b}$$

$$= -[\vec{r}'_e \, \vec{r} \, \vec{v}] \tag{4-16c}$$

where $[\vec{r}'_e \, \vec{r} \, \vec{v}]$ denotes the scalar triple product of $\vec{r}'_e$, $\vec{r}$, and $\vec{v}$.

This quantity vanishes, and consequently

$$\vec{r}\,'_e \cdot \Delta\vec{L}_e/m = 0 \tag{4-17}$$

if and only if $\vec{r}\,'_e$, $\vec{r}$, and $\vec{v}$ are linearly dependent. Assuming $\vec{r}\,'_e = \vec{r}\,'_a$ is obtained by using one of the methods mentioned in Sect. 3, it was remarked in Sect. 3 that these estimates of $\vec{r}\,'$ preserve the property of the exact solution of being a linear combination of the initial conditions:

$$\vec{r}\,'_e = f_1\,\vec{r} + f_2\,\vec{v} \tag{4-18}$$

Thus, assuming the estimate $\vec{r}\,'_e$ of $\vec{r}\,'$ has the property given in eq. (4-18), $\vec{r}\,'_e$, $\vec{r}$, and $\vec{v}$ are linearly dependent, and eq. (4-17) is satisfied. Coupled with this assumption, the choice of $\delta\vec{v}_e$ given by eq. (4-14) leads to

$$\Delta\vec{L}_e = \vec{0} \tag{4-19}$$

exactly conserving the angular momentum.

In eq. (4-14), only ϵ remains unknown. The energy change over the step Δt is given by

$$\Delta E_e = \frac{1}{2}\, m\, ((v'_e)^2 - v^2) + \Delta\phi_e \tag{4-20}$$

where

$$\Delta\phi_e = \phi(r'_e) - \phi(r) \tag{4-21}$$

Substitution of eq. (4-14) into eq. (4-5b) and squaring gives

$$\begin{aligned}(v'_e)^2 = (v'_a)^2 + [\epsilon^2 + 2(\vec{r}\,'_e \cdot \vec{v}\,'_a)\epsilon + 2(\vec{\beta} \cdot \vec{v}\,'_a) \\ + \beta^2/(r'_e)^2\,]/(r'_e)^2\end{aligned} \tag{4-22}$$

since, by eq. (4-11), $\vec{r}_e' \cdot \vec{\beta} = 0$. Inserting eq. (4-22) into eq. (4-20) gives

$$\Delta E_e = \frac{m}{2(r_e')^2}\left[\epsilon^2 + 2(\vec{r}_e' \cdot \vec{v}_a')\epsilon + 2(\vec{\beta} \cdot \vec{v}_a') + \beta^2/(r_e')^2\right]$$

$$+ \frac{1}{2}m((v_a')^2 - v^2) + \Delta\phi_e \tag{4-23}$$

Requiring that

$$\Delta E_e = 0 \tag{4-24}$$

gives the following quadratic equation for ϵ:

$$\epsilon^2 + 2(\vec{r}_e' \cdot \vec{v}_a')\epsilon +$$

$$[2(\vec{\beta} \cdot \vec{v}_a') + \beta^2/(r_e')^2 + (r_e')^2((v_a')^2 - v^2 + \frac{2\Delta\phi_e}{m})] = 0 \tag{4-25}$$

This equation may be solved for ϵ by several routes, including the quadratic formula, functional iteration, or Newton's method.

If $\vec{v}_a'$ has an associated error of $O[(\Delta t)^n]$ and $\vec{r}_e'$ is exact to at least this order, then $\delta\vec{v}_e$, $\vec{\beta}$, and ϵ are all of $O[(\Delta t)^n]$, and eq. (4-25) is only weakly quadratic for small Δt. Solving for ϵ via the linear term in eq. (4-25),

$$\epsilon = -\frac{A + \epsilon^2}{B} \tag{4-26}$$

where

$$A = 2(\vec{\beta} \cdot \vec{v}_a') + \beta^2/(r_e')^2 + (r_e')^2((v_a')^2 - v^2 + \frac{2\Delta\phi_e}{m}) \tag{4-27a}$$

$$B = 2\, \vec{r}\,'_e \cdot \vec{v}\,'_a \tag{4-27b}$$

Since A and B are independent of ϵ and $\epsilon = O[(\Delta t)^n]$, eq. (4-26) may be solved iteratively starting with $\epsilon = 0$.

When $\vec{r}\,'_e \cdot \vec{v}\,'_a = 0$, eq. (4-26) breaks down, and there are two equal and opposite solutions to eq. (4-25), given by

$$\epsilon = \pm \sqrt{-A} \tag{4-28}$$

Both solutions are physically meaningful, one corresponding to the forward motion and the other to the backward motion. The correct root can only be determined by consideration of the absolute time scale. This phenomenon occurs because the computed solution contacts a caustic surface of the equations of motion: the exact solution has the same type of ambiguity at the similar point $\vec{r}\,' \cdot \vec{v}\,' = 0$, i.e., a radial turning-point of the motion.

Taken together then, eqs. (4-5b), (4-14), and (4-26) represent a complete system of equations from which $\vec{v}\,'_e$ may be found, given $\vec{r}\,'_e$, which a fortiori conserves energy and angular momentum. Since the exact $\vec{r}\,'$ and $\vec{v}\,'$ also satisfy $\Delta E = 0$ and $\Delta\vec{L} = \vec{0}$, the equations given above also represent a complete system for determining $\vec{v}\,'$ given $\vec{r}\,'$: i.e., if $\vec{r}\,'_e = \vec{r}\,'$ then $\vec{v}\,'_e = \vec{v}\,'$. This absence of truncation error in $\vec{v}\,'_e$ will be discussed more fully in Sect. 5.B.

As an illustration of the use of the above equations, suppose that $\vec{r}\,'_e$ is given by the n = 3 Adams' corrector $\vec{r}\,'_c$ of eq. (3-11a), and $\vec{v}\,'_a$ is given by the corresponding Adams' predictor of eq. (3-9b):

$$\vec{r}\,'_e = \vec{r} + \vec{v}\,\Delta t + \frac{\vec{F}}{m}\frac{(\Delta t)^2}{2} + \frac{(\Delta t)^2}{6}\,(\vec{F}(\vec{r}\,'_e) - \vec{F})/m \tag{4-29a}$$

$$\vec{v}'_a = \vec{v} + \frac{\vec{F}}{m}\Delta t \qquad (4\text{-}29b)$$

The first step in the calculation is the determination of the value for $\vec{r}\,'_e$ by iteration of eq. (4-29a), which requires reevaluation of $\vec{F}(\vec{r}\,'_e)$ every iteration. Once $\vec{r}\,'_e$ is found, and $\vec{v}\,'_a$ formed from eq. (4-29b), $\vec{\beta}$ is calculated from eq. (4-11). The potential term $\phi(r'_e)$ is then evaluated, and the constants A and B calculated from eqs. (4-27).

After these preliminaries, the quantity ϵ is now found by iterating eq. (4-26) to convergence using the starting value $\epsilon = 0$. The iteration to convergence is important because energy is conserved (i.e., $\Delta E = 0$) only to the extent ϵ is found. The value of $\vec{v}\,'_e$ is finally obtained via substitution of ϵ and $\vec{\beta}$ into eqs. (4-5b) and (4-14):

$$\vec{v}\,'_e = \vec{v}\,'_a + \delta\vec{v}_e \qquad (4\text{-}30a)$$

$$= \vec{v}\,'_a + [\epsilon\, \vec{r}\,'_e + \vec{\beta}]/(r'_e)^2 \qquad (4\text{-}30b)$$

The quantities $\vec{r}\,'_e$ and $\vec{v}\,'_e$ represent the computed solution at the time $t' = t + \Delta t$.

B. Second Formulation

In Sect. 4.A above, it was shown that the energy and angular momentum conservation conditions (4-24) and (4-19) were sufficient to determine $\delta\vec{v}_e$ (and hence $\vec{v}\,'_e$) given an estimate $\vec{r}\,'_e$ for $\vec{r}\,'$. This assumption that $\vec{r}\,'_e$ is obtained by some outside numerical approximation, e.g., an Adams' corrector $\vec{r}\,'_c$, led to eqs. (4-14) and (4-26) for $\delta\vec{v}_e$. Since, however, $\delta\vec{v}_e$ is an estimate of the error of $\vec{v}\,'_a$, it is also an estimate of the next term in the Taylor series of $\vec{v}\,'$, i.e., proportional to $d^n\vec{r}/dt^n$. This information could have been used to improve the order of approximation of $\vec{r}\,'$ as follows.

Suppose that $\vec{r}\,'_a$ and $\vec{v}\,'_a$ are approximations to $\vec{r}\,'$ and $\vec{v}\,'$ with errors of $O[(\Delta t)^n]$ and $O[(\Delta t)^{n-1}]$, respectively. Let $\delta\vec{r}_a$ and $\delta\vec{v}_a$ denote the truncation errors of $\vec{r}\,'_a$ and $\vec{v}\,'_a$, respectively:

$$\vec{r}\,' = \vec{r}\,'_a + \delta\vec{r}_a \tag{4-31a}$$

$$\vec{v}\,' = \vec{v}\,'_a + \delta\vec{v}_a \tag{4-31b}$$

Expanding $\delta\vec{r}_a$ and $\delta\vec{v}_a$ in powers of Δt gives

$$\delta\vec{r}_a = A_n(\Delta t)^n \frac{d^n\vec{r}}{dt^n} + O[(\Delta t)^{n+1}] \tag{4-32a}$$

$$\delta\vec{v}_a = B_n(\Delta t)^{n-1} \frac{d^n\vec{r}}{dt^n} + O[(\Delta t)^n] \tag{4-32b}$$

where A_n and B_n are the truncation error coefficients of $\vec{r}\,'_a$ and $\vec{v}\,'_a$, respectively. Examination of eqs. (4-32) shows that, because $\vec{r}\,'$ and $\vec{v}\,'$ have related Taylor series expansions,

$$\delta\vec{r}_a = \gamma(\Delta t)\delta\vec{v}_a + O[(\Delta t)^{n+1}] \tag{4-33}$$

where

$$\gamma = A_n/B_n \tag{4-34}$$

is the ratio of the truncation error coefficients. (Note that if $\vec{r}\,'_a = \vec{r}\,'_p$ and $\vec{v}\,'_a = \vec{v}\,'_p$ are a pair of "predictors", $\gamma = \gamma_0/\gamma_1$ is the ratio of the corresponding "corrector" coefficients). Eq. (4-33) thus gives a lowest-order approximation for $\delta\vec{r}_a$ given $\delta\vec{v}_a$.

