Internat. J. Math. & Math. Sci.
Vol. 10 No. 1 (1987) 173-179

AN ENERGY CONSERVING MODIFICATION OF NUMERICAL METHODS FOR THE INTEGRATION OF EQUATIONS OF MOTION

ROBERT A. LABUDDE

821 Hialeah Dr.
Virginia Beach, VA 23462

and

DONALD GREENSPAN

Department of Mathematics
University of Texas
Arlington, Texas 76019

(Received July 2, 1982)

ABSTRACT. In the integration of the equations of motion of a system of particles, conventional numerical methods generate an error in the total energy of the same order as the truncation error. A simple modification of these methods is described, which results in exact conservation of the energy.

KEY WORDS AND PHRASES. Conventional numerical methods, ordinary differential equations, multiplicative parameters, truncation error.
1980 AMS SUBJECT CLASSIFICATION CODE. 70F34.

1. INTRODUCTION.

When applied to the motion of a system of particles, conventional numerical methods for the integration of ordinary differential equations only approximately conserve the total energy of the system. The error in the calculated value of the energy is of the same order as the truncation error in the velocities. In previous work [1]-[5], a new class of methods was described, which maximally conserve the constants of motion. These methods exactly conserve the total energy and linear momentum, and conserve the total angular momentum to at least one higher order than the corresponding conventional methods.

In what follows, our purpose is to show how conventional numerical methods--exemplified by the third-order Taylor series and Adams' formulae--can be modified so that exact conservation of energy occurs. This modification simply involves the introduction of adjustable, multiplicative parameters, whose values are unity for the conventional case.

2. EQUATIONS OF MOTION.

The following is a brief description of the equations of motion of a system of n particles, interacting according to a pairwise-additive potential. For more details, see [1] or [5].

Suppose particle i has mass m_i, position vector

$$\vec{r}_i = (x_i, y_i, z_i), \tag{2.1}$$

velocity vector

$$\vec{v}_i = \left(\frac{dx_i}{dt}, \frac{dy_i}{dt}, \frac{dz_i}{dt}\right) , \tag{2.2}$$

and acceleration

$$\vec{a}_i = \left(\frac{d^2x_i}{dt^2}, \frac{d^2y_i}{dt^2}, \frac{d^2z_i}{dt^2}\right) . \tag{2.3}$$

Newton's law of motion

$$m_i\vec{a}_i = \vec{F}_i \tag{2.4}$$

relates the acceleration $\vec{a}_i$ to the force $\vec{F}_i$, given by

$$\vec{F}_i = -\frac{\partial\phi}{\partial\vec{r}_i} \tag{2.5}$$

where ϕ is the potential of interaction. It will be assumed that ϕ has the pairwise-additive form

$$\phi(\vec{r}_1,\vec{r}_2,\ldots,\vec{r}_n) = \sum_{i<j} \phi_{ij}(r_{ij}) \tag{2.6}$$

where r_{ij} is the magnitude of the vector distance $\vec{r}_{ij}$ between particles i and j :

$$\vec{r}_{ij} = \vec{r}_j - \vec{r}_i \tag{2.7}$$

As a consequence of equation (2.6),

$$\vec{F}_i = \sum_{j=1}^{n} \vec{F}_{ji} \tag{2.8}$$

where

$$\vec{F}_{ji} = -\vec{F}_{ij} = -\frac{d\phi_{ji}}{dr_{ji}}\frac{\vec{r}_{ji}}{r_{ji}} = \frac{d\phi_{ij}}{dr_{ij}}\frac{\vec{r}_{ij}}{r_{ij}} \tag{2.9}$$

and $\vec{F}_{ji} = \vec{0}$ if $j = i$. The introduction of equation (2.8) into equation (2.4) gives the equation of motion

$$m_i\vec{a}_i = \sum_{j=1} \vec{F}_{ji} = \sum_{j=1}^{n} \frac{d\phi_{ij}}{dr_{ij}}\frac{\vec{r}_{ij}}{r_{ij}} \tag{2.10}$$

For n particles, equation (2.10) yields a system of second-order ordinary differential equations for the $\vec{r}_i$. This system may be used to solve for the $\vec{r}_i'$ and $\vec{v}_i'$ at any later time $t' = t + \Delta t$, given the $\vec{r}_i$ and $\vec{v}_i$ at time t.

