

A Numerical Integration Scheme for the N-Body Gravitational Problem

A. AHMAD* AND L. COHEN†

Hunter College of the City University of New York, New York, New York 10021

Received March 7, 1972

A method for numerically integrating the N -body gravitational problem is described. We take advantage of the fact that the force on a star can be divided into two parts which operate on different time scales. One part is due to the stars in the immediate vicinity of the star in question and another part due to the distant stars. The part of the force due to the far away stars changes much more slowly than the component due to the nearby stars. Hence that part of the force does not have to be recalculated as frequently as that due to the nearby stars. For systems with large N , most of the stars constitute the "distant" stars and the considerable saving of computing time allows us to integrate systems with up to 1000 particles.

1. INTRODUCTION

Computer simulation of the N -body self gravitating problem has become a powerful tool for the investigation of stellar systems. The basic approach is that starting with the masses, positions, and velocities of N stars at a given time, one can calculate, by numerically solving the equation of motion, the state of the system at any later time. From a knowledge of the positions and velocities as a function of time one can study particular physical phenomena. Solving the equations of motion directly, as contrasted to using Monte Carlo schemes, has the advantage that it is assumption-free and includes all possible interactions. Unfortunately, at present the number of particles that can be evolved by directly integrating the equations of motion is small, although the improvement of both technique and computer speed has increased significantly the number of particles manageable.

* Presented as part of a thesis to the City University of New York in partial fulfillment of the requirements for the Ph.D. degree.

† Work supported by a grant from the National Science Foundation.

Mathematically, the N -body gravitational problem involves the solution of N second order differential equations,

$$\frac{d^2 \mathbf{r}_i}{dt^2} = - \sum_{\substack{j=1 \\ i \neq j}}^N G \frac{m_j \mathbf{r}_{ij}}{|\mathbf{r}_{ij}|^3} \quad i = 1, 2, \dots, N, \quad (1)$$

$$\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j,$$

where \mathbf{r}_i and m_i are the position and mass of the i th star, respectively, and G is the gravitational constant. This system is nonlinear and strongly coupled. There are two basic factors which lead to difficulty in integrating the system given by (1). First, close encounters lead to instabilities. Secondly, since the force on each star depends on the position of all other stars the time needed in calculating the force increases as the square of the number of particles being integrated. Computing time, therefore, increases at least as N^2 and this has been the basic limitation of studying systems with large N . These difficulties limit the usefulness of conventional integration schemes, such as the Runge-Kutta method, for the numerical integration of the gravitational problem and have forced investigators to look for new methods.

A major advance was made by Aarseth [1] who introduced variable and individual time steps for each star. This means that all the particles are treated separately in the integration. The only coupling occurs at the time when the force on a particle is calculated. Also, the assignment of a time step to each star avoids the calculation of the force on every particle when only the recalculating of the force on a particular particle is required. This is particularly important during a close encounter when the force on a particle is changing very quickly and, hence, necessitates recalculation at much higher frequency than for the rest of the system. We shall describe below a further improvement on this idea by incorporating what Wielen [2] has called "double individual" time steps. It has been realized by a number of authors [2, 3, 4] that not all particles in the system are indeed responsible for the necessity of recalculating the force on the particular particle being considered. Our scheme takes advantage of this fact by dividing the force on a particle into two parts: a slowly varying part which is due to the "distant" stars, the regular force, and another component, the irregular force, due to the stars in the immediate neighborhood of the star in question. From the theoretical study of the gravitational force it is known that the highly fluctuating part of the force comes from few close neighbors [5]. As a practical working definition of the irregular force we have chosen the following method. Each particle is surrounded by a sphere of a given radius and the force due to the particles within that sphere on the particle in question is the irregular force. The size of the sphere and, hence, the number of neighbors is made variable depending on the local density.