Suppose $\vec{r}'_e$ and $\vec{v}'_e$ are to be constructed via

$$\vec{r}'_e = \vec{r}'_a + \delta\vec{r}_e \tag{4-35a}$$

$$\vec{v}'_e = \vec{v}'_a + \delta\vec{v}_e \tag{4-35b}$$

where $\vec{r}'_a$ and $\vec{v}'_a$ are as above. Clearly if $\delta\vec{r}_e = \delta\vec{r}_a$ and $\delta\vec{v}_e = \delta\vec{v}_a$, eqs. (4-35) correspond to eqs. (4-31), and $\vec{r}'_e$ and $\vec{v}'_e$ are exact. Suppose $\delta\vec{v}_e$ is left undetermined, and $\delta\vec{r}_e$ is determined via

$$\delta\vec{r}_e = \gamma\,\Delta t\,\delta\vec{v}_e \tag{4-36}$$

Comparison of eqs. (4-36) and (4-33) shows that if $\delta\vec{r}_e$ approximates $\delta\vec{r}_a$ correctly except for terms of $O[(\Delta t)^n]$ or higher, satisfying eq. (4-36) exactly generates an error of $O[(\Delta t)^{n+1}]$ in $\vec{r}'_e$, since eq. (4-36) corresponds to truncating the error expansion of $\delta\vec{r}_a$:

$$\vec{r}' = \vec{r}'_e + O[(\Delta t)^{n+1}] \tag{4-37}$$

The use of eq. (4-36) will now allow simultaneous determination of $\vec{r}'_e$ and $\vec{v}'_e$ via conservation of energy and angular momentum.

Substitution of eqs. (4-35) and (4-36) into eq. (4-4) gives

$$\Delta\vec{L}_e/m = [\vec{r}'_a \times \vec{v}'_a - \vec{r} \times \vec{v}] + \vec{\alpha} \times \delta\vec{v}_e \tag{4-38}$$

where

$$\vec{\alpha} = \vec{r}'_a - \gamma\,\Delta t\,\vec{v}'_a \tag{4-39}$$

Eq. (4-38) is the analog to eq. (4-7b). Since the form of eq. (4-38) is identical to that of eq. (4-7b), the same method of solution for $\delta\vec{v}_e$ suffices.

Now, since

$$\vec{\alpha} \cdot \Delta\vec{L}_e/m = \vec{0} \tag{4-40}$$

because $\vec{\alpha}$, $\vec{r}_a$, and $\vec{v}_a$ are all linear combinations of $\vec{r}$ and $\vec{v}$ (see discussion following eq. (4-17)), the value of $\delta\vec{v}_e$ which conserves angular momentum is given as the solution to the equation

$$\vec{\alpha} \times \Delta\vec{L}_e/m = \vec{0} \tag{4-41}$$

Similar to eq. (4-10),

$$\vec{\alpha} \times \Delta\vec{L}_e/m = \vec{\beta} - \alpha^2 \delta\vec{v}_e + \epsilon\vec{\alpha} \tag{4-42}$$

where

$$\vec{\beta} = \vec{\alpha} \times [\vec{r}\,'_a \times \vec{v}\,'_a - \vec{r} \times \vec{v}] \tag{4-43a}$$

$$= [(\vec{\alpha} \cdot \vec{v}\,'_a)\,\vec{r}\,'_a - (\vec{\alpha} \cdot \vec{v})\vec{r}\,] - [(\vec{\alpha} \cdot \vec{r}\,'_a)\,\vec{v}\,'_a - (\vec{\alpha} \cdot \vec{r}\,)\,\vec{v}] \tag{4-43b}$$

and

$$\epsilon = \vec{\alpha} \cdot \delta\vec{v}_e \tag{4-44}$$

is unknown. The solution to eq. (4-41) is then given by

$$\delta\vec{v}_e = \frac{1}{\alpha^2} [\epsilon\vec{\alpha} + \vec{\beta}] \tag{4-45}$$

completely analogous to eq. (4-14).

The unknown part of $\delta\vec{v}_e$ is now localized in the quantity ϵ, which will be found from the energy condition. Substitution of eqs. (4-35), (4-36), and (4-45) into eq. (4-2) gives (after setting $\Delta E = 0$) the following equation for ϵ:

$$\epsilon^2 + 2(\vec{\alpha} \cdot \vec{v}'_a)\epsilon + $$
$$+ [2(\vec{\beta} \cdot \vec{v}'_a) + \beta^2/\alpha^2 + \alpha^2((v'_a)^2 - v^2 + \frac{2\Delta\phi}{m})] = 0 \quad (4\text{-}46)$$

where

$$\Delta\phi = \phi(r'_e) - \phi(r) \quad (4\text{-}47)$$

Note that $\Delta\phi$ is implicit in ϵ through r'_e: in contrast to eq. (4-25), eq. (4-46) can be solved <u>only</u> by iterative means. Since $\delta\vec{v}_e$ is $O[(\Delta t)^{n-1}]$ from eq. (4-32b), so are ϵ and $\vec{\beta}$. Therefore eq. (4-46) is only weakly implicit in ϵ (<u>i.e.</u>, to $O[(\Delta t)^{2n-2}]$) and

$$\epsilon = - \frac{A + C(\epsilon)}{B} \quad (4\text{-}48)$$

may be solved via functional iteration starting with $\epsilon = 0$, where

$$A = 2(\vec{\beta} \cdot \vec{v}'_a) + \beta^2/\alpha^2 + \alpha^2(v_a^2 - v^2) \quad (4\text{-}49a)$$

$$B = 2(\vec{\alpha} \cdot \vec{v}'_a) \quad (4\text{-}49b)$$

$$C = 2\alpha^2 \Delta\phi/m + \epsilon^2 \quad (4\text{-}49c)$$

The sequence of steps necessary for computing the conservative solutions $\vec{r}'_e$ and $\vec{v}'_e$ are as follows: first the approximate values $\vec{r}'_a$

and $\vec{v}'_a$ are calculated, and then $\vec{\alpha}$ and $\vec{\beta}$ are evaluated from eqs. (4-39) and (4-43b). The constants A and B are obtained via eqs. (4-49), and C is evaluated using r'_a for r'_e and $\epsilon = 0$. A new approximation is constructed for ϵ using eq. (4-48), $\delta \vec{v}_e$ is formed from eq. (4-45), $\delta \vec{r}_e$ is calculated from eq. (4-36), and $\vec{r}'_e$ from eq. (4-35a). Using the new values for ϵ and r'_e, C is reevaluated, and a new iteration of eq. (4-48) started. This process is continued until a convergent value for ϵ is obtained. The corresponding values of $\vec{r}'_e$ and $\vec{v}'_e$ will conserve angular momentum to within round-off error, and energy to within the degree of convergence in ϵ.

For the special case when n = 2 and

$$\vec{r}'_a = \vec{r} + \vec{v}\,\Delta t \tag{4-50a}$$

$$\vec{v}'_a = \vec{v} \tag{4-50b}$$

the truncation errors are given by

$$\vec{r}' = \vec{r}'_a + \frac{(\Delta t)^2}{2} \frac{d^2\vec{r}}{dt^2} + O[(\Delta t)^3] \tag{4-51a}$$

$$\vec{v}' = \vec{v}'_a + \Delta t \frac{d^2\vec{r}}{dt^2} + O[(\Delta t)^2] \tag{4-51b}$$

and so $\gamma = 1/2 \div 1 = 1/2$, and

$$\vec{\alpha} = \vec{r} + \frac{\Delta t}{2} \vec{v} \tag{4-52a}$$

$$\vec{\beta} = \vec{0} \tag{4-52b}$$

These values of $\vec{\alpha}$ and $\vec{\beta}$ lead to

$$\delta\vec{v}_e = \epsilon^* \vec{\alpha} \tag{4-53}$$

where

$$\epsilon^* = \epsilon/\alpha^2 \tag{4-54}$$

The value of ϵ^* is obtained via iteration of

$$\epsilon^* = -\frac{2\Delta\phi/m + \alpha^2(\epsilon^*)^2}{2(\vec{\alpha}\cdot\vec{v})} \tag{4-55}$$

This n = 2 conservative method represented by eqs. (4-50), (4-52), (4-53), and (4-55) is equivalent to the "second-order discrete mechanics" of [4], where instead of $\delta\vec{v}_e$ the discrete mechanics "force" $\vec{F}^*$ was defined: the transformation is given by

$$\delta\vec{v}_e = \Delta t\, \vec{F}^*/m \tag{4-56}$$

As an additional illustration, consider $\vec{r}\,'_a$ and $\vec{v}\,'_a$ given by the n = 3 predictors $\vec{r}_p$ and $\vec{v}_p$ of eqs. (3-9):

$$\vec{r}\,'_a = \vec{r} + \vec{v}\,\Delta t + \frac{\vec{F}}{m}\frac{(\Delta t)^2}{2} \tag{4-57a}$$

$$\vec{v}\,'_a = \vec{v} + \frac{\vec{F}}{m}\Delta t \tag{4-57b}$$

For this case, $\gamma = 1/6 \div 1/2 = 1/3$ from eqs. (3-10), and

$$\vec{\alpha} = \vec{r}\,'_a - \frac{\Delta t}{3}\vec{v}\,'_a \tag{4-58a}$$

$$= \vec{r} + \frac{2\Delta t}{3}\vec{v} + \frac{(\Delta t)^2}{6}\frac{\vec{F}}{m} \tag{4-58b}$$

The value of $\vec{\beta}$ is best obtained via substitution of eqs. (4-57) and (4-58) into eq. (4-43b). The values of A and B are obtained from eqs. (4-49), and eq. (4-48) is iterated to find ϵ. At each stage the iterate for ϵ is found, $\delta\vec{v}_e$ formed, and $\vec{r}\,'_e$ calculated from

$$\delta\vec{r}_e = \frac{\Delta t}{3}\,\delta\vec{v}_e \tag{4-59}$$

and eq. (4-35a). The magnitude r'_e is then found, and $\phi(r'_e)$ evaluated. A new iteration is then performed.

5. Related Theory and Remarks

The formulations given in both Sects. 4.A and 4.B conserve the energy and angular momentum at their initial values - an advantage not enjoyed by conventional methods. In addition, these conservative methods enjoy improved truncation error properties, and close ties with the exact solution, which will be discussed in the next sections.

A. Truncation Error

For the first conservative formulation of Sect. 4.A, the value of $\vec{r}\,'_e$ obtained is of course identical to that of the underlying approximation $\vec{r}\,'_a$, and has the associated truncation error. For example, if $\vec{r}\,'_a = \vec{r}\,'_c$, the n = 3 corrector of eq. (3-11a), then by eq. (3-12a),

$$\vec{r}\,' = \vec{r}\,'_e - \frac{(\Delta t)^4}{24} \frac{d^4\vec{r}}{dt^4} + O[(\Delta t)^5] \qquad (5\text{-}1)$$

Note, that since no approximations were made in deriving eqs. (4-14) and (4-26) for $\vec{v}\,'_e$, except that the $\vec{r}\,'_e$ used was an inexact estimate of $\vec{r}\,'$, there is no associated truncation error in $\vec{v}\,'_e$. The only error in $\vec{v}\,'_e$ is that propagated from that of the previous steps and the truncation error in $\vec{r}\,'_e$: this will be discussed more fully in Sect. 5.B.