Conservation of the total energy E occurs because of the existence of the potential ϕ. Here,

$$E = \sum_{i=1}^{n} \frac{1}{2} m_i(\vec{v}_i\cdot\vec{v}_i) + \phi = \sum_{i=1}^{n} \frac{1}{2} m_i(\vec{v}_i\cdot\vec{v}_i) + \sum_{i<j} \phi_{ij}$$

where $\vec{a}\cdot\vec{b}$ denotes the scalar product of two vectors $\vec{a}$ and $\vec{b}$. Conservation of energy is expressed by the equation

$$E(t') = E(t)$$

for any two times t and t', with E evaluated along the trajectory.

3. CONVENTIONAL NUMERICAL METHODS

A simple example of a conventional approximation method for the numerical solution of equations (2.10) is provided by the truncated Taylor-series formulae

$$\vec{r}'_{c,i} = \vec{r}_i + \vec{v}_i \Delta t + \frac{1}{m_i} \sum_{j=1}^{n} \left[\vec{F}_{ji} \frac{(\Delta t)^2}{2} + \vec{G}_{ji} \frac{(\Delta t)^3}{6} \right] \tag{3.1}$$

and

$$\vec{v}'_{c,i} = \vec{v}_i + \frac{1}{m_i} \sum_{j=1}^{n} \left[\vec{F}_{ji} \Delta t + \vec{G}_{ji} \frac{(\Delta t)^2}{2} \right] \tag{3.2}$$

where the $\vec{r}'_{c,i}$ and $\vec{v}'_{c,i}$ are the calculated values for the $\vec{r}'_i$ and $\vec{v}'_i$ at time $t' = t + \Delta t$, and

$$\vec{G}_{ji} = \frac{d\vec{F}_{ji}}{dt} = - \frac{d\phi_{ji}}{dr_{ji}} \frac{\vec{v}_{ji}}{r_{ji}} - \left[\frac{d^2\phi_{ji}}{dr_{ji}^2} - \frac{1}{r_{ji}} \frac{d\phi_{ji}}{dr_{ji}} \right] \frac{\dot{r}_{ji} \vec{r}_{ji}}{r_{ji}} \tag{3.3}$$

where

$$\dot{r}_{ji} = \frac{dr_{ji}}{dt} = \frac{\vec{r}_{ji} \cdot \vec{v}_{ji}}{r_{ji}}$$

and

$$\vec{v}_{ji} = \vec{v}_i - \vec{v}_j$$

The method of equations (3.1) and (3.2) is of third-order, since

$$\vec{r}_i' = \vec{r}'_{c,i} + 0[(\Delta t)^4] \tag{3.4}$$

and

$$\vec{v}_i' = \vec{v}'_{c,i} + 0[(\Delta t)^3] \tag{3.5}$$

due to the neglect of the succeeding Taylor-series terms. These errors generate an error of $0[(\Delta t)^3]$ in the value of the energy E'_c calculated using the $\vec{r}'_{c,i}$ and $\vec{v}'_{c,i}$:

$$\Delta E_c = E_c' - E = 0[(\Delta t)^3] \tag{3.6}$$

The third-order Adams' method arises via equations (3.1) and (3.2) and the approximation

$$\vec{G}_{ij} = \vec{G}^a_{ij} + 0[\Delta t] \tag{3.7}$$

where

$$\vec{G}^a_{ij} = \frac{\vec{F}'_{c,ij} - \vec{F}_{ij}}{\Delta t} \tag{3.8}$$

In equation (3.8), $\vec{F}'_{c,ij}$ denotes the value of $\vec{F}_{ij}$ obtained from equation (2.9) using the $r'_{c,ij}$. Equations (3.4), (3.5), and (3.6) also hold when the $\vec{G}^a_{ij}$ are used for the $\vec{G}_{ij}$.