For the case of the second conservative formulation of Sect. 4.B, eq. (4-36) holds for the exact $\delta\vec{r}_a$ and $\delta\vec{v}_a$ only if $\vec{r}\,'$ is a polynomial of degree n or less in Δt (compare eqs. (4-33) and (4-36)). Therefore the use of eq. (4-36) in general corresponds to truncating the Taylor series expansion of $\vec{r}\,'$ at the n-th term, and generates a truncation error of $O[(\Delta t)^{n+1}]$:

$$\vec{r}\,' = \vec{r}\,'_e + O[(\Delta t)^{n+1}] \tag{5-2}$$

Once eq. (4-36) was assumed, the derivation of $\vec{v}\,'_e$ followed just as in Sect. 4.A , with no further approximations: As in Sect. 4.A , there is no associated truncation error in $\vec{v}\,'_e$. The only error in $\vec{v}\,'_e$ is again that propagated from previous steps or from the error in $\vec{r}\,'_e$.

Knowing that $\vec{r}\,'_e$ as calculated from the second formulation of Sect. 4.B satisfies eq. (5-2), the corresponding local truncation error coefficient E_{n+1} will now be derived, where

$$\vec{r}\,' = \vec{r}\,'_e + E_{n+1}(\Delta t)^{n+1} \frac{d^{n+1}\vec{r}}{dt^{n+1}} + O[(\Delta t)^{n+2}] \tag{5-3}$$

Suppose that $\vec{r}\,'_a$ and $\vec{v}\,'_a$ have truncation errors given by eqs. (4-32), i.e.,

$$\delta\vec{r}_a = \vec{r}\,' - \vec{r}\,'_a \tag{5-4a}$$

$$= A_n(\Delta t)^n \frac{d^n\vec{r}}{dt^n} + A_{n+1}(\Delta t)^{n+1} \frac{d^{n+1}\vec{r}}{dt^{n+1}} + O[(\Delta t)^{n+2}] \tag{5-4b}$$

and

$$\delta\vec{v}_a = \vec{v}\,' - \vec{v}\,'_a \tag{5-5a}$$

$$= B_n(\Delta t)^{n-1} \frac{d^n\vec{r}}{dt^n} + B_{n+1}(\Delta t)^n \frac{d^{n+1}\vec{r}}{dt^{n+1}} + O[(\Delta t)^{n+1}] \tag{5-5b}$$

where A_n, A_{n+1} and B_n, B_{n+1} are the truncation error coefficients of $\vec{r}\,'_a$ and $\vec{v}\,'_a$, respectively. Comparison of eqs. (4-31) and (4-35), together with eqs. (5-4) and (5-5), leads to

$$\delta\vec{r}_e = O[(\Delta t)^n] \tag{5-6a}$$

$$\delta\vec{v}_e = O[(\Delta t)^{n-1}] \tag{5-6b}$$

Now suppose further that the only error introduced in the calculation of $\vec{v}'_e$ and $\delta\vec{v}_e$ in Sect. 4.B was that due to the use of eq. (4-36): i.e., the use of an inexact value for $\vec{r}'_e$. Since by eq. (5-2) the error in $\vec{r}'_e$ is $O[(\Delta t)^{n+1}]$, this induces (by expanding $\phi(r'_e)$ as a Taylor series and by eqs. (4-43), (4-45), and (4-47)) an error of the same order in $\delta\vec{v}_e$:

$$\delta\vec{v}_a = \delta\vec{v}_e + O[(\Delta t)^{n+1}] \tag{5-7}$$

Comparing eqs. (5-7) and (5-5b) gives

$$\delta\vec{v}_e = B_n(\Delta t)^{n-1} \frac{d^n\vec{r}}{dt^n} + B_{n+1}(\Delta t)^n \frac{d^{n+1}\vec{r}}{dt^{n+1}} + O[(\Delta t)^{n+1}] \tag{5-8}$$

The form of $\delta\vec{r}_e$ may now be obtained through the use of eq. (4-36):

$$\delta\vec{r}_e = \gamma(\Delta t)\delta\vec{v}_e \tag{5-9a}$$

$$= \gamma B_n(\Delta t)^n \frac{d^n\vec{r}}{dt^n} + \gamma B_{n+1}(\Delta t)^{n+1} \frac{d^{n+1}\vec{r}}{dt^{n+1}} + O[(\Delta t)^{n+2}] \tag{5-9b}$$

Noting that $\gamma = A_n/B_n$,

$$\delta\vec{r}_e = A_n(\Delta t)^n \frac{d^n\vec{r}}{dt^n} + \frac{A_nB_{n+1}}{B_n}(\Delta t)^{n+1} \frac{d^{n+1}\vec{r}}{dt^{n+1}} + O[(\Delta t)^{n+2}] \tag{5-10}$$

Comparing eq. (5-10) with the exact eq. (5-4b) for $\delta\vec{r}_a$ gives

$$\delta\vec{r}_a = \delta\vec{r}_e + E_{n+1} (\Delta t)^{n+1} \frac{d^{n+1}\vec{r}}{dt^{n+1}} + O[(\Delta t)^{n+2}] \tag{5-11}$$

where

$$E_{n+1} = A_{n+1} - A_n B_{n+1}/B_n \tag{5-12}$$

Finally, eqs. (5-7) and (5-11) may be translated via eqs. (4-35) into error equations in $\vec{r}\,'_e$ and $\vec{v}\,'_e$:

$$\vec{r}\,' = \vec{r}\,'_e + E_{n+1} (\Delta t)^{n+1} \frac{d^{n+1}\vec{r}}{dt^{n+1}} + O[(\Delta t)^{n+2}] \tag{5-13a}$$

$$\vec{v}\,' = \vec{v}\,'_e + O[(\Delta t)^{n+1}] \tag{5-13b}$$

where E_{n+1} is given by eq. (5-12).

As an example of the use of eqs. (5-13), consider the $n = 2$ conservative method of eqs. (4-50), (4-51), and following equations of Sect. 4.B. This method, as remarked previously, corresponds to the "discrete mechanics" of [4]. For $\vec{r}\,'_a$ and $\vec{v}\,'_a$ given by eqs. (4-51),

$$\vec{r}\,' = \vec{r}\,'_a + \frac{(\Delta t)^2}{2} \frac{d^2\vec{r}}{dt^2} + \frac{(\Delta t)^3}{6} \frac{d^3\vec{r}}{dt^3} + O[(\Delta t)^4] \tag{5-14a}$$

$$\vec{v}\,' = \vec{v}\,'_a + \Delta t \frac{d^2\vec{r}}{dt^2} + \frac{(\Delta t)^2}{2} \frac{d^3\vec{r}}{dt^3} + O[(\Delta t)^3] \tag{5-14b}$$

or

$$A_2 = 1/2 \tag{5-15a}$$

$$A_3 = 1/6 \tag{5-15b}$$

$$B_2 = 1 \tag{5-15c}$$

$$B_3 = 1/2 \tag{5-15d}$$

and

$$\gamma = A_2/B_2 \tag{5-16a}$$

$$= 1/2 \tag{5-16b}$$

Substitution of eqs. (5-15) into eqs. (5-12) and (5-13) give ($n = 2$)

$$\vec{r}\,' = \vec{r}\,'_e - \frac{(\Delta t)^3}{12}\frac{d^3\vec{r}}{dt^3} + O[(\Delta t)^4] \tag{5-17a}$$

$$\vec{v}\,' = \vec{v}\,'_e + O[(\Delta t)^3] \tag{5-17b}$$

and so the "discrete mechanics" of [4] is a <u>third</u>- not second-order method. This result explains the effectiveness of "discrete mechanics" when compared to the third-order Adams' method of eqs. (3-11) in [4a].

As a further illustration, consider the $n = 3$ formulae of eqs. (4-57) for $\vec{r}\,'_a$ and $\vec{v}\,'_a$. For these predictors

$$\vec{r}\,' = \vec{r}\,'_a + \frac{(\Delta t)^3}{6}\frac{d^3\vec{r}}{dt^3} + \frac{(\Delta t)^4}{24}\frac{d^4\vec{r}}{dt^4} + O[(\Delta t)^5] \tag{5-18a}$$

$$\vec{v}\,' = \vec{v}\,'_a + \frac{(\Delta t)^2}{2}\frac{d^3\vec{r}}{dt^3} + \frac{(\Delta t)^3}{6}\frac{d^4\vec{r}}{dt^4} + O[(\Delta t)^4] \tag{5-18b}$$

or

$$A_3 = 1/6 \tag{5-19a}$$

$$A_4 = 1/24 \tag{5-19b}$$

$$B_3 = 1/2 \tag{5-19c}$$

$$B_4 = 1/6 \tag{5-19d}$$

Combined with eqs. (5-12) and (5-13), eqs. (5-19) give

$$\vec{r}\,' = \vec{r}\,'_e - \frac{(\Delta t)^4}{72} \frac{d^4 \vec{r}}{dt^4} + O[(\Delta t)^5] \tag{5-20a}$$

$$\vec{v}\,' = \vec{v}\,'_e + O[(\Delta t)^4] \tag{5-20b}$$

for the errors in the corresponding conservative solutions $\vec{r}\,'_e$ and $\vec{v}\,'_e$ of Sect. 4.B. Note that the use of the Adams' corrector formula of eq. (3-10a) leads to a truncation error coefficient of $-1/24$ in eq. (3-12a), while the conservative solution of Sect. 4.B gives a smaller truncation error coefficient of $-1/72$ in eq. (5-20a). In addition, $\vec{v}\,'_e$ is a <u>fourth-order</u> approximation to $\vec{v}\,'$, while $\vec{v}\,'_c$ of eq. (3-11b) is only a <u>third-order</u> approximation to $\vec{v}\,'$. Thus the conservative formulation of Sect. 4.B, when used with the Adams' predictors $\vec{r}\,'_p$ and $\vec{v}\,'_p$ for $\vec{r}\,'_a$ and $\vec{v}\,'_a$, is superior to the Adams' correctors from the standpoint of order, truncation error, and conservation of the constants of motion.

B. Propagation of Error

In the case of the predictor-corrector methods of eqs. (3-5), the explicit dependences of $\delta\vec{r}_c$ and $\delta\vec{v}_c$ on Δt in eqs. (3-6) guarantees that any error (except round-off) made in evaluating these quantities will be unimportant in the limit as $\Delta t \to 0$. Suppose errors $\vec{\eta}(\vec{r}\,'_p) = \vec{r}\,' - \vec{r}\,'_p$ and $\vec{\eta}(\vec{v}\,'_p) = \vec{v}\,' - \vec{v}\,'_p$ are introduced (via truncation error, error in the initial conditions, etc.) in $\vec{r}\,'_p$ and $\vec{v}\,'_p$, respectively. Then the propagated errors in $\vec{r}\,'_c$ and $\vec{v}\,'_c$, respectively, are given by

$$\vec{\eta}(\vec{r}\,'_c) = \vec{\eta}(\vec{r}\,'_p) + O[(\Delta t)^2] \tag{5-21a}$$

$$\vec{\eta}(\vec{v}\,'_c) = \vec{\eta}(\vec{v}\,'_p) + O[\Delta t] \tag{5-21b}$$

In the case of an Adams' method

$$\vec{r}\,'_p = \vec{r} + \vec{v}\Delta t + O[(\Delta t)^2] \tag{5-22a}$$

$$\vec{v}\,'_p = \vec{v} + O[\Delta t] \tag{5-22b}$$

and hence

$$\vec{\eta}(\vec{r}\,'_p) = \vec{\eta}(\vec{r}) + O[\Delta t] \tag{5-23a}$$

$$\vec{\eta}(\vec{v}\,'_p) = \vec{\eta}(\vec{v}) + O[\Delta t] \tag{5-23b}$$

so that, from eqs. (5-21),

$$\vec{\eta}(\vec{r}\,'_c) = \vec{\eta}(\vec{r}) + O[\Delta t] \tag{5-24a}$$

$$\vec{\eta}(\vec{v}\,'_c) = \vec{\eta}(\vec{v}) + O[\Delta t] \tag{5-24b}$$

giving the errors in $\vec{r}'_c$ and $\vec{v}'_c$ in terms of those in $\vec{r}$ and $\vec{v}$, respectively. Therefore, for the case of an Adams' method, in the limit $\Delta t \to 0$ the errors in $\vec{r}$ and $\vec{v}$ are propagated with unit coefficients in $\vec{r}'_c$ and $\vec{v}'_c$, respectively. This is usually phrased in terms of the statement that for an Adams' method, the principal roots of the asymptotic ($\Delta t \to 0$) stability matrix have value unity.

Because the forms of the Taylor series and Runge-Kutta methods are similar to eqs. (5-22), this property of unit feedback of the errors in $\vec{r}$ and $\vec{v}$ into the calculated values for $\vec{r}'$ and $\vec{v}'$ (for small Δt) also holds for these methods.

The question is now whether this desirable property of unit (and no more) propagation of the errors in $\vec{r}$ and $\vec{v}$ occurs with the conservative formulations of Sect. 4 for small Δt. Suppose, as above, that $\vec{r}'_a$ and $\vec{v}'_a$ satisfy equations similar to (5-22):

$$\vec{r}'_a = \vec{r} + \vec{v}\Delta t + O[(\Delta t)^2] \tag{5-25a}$$

$$\vec{v}'_a = \vec{v} + O[\Delta t] \tag{5-25b}$$

and errors $\vec{\eta}(\vec{r})$ and $\vec{\eta}(\vec{v})$ are introduced into $\vec{r}$ and $\vec{v}$, respectively.

Consider first the formulation of Sect. 4.A. Then $\vec{r}'_e = \vec{r}'_a$, and by expanding in Taylor series,

$$\vec{\eta}(\vec{r}'_e) = \left(\frac{\partial \vec{r}'_e}{\partial \vec{r}}\right)_{\vec{v}} \cdot \vec{\eta}(\vec{r}) + \left(\frac{\partial \vec{r}'_e}{\partial \vec{v}}\right)_{\vec{r}} \cdot \vec{\eta}(\vec{v}) + O[\eta^2] \tag{5-26a}$$

$$= \left(\frac{\partial \vec{r}'_a}{\partial \vec{r}}\right)_{\vec{v}} \cdot \vec{\eta}(\vec{r}) + \left(\frac{\partial \vec{r}'_a}{\partial \vec{v}}\right)_{\vec{r}} \cdot \vec{\eta}(\vec{v}) + O[\eta^2] \tag{5-26b}$$

where $\partial\vec{a}/\partial\vec{b}$ denotes the Jacobian matrix of $\vec{a}$ with respect to $\vec{b}$, with elements

$$\left(\frac{\partial\vec{a}}{\partial\vec{b}}\right)_{ij} = \frac{\partial a_i}{\partial b_j} \tag{5-27}$$

and $(\partial\vec{r}\,'_a/\partial\vec{r})_{\vec{v}}$ and $(\partial\vec{r}\,'_a/\partial\vec{v})_{\vec{r}}$ denote derivatives of $\vec{r}\,'_a$ with respect to $\vec{r}$ and $\vec{v}$, respectively, with $\vec{r}\,'_a$ and $\vec{v}\,'_a$ treated as functions of $\vec{r}$ and $\vec{v}$. From eq. (5-25a):

$$\left(\frac{\partial\vec{r}\,'_a}{\partial\vec{r}}\right)_{\vec{v}} = I \tag{5-28a}$$

$$\left(\frac{\partial\vec{r}\,'_a}{\partial\vec{v}}\right)_{\vec{r}} = (\Delta t)\, I + O[(\Delta t)^2] \tag{5-28b}$$

where I is the identity matrix. Eq. (5-26b) may be written, then,

$$\vec{\eta}(\vec{r}\,'_e) = \vec{\eta}(\vec{r}) + O[\Delta t + \eta^2] \tag{5-29}$$

Eq. (5-29) is still satisfied for $\vec{r}\,'_e$ given by the second conservative formulation of Sect. 4.B, since

$$\vec{r}\,'_e = \vec{r}\,'_a + O[\Delta t] \tag{5-30}$$

by virtue of eqs. (4-35a) and (4-36). Therefore <u>both</u> conservative formulations preserve the property of unit propagation of error of $\vec{r}$ into $\vec{r}\,'_e$. The asymptotic ($\Delta t \to 0$) stability properties of the conservative $\vec{r}\,'_e$ are thus identical to those of the underlying approximation $\vec{r}\,'_a$.

For the case of $\vec{v}\,'_e$ of Sect. 4.A, the error $\vec{\eta}(\vec{v}\,'_e)$ induced by errors $\vec{\eta}(\vec{r})$ and $\vec{\eta}(\vec{v})$ in $\vec{r}$ and $\vec{v}$ is given by

$$\vec{\eta}(\vec{v}\,'_e) = \left(\frac{\partial\vec{v}\,'_e}{\partial\vec{r}}\right)_{\vec{v}} \cdot \vec{\eta}(\vec{r}) + \left(\frac{\partial\vec{v}\,'_e}{\partial\vec{v}}\right)_{\vec{r}} \cdot \vec{\eta}(\vec{v}) + O[\eta^2] \tag{5-31}$$

Since $\vec{v}'_e$ is given by eq. (4-5b),

$$\left(\frac{\partial \vec{v}'_e}{\partial \vec{r}}\right)_{\vec{v}} = \left(\frac{\partial \vec{v}'_a}{\partial \vec{r}}\right)_{\vec{v}} + \left(\frac{\partial\, \delta\vec{v}_e}{\partial \vec{r}}\right)_{\vec{v}} \tag{5-32a}$$

$$\left(\frac{\partial \vec{v}'_e}{\partial \vec{v}}\right)_{\vec{r}} = \left(\frac{\partial \vec{v}'_a}{\partial \vec{v}}\right)_{\vec{r}} + \left(\frac{\partial\, \delta\vec{v}_e}{\partial \vec{v}}\right)_{\vec{r}} \tag{5-32b}$$

Because eqs. (5-25) are assumed to hold,

$$\left(\frac{\partial \vec{v}'_a}{\partial \vec{r}}\right)_{\vec{v}} = O[\Delta t] \tag{5-33a}$$

$$\left(\frac{\partial \vec{v}'_a}{\partial \vec{v}}\right)_{\vec{r}} = I + O[\Delta t] \tag{5-33b}$$

and eqs. (5-32) become

$$\left(\frac{\partial \vec{v}'_e}{\partial \vec{r}}\right)_{\vec{v}} = \left(\frac{\partial \delta\vec{v}_e}{\partial \vec{r}}\right)_{\vec{v}} + O[\Delta t] \tag{5-34a}$$

$$\left(\frac{\partial \vec{v}'_e}{\partial \vec{v}}\right)_{\vec{r}} = I + \left(\frac{\partial \delta\vec{v}_e}{\partial \vec{v}}\right)_{\vec{r}} + O[\Delta t] \tag{5-34b}$$

Relations (5-34), when substituted into eq. (5-31) give

$$\vec{\eta}(\vec{v}'_e) = \vec{\eta}(\vec{v}) + \left(\frac{\partial \delta\vec{r}_e}{\partial \vec{r}}\right)_{\vec{v}} \cdot \vec{\eta}(\vec{r}) + \left(\frac{\partial \delta\vec{v}_e}{\partial \vec{v}}\right)_{\vec{r}} \cdot \vec{\eta}(\vec{v}) + O[\Delta t + \eta^2] \tag{5-35}$$

Since the objective is to show

$$\vec{\eta}(\vec{v}'_e) = \vec{\eta}(\vec{v}) + O[\Delta t + \eta^2] \tag{5-36}$$

(preserving the asymptotic stability properties of $\vec{v}'_a$), all that remains is to prove

$$\left(\frac{\partial \delta \vec{v}_e}{\partial \vec{r}}\right)_{\vec{v}} = O[\Delta t] \tag{5-37a}$$

$$\left(\frac{\partial \delta \vec{v}_e}{\partial \vec{v}}\right)_{\vec{r}} = O[\Delta t] \tag{5-37b}$$

By virtue of the chain-rule and eqs. (5-28),

$$\left(\frac{\partial}{\partial \vec{r}}\right)_{\vec{v}} = \frac{\partial}{\partial \vec{r}} + \frac{\partial}{\partial \vec{r}'_e} \cdot \frac{\partial \vec{r}'_e}{\partial \vec{r}} + \frac{\partial}{\partial \vec{r}'_a} \cdot \frac{\partial \vec{r}'_a}{\partial \vec{r}} \tag{5-38a}$$

$$= \frac{\partial}{\partial \vec{r}} + \frac{\partial}{\partial \vec{r}'_e} + O[\Delta t] \tag{5-38b}$$

(where multiplications occur after differentiations).
Similarly,

$$\left(\frac{\partial}{\partial \vec{v}}\right)_{\vec{r}} = \frac{\partial}{\partial \vec{v}} + \frac{\partial}{\partial \vec{r}'_e} \cdot \frac{\partial \vec{r}'_e}{\partial \vec{v}} + \frac{\partial}{\partial \vec{v}'_a} \cdot \frac{\partial \vec{v}'_a}{\partial \vec{v}} \tag{5-39a}$$

$$= \frac{\partial}{\partial \vec{v}} + \frac{\partial}{\partial \vec{v}'_a} + O[\Delta t] \tag{5-39b}$$