4. ENERGY CONSERVING MODIFICATION OF CONVENTIONAL METHODS.

Consider the third-order methods of Section 3, with $\vec{G}^{*}_{ij}$ replacing either the $\vec{G}_{ij}$ or $\vec{G}^{a}_{ij}$:

$$\vec{r}'_{c,i} = \vec{r}_i + \vec{v}_i \Delta t + \frac{1}{m_i} \sum_{j=1}^{n} \left[\vec{F}_{ji} \frac{(\Delta t)^2}{2} + \vec{G}^{*}_{ji} \frac{(\Delta t)^3}{6} \right] \tag{4.1}$$

and

$$\vec{v}'_{c,i} = \vec{v}_i + \frac{1}{m_i} \sum_{j=1}^{n} \left[\vec{F}_{ji} \Delta t + \vec{G}^{*}_{ji} \frac{(\Delta t)^2}{2} \right] \tag{4.2}$$

When equations (4.1) and (4.2) are used to obtain estimates for $\vec{r}'_i$ and $\vec{v}'_i$, an error ΔE_c is made in the total energy, which is given by

$$\begin{aligned}
\Delta E_c = E'_c - E &= \sum_{i=1}^{n} \frac{1}{2} m_i (\vec{v}'_{c,i} \cdot \vec{v}'_{c,i} - \vec{v}_i \cdot \vec{v}_i) + \phi'_c - \phi \\
&= \Delta t \sum_{i=1}^{n} \sum_{j=1}^{n} \left[(\vec{v}_i + \frac{\Delta t}{2m_i} \sum_{k=1}^{n} \vec{F}_{ki}) \cdot \vec{F}_{ji} \right. \\
&\quad \left. + \frac{\Delta t}{2} (\vec{v}_i + \frac{\Delta t}{m_i} \sum_{k=1} \left\{ \vec{F}_{ki} + \vec{G}^{*}_{ki} \frac{\Delta t}{4} \right\}) \cdot G^{*}_{ji} \right] + \Delta\phi \\
&= \Delta t \sum_{i<j} \left[\left\{ \vec{v}_{ij} + \vec{a}_{ij} \Delta t + \vec{b}_{ij} \frac{(\Delta t)^2}{4} \right\} \cdot \vec{G}^{*}_{ij} \frac{\Delta t}{2} \right. \\
&\quad \left. + (\vec{v}_{ij} + \vec{a}_{ij} \frac{\Delta t}{2}) \cdot \vec{F}_{ij} + \frac{\Delta\phi_{ij}}{\Delta t} \right]
\end{aligned} \tag{4.3}$$

where

$$\vec{a}_{ij} = \sum_{k=1}^{n} \left[\frac{\vec{F}_{kj}}{m_j} - \frac{\vec{F}_{ki}}{m_i} \right],$$

$$\vec{b}_{ij} = \sum_{k=1}^{n} \left[\frac{\vec{G}^{*}_{kj}}{m_j} - \frac{\vec{G}^{*}_{ki}}{m_i} \right],$$

$$\Delta\phi_{ij} = \phi'_{c,ij} - \phi_{ij} = \phi_{ij}(r'_{c,ij}) - \phi_{ij}(r_{ij}),$$

and

$$\vec{r}'_{c,ij} = \vec{r}'_{c,j} - \vec{r}'_{c,i} \tag{4.4}$$

Suppose now, instead of using $\vec{G}^{*}_{ij} = \vec{G}_{ij}$ or $\vec{G}^{a}_{ij}$ in equation (4.3)--which leads to an error ΔE_c of $0[(\Delta t)^3]$--that adjustable $\vec{G}^{*}_{ij}$ given by

$$\vec{G}^{*}_{ij} = \varepsilon_{ij} \vec{G}_{ij} \tag{4.5}$$

or

$$\vec{G}^{*}_{ij} = \varepsilon_{ij} \vec{G}^{a}_{ij}$$

is used. The ε_{ij} are to be chosen so that

$$\varepsilon_{ij} = 1 + O[\Delta t] \tag{4.6}$$

(preserving the order of the method) and so that exact conservation of energy occurs. Solving