Carrying out the differentiations of $\vec{\beta}$ via eq. (4-11) gives

$$\frac{\partial \vec{\beta}}{\partial \vec{r}'_e} = -\frac{\partial \vec{\beta}}{\partial \vec{r}} + O[\Delta t] \tag{5-40a}$$

$$= \vec{v}\vec{r} - r\dot{r}\,I + O[\Delta t] \tag{5-40b}$$

where

$$r\dot{r} = \vec{r} \cdot \vec{v} \tag{5-41}$$

and $\vec{a}\vec{b}$ denotes the dyadic matrix with components

$$(\vec{a}\,\vec{b})_{ij} = a_i b_j \tag{5-42}$$

Using eqs. (5-38) for $(\partial\vec{\beta}/\partial\vec{r})_{\vec{v}}$ gives, together with eqs. (5-40),

$$\left(\frac{\partial\vec{\beta}}{\partial\vec{r}}\right)_{\vec{v}} = \frac{\partial\vec{\beta}}{\partial\vec{r}} + \frac{\partial\vec{\beta}}{\partial\vec{r}\,'_e} + O[\Delta t] \tag{5-43a}$$

$$= O[\Delta t] \tag{5-43b}$$

A similar effect occurs for $(\partial\vec{\beta}/\partial\vec{v})_{\vec{r}}$ because of antisymmetric contributions from terms in $\vec{v}$ and $\vec{v}\,'_a$:

$$\frac{\partial\vec{\beta}}{\partial\vec{v}\,'_a} = -\frac{\partial\vec{\beta}}{\partial\vec{v}} + O[\Delta t] \tag{5-44a}$$

$$= \vec{r}\,\vec{r} - r^2 I + O[\Delta t] \tag{5-44b}$$

and so

$$\left(\frac{\partial\vec{\beta}}{\partial\vec{v}}\right)_{\vec{r}} = \frac{\partial\vec{\beta}}{\partial\vec{v}} + \frac{\partial\vec{\beta}}{\partial\vec{r}\,'_a} + O[\Delta t] \tag{5-45a}$$

$$= O[\Delta t] \tag{5-45b}$$

Implicit differentiation of eq. (4-25) gives for the contributions from ϵ

$$\frac{\partial \epsilon}{\partial \vec{r}'_e} = - \frac{\partial \epsilon}{\partial \vec{r}} + O[\Delta t] \tag{5-46a}$$

$$= \frac{r}{\dot{r}} \frac{\vec{F}}{m} + O[\Delta t] \tag{5-46b}$$

and

$$\frac{\partial \epsilon}{\partial \vec{v}'_a} = - \frac{\partial \epsilon}{\partial \vec{v}} + O[\Delta t] \tag{5-47a}$$

$$= - \frac{r}{\dot{r}} \vec{v} + O[\Delta t] \tag{5-47b}$$

Substitution of eqs. (5-46) and (5-47) into eqs. (5-38) and (5-39) for ϵ gives

$$\left(\frac{\partial \epsilon}{\partial \vec{r}}\right)_{\vec{v}} = \frac{\partial \epsilon}{\partial \vec{r}} + \frac{\partial \epsilon}{\partial \vec{r}'_e} + O[\Delta t] \tag{5-48a}$$

$$= O[\Delta t] \tag{5-48b}$$

and

$$\left(\frac{\partial \epsilon}{\partial \vec{v}}\right)_{\vec{r}} = \frac{\partial \epsilon}{\partial \vec{v}} + \frac{\partial \epsilon}{\partial \vec{v}'_a} + O[\Delta t] \tag{5-49a}$$

$$= O[\Delta t] \tag{5-49b}$$

Finally, direct differentiation of eq. (4-14) gives

$$\left(\frac{\partial \delta \vec{v}_e}{\partial \vec{r}}\right)_{\vec{v}} = \frac{1}{(r'_e)^2} \left[\vec{r}'_e \left(\frac{\partial \epsilon}{\partial \vec{r}}\right)_{\vec{v}} + \left(\frac{\partial \vec{\beta}}{\partial \vec{r}}\right)_{\vec{v}}\right] + O[(\Delta t)^n] \tag{5-50a}$$

and

$$\left(\frac{\partial \delta \vec{v}_e}{\partial \vec{v}}\right)_{\vec{r}} = \frac{1}{(r_e')^2}\left[\vec{r}\,'_e\left(\frac{\partial \epsilon}{\partial \vec{v}}\right)_{\vec{r}} + \left(\frac{\partial \vec{\beta}}{\partial \vec{v}}\right)_{\vec{r}}\right] + O[(\Delta t)^{n+1}] \tag{5-50b}$$

Substitution of eqs. (5-43b), (5-45b), (5-48b) and (5-49b) into eqs. (5-50) give (since $n \geq 2$)

$$\left(\frac{\partial \delta \vec{v}_e}{\partial \vec{r}}\right)_{\vec{v}} = O[\Delta t] \tag{5-51a}$$

$$\left(\frac{\partial \delta \vec{v}_e}{\partial \vec{v}}\right)_{\vec{r}} = O[\Delta t] \tag{5-51b}$$

These relations ensure that (via eq. (5-35)), eq. (5-36) holds, and there is only unit reintroduction of the error of $\vec{v}$ into $\vec{v}\,'_e$.

The development of eqs. (5-26) through (5-51) also holds for the second formulation of Sect. 4.B with substitution of $\vec{r}\,'_a$ for $\vec{r}\,'_e$. Thus both conservative formulations preserve the desirable asymptotic ($\Delta t \to 0$) stability properties of the underlying approximations $\vec{r}\,'_a$ and $\vec{v}\,'_a$ of only unit propagation of the errors in $\vec{r}$ and $\vec{v}$. This occurs because of the antisymmetry with respect to interchange of $\vec{r}\,'_e$ and $\vec{v}\,'_e$ with $\vec{r}$ and $\vec{v}$ in the conservative relations.

C. Comparison of the Two Formulations

Two conservative formulations have been proposed, those of Sects. 4.A and 4.B. The two methods differ in the handling of the estimate $\vec{r}\,'_e$ for $\vec{r}\,'$, and have correspondingly different characteristics.

The formulation of Sect. 4.A enjoys the advantage of using a conventional method to find an approximation $\vec{r}\,'_e = \vec{r}\,'_a$ for $\vec{r}\,'$, and is thus "explicit" in $\vec{r}\,'_e$, maintaining the "without first-derivatives present" nature of the equations of motion. As a consequence, eq. (4-25) is a simple, explicit quadratic equation in ϵ, and is easily solved explicitly by, e.g., the quadratic formula.

On the other hand, the version of Sect. 4.B is <u>implicit</u> in $\vec{r}\,'_e$, and <u>both</u> $\vec{r}\,'_e$ and $\vec{v}\,'_e$ must be determined via joint iteration of the <u>implicit</u> quadratic eq. (4-46). This dual iteration destroys the property of "first-derivatives absent", and requires the extra work of redetermining $\phi(r'_e)$ at each iteration. On the other hand, when the same approximation $\vec{r}\,'_a$ is used in Sect. 4.B as in Sect. 4.A, the second formulation results in one higher order approximations for both $\vec{r}\,'_e$ and $\vec{v}\,'_e$.

Thus the first formulation, being "explicit" in $\vec{r}\,'_e$ has the advantages and disadvantages of a "predictor": the method is relatively easy to implement, requires only one potential evaluation $\phi(r'_e)$ per step, and requires no iteration. The second formulation, being "implicit" in $\vec{r}\,'_e$, enjoys the advantages, and suffers from the disadvantages, of a "corrector": an <u>implicit</u> (quadratic) equation must be solved, with the accompanying reevaluation of the potential, but the final results are of higher accuracy.

6. Numerical Examples

In this section, the results of Sect. 4 are compared with conventional procedures, with the underlying approximations $\vec{r}_a$ and $\vec{v}_a$ being given by Adams' methods. In general with predictor-corrector methods special starting and step-size changing procedures are necessary: in the calculations below the Adams-Nordsieck formulation for second-order differential equations was used [8]-[9], with the control algorithms as given by Nordsieck [9]. (Programs for implementing this method for n = 3 to n = 8 are given in the Appendices of [10]).

A comparison between the conventional Adams-Nordsieck methods for n = 3 to n = 8 with the corresponding versions of the conservative methods of Sect. 4 are given in Tables I and II, for the example of Sect. 7.a of [4]. Here $\phi(r)$ is in the Lennard-Jones form of eq. (2-9) with $\epsilon = 1$ and $\sigma = 1$. The initial conditions were

$$\vec{r}(0) = \langle 0, 1, -20 \rangle \tag{6-1a}$$

$$\vec{v}(0) = \langle 0, 0, \sqrt{2} \rangle \tag{6-1b}$$

The mass m was chosen to be unity, giving

$$E = 1.000000 \tag{6-2a}$$

$$\vec{L} = \langle \sqrt{2}, 0, 0 \rangle \tag{6-2b}$$

The Adams' corrector $\vec{r}_c$ was used for the approximation $\vec{r}_a$ of $\vec{r}\,'$ for the method of Sect. 4.A ; the Adams' predictor $\vec{r}_p$ was used for $\vec{r}_a$ in the method of Sect. 4.B. Thus both conservative formulations gave errors of $O[(\Delta t)^{n+1}]$ in $\vec{r}\,'$ and $\vec{v}\,'$. Instead of changing the step-control procedure to take advantage of the extra order of accuracy in $\vec{v}\,'$ in the conservative formulations (which would result in fewer steps), identical step-control procedures were used as in the Adams' methods to control truncation error. (Programs are given in the Appendices of [10]). The superiority of the conservative methods is then reflected in the higher accuracy of the results. Occasionally the conservation of E required a slightly smaller value of Δt in the version of Sect. 4.B: this is reflected in the variations in the total number of steps in Table II.

For comparison, the calculated values of E and L at the last steps ($r > 20$) of the computation are given, along with the maximum observed derivations in these quantities. Finally, the calculated value of the angle of deflection χ, is defined by

$$\chi = \cos^{-1} \frac{\vec{v} \cdot \vec{v}(0)}{\|\vec{v}\| \; \|\vec{v}(0)\|} \tag{6-3}$$

where $\vec{v}(0)$ is given by eq. (6-1a), and for $\vec{v}$ evaluated at the last step. The correct value of χ is 0.996932, and is also listed in Tables I and II. For further details of the problem and the evaluation of χ, see [4].

Table I compares the Adams' methods with the corresponding methods of Sect. 4.A., with iteration of the implicit equation in $\vec{r}\,'$ to a relative convergence in $\|\vec{r}\,'\|$ of 10^{-3}. Table II compares the Adams' and Sec. 4.B methods, with iteration of the implicit equations in $\vec{r}\,'$ and $\vec{v}\,'$ to a relative convergence in $\|\vec{r}\,'\|$ and $\|\vec{v}\,'\|$ of 10^{-10}. Both Tables I and II show the superior accuracy of the conservative formulations in computing χ because of the extra order of accuracy in $\vec{v}\,'$. In addition, of course, the conservative formulations give excellent values for E and $\vec{L}$.

TABLE I.[a] Comparison of Adams' Methods vs. First Conservative Formulation for n = 3 to n = 8 .

Method	Order	No. Steps[b]	Final[c] E	max[d] $\Delta E \times 10^6$	χ[e]	Final[f] ΔI_χ
Adams'	n = 3	1895	1.001251	-3346	.996854	.87(-3)
	4	383	.986616	-293338	.981370	.16(-2)
	5	197	.981625	-29975	.986446	-.21(-2)
	6	190	1.047371	+52422	1.01762	.12(-2)
	7	204	.984948	-18296	.990096	.49(-5)
	8	235	1.152245	+153832	1.05494	-.81(-5)
Sect. 4.A[g]	n = 3	1892	1.000000	0	.996932	.83(-16)
	4	392	1.000000	0	.996932	-.49(-16)
	5	196	1.000000	0	.996932	-.42(-16)
	6	141	1.000000	0	.996931	-.49(-16)
	7	117	1.000000	0	.996934	-.42(-16)
	8	106	1.000000	0	.996934	-.21(-16)
Exact			1.000000	0	0.996932	0

TABLE I - FOOTNOTES

[a]Adams' correctors and implicit equations iterated to a relative convergence of 10^{-3} in $\|\vec{r}'\|$. Numbers in parentheses denote powers of ten.

[b]Total number of steps necessary to reach final state when $r > 20$.

[c]Calculated using $\vec{r}$ and $\vec{v}$ at final step.

[d]Maximum observed deviation between calculated value of E and true value $E = 1.000000$ during any step of the calculation.

[e]Angle of deflection calculated at final step.

[f]Error in X component of the angular momentum $\vec{L}$ at the final step.

[g]First conservative formulation of Section 4.A using the Adams' correctors for $\vec{r}_a{}'$. (Same step size control algorithm as in the Adams' methods.

TABLE II.[a] Comparison of Adams' Methods vs. Second Conservative Formulation for n = 3 to n = 8

Method	Order	No. Steps[b]	Final[c] E	max[d] $\Delta E \times 10^6$	χ[e]	Final[f] ΔL_X
Adams'	n = 3	1893	1.000003	+26	.996957	-.15(-6)
	4	393	1.000016	+19	.996957	-.12(-4)
	5	197	1.000004	+21	.996957	-.55(-5)
	6	140	1.000046	+48	.996993	-.16(-4)
	7	120	1.000038	+42	.997001	-.19(-4)
	8	106	1.000046	+97	.997102	-.73(-4)
Sect. 4.B[g]	n = 3	1892	1.000000	0	.996932	.62(-16)
	4	395	1.000000	0	.996932	.35(-16)
	5	202	1.000000	0	.996932	.14(-16)
	6	151	1.000000	0	.996932	.69(-17)
	7	120	1.000000	0	.996932	.21(-16)
	8	132	1.000000	0	.996932	.14(-16)
Exact			1.000000	0	.996932	0

TABLE II - FOOTNOTES

[a] Adams' correctors and implicit equations iterated to a relative convergence of 10^{-10} in $\|\vec{r}'\|$. Numbers in parentheses denote powers of ten.

[b, c, d, e, f] See Table I.

[g] Second conservative formulation of Sect. 4.B using the Adams' predictors for $\vec{r}_a{}'$. (Same step-size control as in the Adams' methods.)

7. Conclusion

A numerical method for the solution for the motion of a particle subject to a central potential has been given which conserves exactly the additive constants of motion. Two formulations of the method were given, one "explicit", and the other "implicit", in the final position vector $\vec{r}\,'_e$. The method can be applied as a modification to conventional approximation schemes, such as truncated Taylor-series, Runge-Kutta formulae, or predictor-corrector methods. In direct numerical comparison with Adams' methods of order $n = 3$ to $n = 8$, the conservative formulations were found to be superior from the standpoint of exact conservation of energy and angular momentum, smaller truncation error in $\vec{r}\,'$, and an extra order of accuracy in $\vec{v}\,'$.

References

1. L. D. Landau and E. M. Lifschitz, Mechanics, Chap. II, Addison-Wesley, Reading, Mass., 1960.

2. E. J. Saletan and A. H. Cromer, Theoretical Mechanics, Chap. III., Wiley, N.Y., 1971.

3. D. Greenspan:
   (a) "Topics in the Computer Simulation of Discrete Physics", Univ. of Wis., Computer Sciences Dept.,Report WIS-CS-130 (1971);
   (b) "Discrete Newtonian Gravitation and the Three-Body Problem", Univ. of Wis., Computer Sciences Dept., Report WIS-CS-133 (1971), to appear in Foundations of Physics;
   (c) "Discrete Bars, Conductive Heat Transfer, and Elasticity", Univ. of Wis., Computer Sciences Dept., Report WIS-CS-164 (1972), to appear in Computers and Structures;
   (d) "Discrete Newtonian Gravitation and the N-Body Problem", Utilitas Math., 2, 105 (1972);
   (e) "New Forms of Discrete Mechanics", Kybernetes, 1, 87 (1972);
   (f) "An algebraic, energy conserving formulation of classical molecular and Newtonian n-body interaction", Bull. Amer. Math. Soc. 79, 423 (1973);
   (g) Discrete Models, Addison-Wesley, Reading, Mass., 1973.

4. R. A. LaBudde and D. Greenspan:
   (a) "Discrete Mechanics - A General Treatment", Univ. of Wis., Computer Sciences Dept., Report WIS-CS-192, 1973, to appear in J. Computational Physics;

(b) "Discrete Mechanics for Anisotropic Potentials", Univ. of Wis., Computer Sciences Dept., Report WIS-CS-203 (1974).

5. R. A. LaBudde and D. Greenspan, "Energy and Momentum Conserving Methods for the Numerical Integration of Equations of Motion. II. Motion of a System of Particles", Univ. of Wis., Computer Sciences Dept., Tech. Rpt. to appear.

6. For a reduction of the two-body problem to the equivalent central one-body problem, see, e.g., H. Goldstein, Classical Mechanics, pp. 58-59, Addison-Wesley, Reading Mass., 1965.

7. J. O. Hischfelder, C. F. Curtiss, and R. B. Bird, Molecular Theory of Gases and Liquids, Wiley, N.Y., 1965.

8. R. A. LaBudde, "Extension of Nordsieck's Method to the Numerical Integration of Higher-Order Ordinary Differential Equations", Univ. of Wis. Theoretical Chem. Institute Report WIS-TCI-443 (1971).

9. A. Nordsieck, "On Numerical Integration of Ordinary Differential Equations", Math. Comp., <u>16</u>, 22 (1962).

10. R. A. LaBudde and D. Greenspan, "Energy and Momentum Conserving Methods for the Numerical Integration of Equations of Motion. I. Motion of a Single Particle", Univ. of Wis., Computer Sciences Department, Report WIS-CS-208 (1974).