$$\Delta E_c = 0$$

for the ε_{ij} gives, for example, for (4.5), the equation (cf. [1] and [5])

$$\left\{\vec{v}_{ij} + \vec{a}_{ij}\Delta t + \vec{b}_{ij}\frac{(\Delta t)^2}{4}\right\} \cdot \vec{G}_{ij}\frac{\Delta t}{2}\,\varepsilon_{ij} + \left(\vec{v}_{ij} + \vec{a}_{ij}\frac{\Delta t}{2}\right) \cdot \vec{F}_{ij} + \frac{\Delta\phi_{ij}}{\Delta t} = 0 \tag{4.7}$$

For n particles, equation (4.7) yields a set of implicit, coupled equations in the ε_{ij}, since the b_{ij} and $\phi'_{c,ij}$ depend upon the values of the ε_{ij}.

For small Δt, the equations of (4.7) are strongly linear in the ε_{ij}. The only nonlinear dependences on the ε_{ij} occur through the $\vec{b}_{ij}$ and $\phi'_{c,ij}$ (through the $r'_{c,ij}$). In both these cases, the terms involving the ε_{ij} occur with coefficients proportional to $(\Delta t)^3$. (Compare equations (4.1), (4.2), (4.4), and (4.7).) In contrast, the coefficients of the linear terms in ε_{ij}, namely

$$\left(\vec{v}_{ij} + \vec{a}_{ij}\Delta t\right) \cdot \vec{G}_{ij}\frac{\Delta t}{2}$$

are of $O[\Delta t]$.

Because the equations of (4.7) are linear except for terms of $O[(\Delta t)^3]$, they may be easily solved via the iteration formula

$$\varepsilon_{ij} = -\frac{2}{\Delta t}\,\frac{\dfrac{\Delta\phi_{ij}}{\Delta t} + \left(\vec{v}_{ij}+\vec{a}_{ij}\dfrac{\Delta t}{2}\right)\cdot\vec{F}_{ij}}{\vec{G}_{ij}\cdot\left(\vec{v}_{ij}+\vec{a}_{ij}\Delta t+\vec{b}_{ij}\dfrac{(\Delta t)^2}{4}\right)} \tag{4.8}$$

For small Δt, the equations of (4.8) are solved via successive substitutions, starting with

$$\varepsilon_{ij} = 1 \tag{4.9}$$

Iteration to convergence of the ε_{ij} guarantees <u>exact</u> conservation of energy in the method.

Higher-order formulae may be obtained directly in the same way as equation (4.7). If the highest-order terms involve

$$\vec{F}^{(m)}_{ij} = \frac{d^m\vec{F}_{ij}}{dt^m},$$

then these are replaced by

$$\vec{F}^{(m)*}_{ij} = \varepsilon_{ij}\,\vec{F}^{(m)}_{ij}$$

where the ε_{ij} satisfy equation (4.6). The formulae for the $\vec{v}'_{c,i}$ are substituted in equation (4.3), the sum transformed to $i < j$, and the ij terms set individually to zero. These resulting implicit equations in the ε_{ij} are then solved by standard methods, with the first approximations given by equations (4.9).

For very high order methods, the extra algebra needed to obtain the ε_{ij} is considerable, and substantially reduces the relative efficiency of the method. However, it should be noted that conservation of energy guarantees stability in the usual sense (bounded motion), which is always a desirable computational property.

5. NUMERICAL EXAMPLE.

As an illustration of the affect of the modification described in Section 4, the modified and unmodified forms of the third-order Adams' method are compared numerically on a sample two-dimensional problem involving two particles.