APPENDIX

```
VTEST
ACC 1.10S-02/24/74-21:23:31 (,0)        DVTEST
      C***********************************************************************
      C
      C     PROGRAM FOR CALCULATING TWO-BODY ORBITS OR ANGLES OF SCATTERING
      C     USING ADAMS/CONSERVATIVE ADAMS METHODS OF ADJUSTABLE ORDER.
      C     THE CALCULATIONS ARE CARRIED OUT IN THE XY PLANE IN RELATIVE
      C     COORDINATES (ORIGIN AT ONE PARTICLE).
      C
      C***********************************************************************
            IMPLICIT DOUBLE PRECISION (A-H,O-Z)
            DIMENSION Y(3,9),AUX(3,10),R(3),DR(3)
            DATA AUX(1,1)/0.5D0/,AUX(2,1)/0.5D0/
      C***** NOTE THE IMPACT OF THIS EQUIVALENCE STATEMENT ...
            EQUIVALENCE (Y(1,1),R(1)),(DR(1),Y(1,2)),(T,Y(3,1)),(H,Y(3,2))
           1 ,(HMAX,Y(3,4)),(EPSIT,AUX(3,1)), (YO,Y(2,1)),(B,Y(1,1))
            EXTERNAL FHODM
      50    READ (5,10) MSIGN,METHOD,IORD,IBIT,ISW,EPSIT,B,E,YO,H,HMAX,STEP
      10    FORMAT (5I2,D10.1,6F10.0)
      C***********************************************************************
      C     MSIGN - IF GE 0, INITIAL CONDITIONS ARE OBTAINED FROM B,E, AND YO.
      C         IF MSIGN .LT. 0, I.C. ARE READ IN.
      C     METHOD = 1, ADAMS TYPE. IF = 2, CONSERVATIVE TYPE I. IF = 3, TYPE2
      C     METHOD = 0 IF ADAMS TYPE, = 1 IF ENERGY CONSERVING
      C     IORD - DESIRED ORDER OF METHOD (3 LE IORD LE 8)
      C     IBIT - NO. OF BINARY BITS OF ACCURACY DESIRED IN SOLUTION
      C     ISW - STEP CONTROL SWITCH ... IF -1, NO MOD CONTROL. IF 0 NO MOD
      C         OR TRUNCATION CONTROL.  IF +1 MOD AND TRUNCATION CONTROL.
      C     EPSIT - CONVERGENCE CRITERION FOR CORRECTOR ITERATION
      C     B AND E - INITIAL CONDITIONS OF THE IMPACT PARAMETER B AND ENERGY
      C     YO - INITIAL Y-COORDINATE WHEN MSIGN .GE. 0
      C     H - INITIAL STEPSIZE
      C     HMAX - UPPER BOUND ON STEPSIZE
      C     STEP - INCREMENT IN T AT WHICH POINTS ARE PRINTED
      C***********************************************************************
            IF (MSIGN .GE. 0) GO TO 51
            READ (5,11) (R(I),DR(I),I=1,2)
      11    FORMAT (4F10.4)
            E = 0.5D0*(DR(1)*DR(1) + DR(2)*DR(2))
      51    VREL = DSQRT(2.D0*E)
            NSTEP = 10
            IF (IORD .GE. 5) NSTEP = 5
            WRITE (6,20) METHOD,IORD,B,E,VREL,IBIT,YO,ISW,EPSIT
      20    FORMAT (1H1,5X,'TEST TRAJECTORY USING METHOD',I5,3X,'OF ORDER',I5/
           2 10X,'WITH B =',F10.5,5X,'E =',F10.5/10X,'RELATIVE VELOCITY =',
           3F10.5/10X,'NO. OF BITS =',I5,5X,'YO =',F10.5/10X,'STEP CONTROL SWI
           4TCH =',I5,5X,'ITERATION CRITERION =',D15.6//)
            TIME = 20.D0/DMAX1(1.D0,VREL)
            IF (MSIGN .LT. 0) GO TO 52
            DR(1) = 0.D0
            DR(2) = VREL
      52    Y(3,3) = HMAX/1.04857D06
            ANGO = R(1)*DR(2) - R(2)*DR(1)
            CALL FSTART (METHOD,IORD,1.D-6)
            CALL FADAM(Y)
            WRITE (6,13) ((Y(I,J),J=1,4),I=1,3)
      13    FORMAT (10X,'Y MATRIX ...'//(10X,4D18.8))
```

```
          WRITE (6,14) (AUX(I,1),I=1,3)
    14    FORMAT (//10X,'FIRST COLUMN OF AUX IS',3D18.8//)
          T = 0.D0
          TEST = STEP
          WRITE (6,15) H,HMAX,STEP
    15    FORMAT (10X,'INITIAL STEPSIZE =',D15.6,5X,'MAX. STEPSIZE =',D15.6,
         1 5X,'TIME STEP =',D23.10//)
          CALL FINAL (KOUNT)
          WRITE (6,25)
    25    FORMAT (5X,'STEP',5X,'TIME',5X,'DELTA T',5X,'ENERGY',5X,'RADIUS',
         1 9X,'X',9X,'XDOT',9X,'Y',7X,'YDOT',7X,'VREL'/)
          I = 0
          RR = DSQRT(R(1)*R(1) + R(2)*R(2))
          WRITE (6,30) I,T,H,E,RR,(R(J),DR(J),J=1,2),VREL
    30    FORMAT (4X,I5,F11.6,F11.7,7F11.6)
          CALL TIMSET (0.)
          CALL HODMIN (Y,AUX,IORD,IBIT,2,FHODM,ISW,3)
          DO 100 I = 1,20000
          RR = DSQRT(R(1)*R(1) + R(2)*R(2))
          IF (RR .GT. 10.D0 .AND. T .GE. TIME) GO TO 200
          IF (T .LT. TEST - 1.D-4) GO TO 100
    90    TEST = TEST + STEP
          IF (T .GT. TEST - 1.D-4) GO TO 90
          VREL = DR(1)*DR(1) + DR(2)*DR(2)
          CALL DFPOT(RR,FORCE,POT)
          E = 0.5D0*VREL + POT
          VREL = DSQRT(VREL)
          WRITE (6,30) I,T,H,E,RR,(R(J),DR(J),J=1,2),VREL
   100    CALL HODM (Y,AUX,FHODM,3)
          I = I + 1
   200    VREL = DSQRT(DR(1)*DR(1) + DR(2)*DR(2))
          CALL TIMGET('END')
          CALL FINAL (KOUNT)
          CHI = DACOS(DR(2)/VREL)*DSIGN (1.D0,DR(1))
          RR = DSQRT(R(1)*R(1) + R(2)*R(2))
          CALL DISPOT (RR,POT)
          E = 0.5D0*VREL*VREL + POT
          WRITE (6,30) I,T,H,E,RR
          ANG = R(1)*DR(2) - R(2)*DR(1)
          EANG = ANG - ANG0
          WRITE (6,40) CHI,KOUNT,ANG,EANG
    40    FORMAT (5X,'ANGLE OF DEFLECTION IS ',D15.6,5X,'NO. POT. EVAL.',I8/
         1 5X,'FINAL ANGULAR MOMENTUM =',D18.8,5X,'ERROR =',D18.8)
          GO TO 50
          END

D OF COMPILATION:        NO  DIAGNOSTICS.
```

```
.DFDM
!ACC 1.10S-02/24/74-21:23:34 (,0)          XDFDM
      C
      C***** SUBROUTINE WHICH GIVES FORCES FOR DHODM
      C
            SUBROUTINE FSTART(METHOD,IORD,ETOL)
      C***************************************************************************
      C     IF METHOD = 0, ADAMS-TYPE. IF METHOD = 1, THEN CONSERVATIVE.
      C     IORD - ORDER OF METHOD DESIRED... (3 LE  IORD LE 8)
      C     ETOL - CONVERGENCE CRITERION IN FUNCTIONAL ITERATION.
      C        ENERGY WILL BE CONSERVED TO WITHIN ETOL PER STEP
      C***************************************************************************
            IMPLICIT DOUBLE PRECISION (A-H,O-Z)
            DIMENSION COEF(2,6),AUX(3,1),Q(3,1)
            DATA COEF/0.16666666666666667DO,0.5DO,0.125DO,0.41666666666666667
           1DO,0.10555555555555556DO,0.375DO,0.09375DO,0.34861111111111111DO
           2, 0.0856150793650793651DO,0.32986111111111111DO,
           3 0.0795717592592592593DO,0.315591931216931217DO/
            DATA XMASS/1.DO/
            FACTOR = 2.DO/XMASS
            C = COEF(2,IORD-2)
            GAMMA = COEF(1,IORD-2)/C
            METH = METHOD
            IND = IORD + 1
            EPS = FACTOR*ETOL
            RETURN
      C
      C***** FUNCTION ENTRY FOR DHODM FORCES
      C
            ENTRY FHODM (AUX,Q,ISW,ITSW)
      C**** FHODM FOR MAXIMALLY CONSERVATIVE METHOD
            IF (ISW .NE. 0) GO TO 100
            RP= DSQRT(Q(1,1)*Q(1,1) + Q(2,1)*Q(2,1))
            CALL DFPOT (RP,FN,PHIP)
            FN = -FN/RP
            DO 80 I = 1,2
      80    Q(I,3)  = FN*Q(I,1)
            EPSLN = 0.DO
            RETURN
      100   R = DSQRT(Q(1,1)*Q(1,1) + Q(2,1)*Q(2,1))
            CALL DFPOT (R,FN,PHI)
            FN = -FN/R
            DO 200 I = 1,3
            Q(I,3) = FN*Q(I,1)
      200   Q(I,IND) = Q(I,3) - AUX(I,10)
            GO TO (210,210,250),METH
      210   RETURN
      C
      C     ENTRY FOINT TO CALCULATE VELOCITIES
      C
            ENTRY VCALC (AUX,Q,ITSW)
            DIMENSION ALPHA(2),BETA(2)
            GO TO (470,250,210),METH
      250   ITSW = 0
            HC = C*Q(3,2)
            HGAM = GAMMA*Q(3,2)
            AVP = 0.DO
```

```
      AV = 0.D0
      AR = 0.D0
      ARP = 0.D0
      ASQ = 0.D0
      DO 280 I = 1,2
      IF (METH .EQ. 3) GO TO 270
      BETA(I) = G(I,1)
      ALPHA(I) = Q(I,1)
      GO TO 275
  270 ALPHA(I) = AUX(I,8) - HGAM*AUX(I,9)
      BETA(I) = AUX(I,8)
  275 ASQ = ASQ + ALPHA(I)*ALPHA(I)
      ARP = ARP + ALPHA(I)*BETA(I)
      AVP = AVP + ALPHA(I)*AUX(I,9)
      AV = AV + ALPHA(I)*AUX(I,6)
  280 AR = AR + ALPHA(I)*AUX(I,5)
      TOP = FACTOR*(PHI - PHIF)
      DO 300 I  = 1,2
      BETA(I) = ((AVP*BETA(I) - AV*AUX(I,5))-(ARP*AUX(I,9)-AR*AUX(I,6)))
     1 /ASQ
      WORK = AUX(I,9) + BETA(I)
  300 TOP = TOP + (WORK - AUX(I,6))*(WORK + AUX(I,6))
      KOUNT = 0
  350 OLDEPS = EPSLN
      EPSLN = -(TOP + EPSLN*EPSLN*ASQ)/(2.D0*AVP)
      IF (METH .EQ. 3 .OR. DABS(EPSLN - OLDEPS) .LT. 1.D-10*DABS(EPSLN))
     1 GO TO 400
      KOUNT = KOUNT + 1
      IF (KOUNT .LT. 10) GO TO 350
C     CALCULATE ERROR IN ENERGY AND NEW VELOCITIES
  400 DELE = FACTOR*(PHI - PHIF)
      DO 450 I = 1,2
      WORK = EPSLN *ALPHA(I) + BETA(I)
      Q(I,IND) = WORK/HC
      G(I,2) = WORK + AUX(I,9)
  450 DELE = DELE + (Q(I,2) - AUX(I,6))*(Q(I,2) + AUX(I,6))
      IF (DABS(DELE) .GT. EPS) ITSW = 1
      RETURN
  470 ITSW = 0
      DO 480 I = 1,2
  480 G(I,2) = AUX(I,9) + HC*G(I,IND)
      RETURN
C
C***** FORCE ENTRY FOR CLASSICAL FORCE
C
      ENTRY FADAM (Q)
      R = DSQRT(G(1,1)*G(1,1) + Q(2,1)*Q(2,1))
      CALL DFPOT (R,FN,V)
      FN = - FN/R
      DO 500 I = 1,2
  500 G(I,3) = FN*G(I,1)
      RETURN
      END
```