Here $n = 2$,

$$m_1 = m_2 = 2 \tag{5.1}$$

and the gravitational interaction

$$\phi_{12}(r_{12}) = - \frac{1}{r_{12}} \tag{5.2}$$

is used. The initial conditions are chosen so that the center-of-mass of the system is at rest with

$$\vec{r}_{12}(0) = (\tfrac{1}{2},0) \tag{5.3}$$

$$\vec{v}_{12}(0) = (0,1.63). \tag{5.4}$$

The value of the energy is then

$$E = - 0.6715500000... \tag{5.5}$$

Because of the form of ϕ_{12} in (5.2), the exact motion that occurs traces out a closed ellipse with major-axis

$$2a = 1.48909\ 23855 \tag{5.6}$$

corresponding to upper and lower bounds on r_{12} of

$$r_> = 0.98909\ 23855$$

and $$r_< = 0.50000\ 00000.$$

The motion repeats itself with period τ equal to

$$\tau = 4.0366\ 15087,$$

The implicit equations of the third-order methods were iterated to a relative convergence of 10^{-8}. A constant step-size of

$$\Delta t = \tau/80$$

was used. In order to focus attention on the errors made in the methods, results were obtained at times t which were multiples of the period τ, where the exact solu-

tion returns to the initial conditions. Measures of the errors at these points are the error in the calculated value of E, the deviations from zero of dX/dt and Y, and the deviation of r_{12} from 1/2.

Table I gives these quatities for several times $t = m\tau$. It can be seen that the unmodified Adams' method makes an error in E as well as larger errors in dX/dt and Y, and compares unfavorably with the modified method. Another simple measure of the error for this problem is the number of steps over which a phase error of 180^o is made: i.e., the time at which $r_{12} = 0.985$ instead of 0.5. For the unmodified methods, this was about 2800 steps (35τ). For the modified methods, at 20000 steps (250τ), a phase error of less than 180^o had been made.

Programs for the methods are given in the Appendix of [6].

m	Method	E	r	$\frac{dX}{dt}$	Y
0	Exact[a]	- 0.67155	0.50000	0.00000	0.00000
1	U[b]	- 0.67140	0.50221	0.20630	-0.08704
	M[c]	- 0.67155	0.49997	0.02164	-0.00462
2	U	- 0.67099	0.50873	0.40254	-0.17213
	M	- 0.67155	0.49997	0.04328	-0.00923
3	U	- 0.67040	0.51924	0.58036	-0.25351
	M	- 0.67155	0.50001	0.06492	-0.01385
5	U	- 0.66905	0.55019	0.86162	-0.39996
	M	- 0.67155	0.50017	0.10818	-0.02311
10	U	- 0.66679	0.65934	1.15127	-0.64976
	M	- 0.67155	0.50116	0.21592	-0.04639
100	U	- 0.66561	0.97998	0.82003	-0.97598
	M	- 0.67155	0.62554	1.35684	-0.57888

At times t = m

[a]Initial conditions

[b]Unmodified third-order Adams' method

[c]Third-order Adams' method modified to give exact energy conservation.

TABLE I.
Comparison of Modified and Unmodified Methods
on a Simple Gravitation Problem

REFERENCES

1. LaBUDDE, R. A. and GREENSPAN, D., "Discrete Mechanics--A General Treatment", to be published in J. Computational Phys.

2. LaBUDDE, R. A. and GREENSPAN, D., "Discrete Mechanics for Anisotropic Potentials", Univ. of Wis. Computer Sciences Dept. Report WIS-CS-203 (1974).

3. LaBUDDE, R. A. and GREENSPAN, D., "Energy and Momentum Conserving Methods of Arbitrary Order for the Numerical Integration of Equations of Motion. I. Motion of a Single Particle", Univ. of Wis. Computer Sciences Dept. Report WIS-CS-208 (1974).

4. LaBUDDE, R. A. and GREENSPAN, D., "Discrete Mechanics for Nonseparable Potentials with Application to the LEPS Form", Univ. of Wis. Computer Sciences Dept. Report WIS-CS-210 (1974).

5. LaBUDDE, R. A. and GREENSPAN, D., "Energy and Momentum Conserving Methods of Arbitrary Order for the Numerical Integration of Equations of Motion. II. Motion of a System of Particles," Univ. of Wis. Computer Sciences Dept. Report WIS-CS-215 (1974).

6. LaBUDDE, R. A., and GREENSPAN, D., "An Energy Conserving Modification of Numerical Methods for the Integration of Equations of Motion," Univ. of Wis. Computer Sciences Dept. Report WIS-CS-217 (1974).