D OF COMPILATION: NO   DIAGNOSTICS.

```
HODM
ACC 1.10S-03/24/74-21:27:38 (,0)        DHODM
      SUBROUTINE HODMIN (Q,AUX,IORD,IBIT,N,F,ISW,NDIM)
C*********************************************************************
C     VARIABLE ORDER ADAMS METHOD FOR NUMERICAL INTEGRATION
C     SOLVES A SECOND-ORDER SYSTEM OF EQUATIONS D2Q(I)/DX2 =
C       F(X,Q,DQ/DX) , I = 1,2,...,N BY 2ND ORDER NORDSIECK METHODS
C     Q - A MATRIX OF SIZE AT LEAST Q(N+1,IORD+1),WHICH CONTAINS Q(I) IN
C     Q(I,1), DQ(I)/DX IN Q(I,2), D2Q(I)/DX2 IN Q(I,3), ETC.
C     AND X,H,HMIN,HMAX IN Q(N+1,I), I = 1,4. (WHERE H IS STEPSIZE, AND
C      HMIN AND HMAX ARE LOWER AND UPPER BOUNDS ON H )
C     AUX - WORK STORAGE OF SIZE AT LEAST 10*(N+1)
C     AUX(I,1) SHOULD CONTAIN WEIGHTS ASSIGNED TO EQTNS (SUM = 1)
C     AUX (I,J), J = 2,4 WILL CONTAIN A RECORD OF THE INITIAL CONDITIONS
C     AUX(I,J ), J = 5,7 WILL CONTAIN STARTING CONDITIONS FOR CURRENT
C      STEP.  AUX(I,J), J=8,10 WILL CONTAIN PREDICTED FINAL CONDITIONS
C      METHOD ASSUMES THAT AUX(N+1,1) CONTAINS ITERATION CONVERGENCE
C      CRITERION FOR FUNCTIONAL ITERATION
C     IORD - ORDER OF METHOD DESIRED (3.LE. IORD .LE. 8)
C     IBIT - NO. OF BITS OF ACCURACY REQUIRED OF DQ(I)/DX.  IF LT 0,
C       Q MATRIX IS ASSUMED ALREADY PREINITIALIZED
C     N - NUMBER OF EQUATIONS TO BE SOLVED
C     CALL F (AUX,Q,J,ITSW) SHOULD STORE D2Q(I)/DX2 IN Q(I,3) AND ETA*
C      (Q(I,3) - AUX(I,10)) IN Q(I,IORD+1). J IS ITERATION NO. AND ITSW
C     SHOULD BE SET TO 1 IF ITERATION CANNOT END THIS TIME.
C      NOTE THAT ITSW IS A VECTOR, WHERE ITSW(1) IS SET, ITSW(2) IS
C         IORD + 1, AND ITSW(3) IS NO. OF EQUATIONS.
C     ISW = -1,0,+1.  IF 0, NO STEP CONTROL (OTHER THAN STABILITY).
C         IF = -1, NO STEP CONTROL MOD INITIAL STEP
C     NDIM - DIMENSIONED COLUMN LENGTH OF Q AND AUX
C*********************************************************************
      IMPLICIT DOUBLE PRECISION (A-H,O-Z)
      DIMENSION     BETA(6),BINV(7),Q(NDIM,1),AUX(NDIM,1),COEF(6,8),
     1 C(8,6)   ,ERVEC(6),ITVEC(3)
      EQUIVALENCE (ITSW,ITVEC(1)),(NORD3,ITVEC(2)),(M,ITVEC(3))
      DATA BETA/2.D0,4.D0,8.D0,1.6D1,3.2D1,6.4D1/,BINV/0.5D0,0.25D0,
     1 0.125D0,0.0625D0,0.03125D0,0.015625D0,.0078125D0/,ITER/5/
      DATA COEF/0.5D0,0.333333333333333333D0,0.25D0,0.2D0,
     1 0.1666666666666666667D0,0.1428571428571428570D0,7*1.D0,2.D0,3.D0,
     2 4.D0,5.D0,6.D0,0.D0,1.D0,3.D0,6.D0,10.D0,15.D0,2*0.D0,1.D0,4.D0,
     3 10.D0,20.D0,3*0.D0,1.D0,5.D0,15.D0,4*0.D0,1.D0,6.D0,5*0.D0,1.D0/
      DATA ERVEC/0.02D0,0.05D0,0.1D0,0.25D0,0.5D0,1.D0/
      DATA C/0.1666666666666666667D0,0.5D0,1.D0,5*0.D0,
     1 0.125D0,0.4166666666666666667D0,1.D0,0.5D0,4*0.D0,
     2 0.1055555555555555556D0,0.375D0,1.D0,0.75D0,0.1666666666666666667D0
     3,3*0.D0,    0.09375D0,0.3486111111111111111D0,1.D0,0.9166666666666666
     4667D0,0.3333333333333333333D0,0.0416666666666666667D0,2*0.D0,
     5 0.0856150793650793651D0,0.3298611111111111111D0,1.D0,
     6 1.0416666666666666667D0,0.4861111111111111111D0,0.1041666666666666667
     7D0,0.8333333333333333333D-2,0.D0,0.7957175925925925926D-1,
     8 0.3155919312169312170D0,1.D0,1.1416666666666666667D0,0.625D0,
     9 0.1770833333333333333D0,0.025D0,0.1388888888888888889D-2/
      MORD = IORD
      NORD3 = IORD + 1
      NORD = IORD - 2
      NORD2 = 2*NORD
      NORD4 = 4*NORD
```

```
              NORD6 = 6*NORD
              NPLUS1 = N + 1
              HLO = 1.00000000001D0*Q(NPLUS1,3)
              HUP = 1.00000000001D0*Q(NPLUS1,4)
              MODH = ISW
              XO = Q(NPLUS1,1)
              M = N
              DO 100 J = 1,3
              JJ = J + 1
              DO 100 I =1,M
      100     AUX(I,JJ) = Q(I,J)
              EPS = ERVEC(NORD)/(DABS(Q(NPLUS1,2))*2**IABS(IBIT))
              IF (ISW .NE. 0) GO TO 250
              EPS = 1.D+200
              HUP = -1.D0
      250     ISTEP = 0
              IF (IBIT .LT. 0) ISTEP = NORD6
              ITOV =0
              IDELAY = NORD
              IF (IBIT .LT. 0) GO TO 400
              DO 260 J = 4,NORD3
              DO 260 I=1,M
      260     Q(I,J) = 0.D0
              GO TO 400
      C       ENTRY POINT FOR REVERSING STEP-SIZE
              ENTRY HCRRVS (Q,AUX,F,NDIM)
      350     IDELAY = NORD
              ASSIGN 400 TO ISTAT
              GO TO 1000
      C
      C*************************************************************************
      C
      C       NORMAL ENTRY POINT FOR INTEGRATING ONE STEP
              ENTRY HCDM (Q,AUX,F,NDIM)
      C
      C*************************************************************************
              IF (IDELAY .LE. 0) GO TO 370
              IDELAY = IDELAY - 1
              GO TO 400
      C       CHECK IF STEP-SIZE CAN BE DOUBLED
      370     IF (ITOV .EQ. 0) GO TO 400
              HW = BETA(1)*Q(NPLUS1,2)
              T = Q(NPLUS1,1) - XO
              I = T/HW
              IF (MODH*DABS(T-I*HW) .GT. HLO .OR. DABS(HW) .GT. HUP) GO TO 400
              ASSIGN  400 TO ISTAT
              GO TO 1200
      400     CALL F(AUX,Q,0,ITSW)
      405     H = Q(NPLUS1,2)
              TC = 0.D0
              DO 450 I = 1,M
      C       STORE STARTING CONDITIONS OF STEP IN AUX(I,J), J = 5,6,7
              DO 420 J = 1,3
              AUX(I,J+4) = Q(I,J)
      C       STORE PREDICTORS IN AUX(I,J ), J = 8,9,10
              W = 0.D0
              DO 410 K = 1,NORD
```

```
410   W = W + COEF(K,J)*Q(I,K+2)
420   AUX(I,J+7) = W
      AUX(I,8) = Q(I,1) + H*(Q(I,2) + H*AUX(I,8) )
      AUX(I,9) = Q(I,2) + H*AUX(I,9)
      Q(I,1) = AUX(I,8)
      Q(I,2) = AUX(I,9)
450   TO = TO + AUX(I,1)*DABS(Q(I,1))
      TO = TO*AUX(NPLUS1,1)
      Q(NPLUS1,1) = Q(NPLUS1,1) + H
      H2 = C(1,NORD)*H*H
      HW = C(2,NORD)*H
C     SOLVE CORRECTORS BY FUNCTIONAL ITERATION
      DO 550 J = 1,ITER
      T = 0.D0
      CALL F (AUX,Q,J,ITSW)
      DO 500 I = 1,M
      W = Q(I,1)
      Q(I,1) = AUX(I,8) + H2*Q(I,NORD3)
500   T = T + AUX(I,1)*DABS(Q(I,1) - W)
      IF (ITSW .EQ. 0 .AND. T .LT. TO) GO TO 560
      IF (ITSW .GT. 1) GO TO 555
550   CONTINUE
C     ITERATION DID NOT CONVERGE
555   ITA = 0
      GO TO 590
560   ITA =1
      CALL F (AUX,Q,1,ITSW)
      CALL VCALC (AUX,Q,ITSW)
      IF (ITSW .NE. 0) GO TO 555
C**** CALCULATE ESTIMATE OF TRUNCATION ERROR
      DODA = 0.D0
      DO 580 I = 1,M
580   DODA = DODA + AUX(I,1)*DABS(Q(I,NORD3))
C     TEST TRUNCATION ERROR
      IF (DODA.GT.EPS) GO TO 590
      ITB =1
      ITOV =1
      IF (DODA - BINV(NORD+1)*EPS) 600,600,595
590   ITB =0
595   ITOV =0
600   IF (ISTEP .GE. NORD6) GO TO 605
      IF (ISTEP .GT. 0) GO TO 650
605   IF (ITA .EQ. 0) GO TO 620
610   IF (ISTEP.EQ.0.OR.ITB.NE.0) GO TO 650
C     HALVE STEP-SIZE
620   IF (DABS(Q(NPLUS1,2)) .LT. HLO) GO TO 1500
      Q(NPLUS1,1) = Q(NPLUS1,1) - Q(NPLUS1,2)
      Q(NPLUS1,2) = BINV(1)*Q(NPLUS1,2)
      EPS = BETA(1)*EPS
      DO 630 I=1,M
      DO 625 J = 1,3
625   Q(I,J) = AUX(I,J+4)
      DO 630 J = 4,NORD3
630   Q(I,J) = BINV(J-3)*Q(I,J)
      IDELAY = NORD
      GO TO 405
C     UPDATE INTERPOLATED DERIVATIVES
```

```
      650  ISTEP = ISTEP +1
           IF (NORD .EQ. 1) GO TO 705
           DO 700 I =1,M
           DODA = Q(I,3) - AUX(I,10)
           DO 700 J = 4,MORD
           W = DODA*C(J,NORD)
           DO 660 K = J,MORD
      660  W = W + COEF(K-2,J)*Q(I,K)
      700  Q(I,J) = W
      705  IF (ISTEP .GT. NORD6) RETURN
      C    TRANSFER TO SPECIAL ROUTINES FOR STARTING PROCEDURE
           IF (ISTEP .NE. (ISTEP/NORD)*NORD) GO TO 400
           IF (ISTEP .EQ. NORD2) GO TO 710
           IF (ISTEP .EQ. NORD4) GO TO 760
           IF (ISTEP .EQ. NORD6) GO TO 770
           ASSIGN 400 TO ISTAT
           GO TO 1000
      C    RESET TO I.C. AND REVERSE STEP DIRECTION
      710  DO 750 J = 1,3
           JJ = J + 1
           DO 750 I = 1,M
      750  Q(I,J) = AUX(I,JJ)
           Q(NPLUS1,1) = X0
           GO TO 350
      C    HALVE STEP-SIZE AND REVERSE DIRECTION
      760  Q(NPLUS1,2) = BINV(1)*Q(NPLUS1,2)
           EPS = BETA(1)*EPS
           DO 765 J = 4,NORD3
           DO 765 I=1,M
      765  Q(I,J) = BINV(J-3)*Q(I,J)
           IF (ITB.NE.0) GO TO 710
      C    IF TRUNCATION TEST FAILED FOR THIS STEP (ITB = 0), THEN RESTART
      C       WITH CURRENT SMALLER STEPSIZE AS ESTIMATE
           ASSIGN 250 TO ISTAT
           GO TO 1000
      C    DOUBLE STEP-SIZE
      770  ASSIGN 710 TO ISTAT
           GO TO 1200
      C    PROCEDURE FOR REVERSING STEP DIRECTION
      1000 Q(NPLUS1,2) = -Q(NPLUS1,2)
           DO 1100 J = 4,NORD3,2
           DO 1100 I=1,M
      1100 Q(I,J) = -Q(I,J)
           GO TO ISTAT,(250,400)
      C    PROCEDURE FOR DOUBLING STEP-SIZE
      1200 Q(NPLUS1,2) = BETA(1)*Q(NPLUS1,2)
           EPS = BINV(1)*EPS
           DO 1300 J = 4,NORD3
           DO 1300 I=1,M
      1300 Q(I,J) = BETA(J-3)*Q(I,J)
           IDELAY = NORD
           GO TO ISTAT,(400,710)
      C    ERROR TERMINATION IF STEP-SIZE IS TOO SMALL
      1500 PRINT 10,ISTEP,(Q(NPLUS1,I),I=1,4),DODA,EPS,T,TO
      10   FORMAT ('0****HODM FAILS AT STEP',I8,2X,'X =',D19.9/5X,'STEPSIZE =
          1',D19.9,5X,'BOUNDS =',2D19.9//5X,'TRUNCATION TEST',D19.9,5X,'CRITE
          2RION',D19.9/5X,'ITERATION TEST',D19.9,5X,'CRITERION',D19.9//5X,
```

```
          3 'SOLUTION MATRIX ...'/)
            DO 1600 I = 1,M
      1600  PRINT 20,I,(Q(I,J),J=1,NORD3)
      20    FORMAT (5X,I5,8D15.7)
            DO 1700 I = 1,M
      1700  PRINT 30,I,(ALX(I,J),J=1,10)
      30    FORMAT (/5X,I5,6D17.9/10X,4D17.9)
            STOP
            END
```

ND OF COMPILATION: NO DIAGNOSTICS.

```
DPOT
MACC 1.10S-02/24/74-21:23:47 (,0)          DPOT
                SUBROUTINE DISPOT (R,POT)
                IMPLICIT DOUBLE PRECISION (A-H,O-Z)
                DATA KOUNT /0/
                ISW = 1
        100     R6 = 1.D0/R**6
                POT = 4.D0*R6*(R6 - 1.D0)
                KOUNT = KOUNT + 1
                GO TO (200,300),ISW
        200     RETURN
                ENTRY DFPOT (R,FN,POT)
                ISW = 2
                GO TO 100
        300     FN = 48.D0*R6*(0.5D0 - R6)/R
                RETURN
                ENTRY FINAL (I)
                I = KOUNT
                KOUNT = 0
                RETURN
                END

ND OF COMPILATION:          NO   DIAGNOSTICS.
```