A difficulty which arises is in the matching of the irregular and regular forces when the neighbors of a particle change. We have circumvented this difficulty by analytically calculating the contributions of the changing neighbors. The efficiency of the scheme arises from the fact that the regular force changes much more slowly than the irregular force and, therefore, the regular force does not have to be calculated as often as that due to the nearby stars. For systems with a large number of particles most of the stars constitute the distant stars and there is, hence, a considerable saving of time in the force calculation. Our basic integration format follows that of Aarseth.

An integration method which uses double individual time steps and has certain similarities to ours is the category scheme of Hayli [4] and Henon. Stars are grouped into categories—the stars in the same category having near equal time steps. The aim of both schemes is to isolate those field particles which are responsible for the highly fluctuating part of the force. In the category scheme this is achieved by not calculating the force due to stars which are in a category higher than the one in which the force is being calculated. Thus, a large number of unnecessary calculations is avoided. But it is clear that not all the particles in the lower categories are responsible for the necessity of recalculating the force. The advantage of our scheme is that it isolates, to a very great extent, only those particles which are indeed responsible for the rapid change in the force.

2. METHOD OF INTEGRATION

We surround each particle with a sphere of radius R_i and divide the total force, \mathbf{F}_i , on the particle into two parts: the irregular force, \mathbf{S}_i , which is the sum of the forces due to the particles within R_i and the regular force, \mathbf{K}_i , due to the rest of the system. If a star is approaching the i th star with a high velocity it will cause a rapid change in the irregular force, and it is, therefore, included as a neighbor even though it lies outside R_i . If computer storage is at a premium it is prudent to limit the maximum number of stars to be neighbors. For systems containing 1000 stars or less, a choice of R_i which gives approximately a maximum of 15 neighbors has been found to be sufficient. Initially, all R_i s are taken equal but in later stages of evolution high density in some regions usually forces a decrease in some of the R_i s. We shall describe in the next section the effects on the integration of choosing different maximum values of R .

As we have divided the force into two parts we shall have two different time steps associated with each particle. The regular time step, T_i , is the time from the last regular force calculation, for which the exact regular force is valid. Similarly,

t_i applies to the irregular force. The regular time step is calculated from

$$\Delta T_i = \beta \frac{|\mathbf{K}_i|}{|\dot{\mathbf{K}}_i|}, \quad (2)$$

where $\dot{\mathbf{K}}_i$ is the time derivative of the regular force and β is a constant which is used to control the accuracy of the integration. In cases where (2) yields a very large value for ΔT_i , which may occur if K_i is very small, we choose [1] the ΔT_i to be

$$\Delta T_i = \beta \frac{|\mathbf{r}_i|}{\langle v^2 \rangle^{1/2}}, \quad (3)$$

where $\langle v^2 \rangle^{1/2}$ is the root mean square velocity of the system. In practice, both (2) and (3) are calculated and the smaller of the two is chosen.

The formulas used by Lecar and Gonzalez [3] in their analytic program are used to find the irregular time steps

$$\delta t_i = \alpha [\min(r_{ij})]^{3/2}, \quad (4)$$

where $\min(r_{ij})$ is the distance between the i th particle and its closest neighbor, j . The constant α , along with β of (2), serves as an error control of the integration.

In case a neighbor of i has a very high velocity the irregular time step is obtained from

$$\delta t_i = \frac{\eta \alpha}{m_i^{1/2}} \min \left[\frac{r_{ij}}{v_{ij}} \right] \quad (5)$$

and is used if the δt_i given by (5) is indeed smaller than that given by (4). The square root of the mass is included in (5) rather than (4) for the sake of convenience. η is constant which we have found to be best taken as 3.3.

To evolve the positions and velocities from T_0 to T a Taylor series up to the fourth order in the force derivatives is used

$$\mathbf{r}_i(T) = \mathbf{r}_i(T_0) + \mathbf{v}_i(T_0) \Delta T + \frac{1}{2} \frac{\mathbf{F}_i(T_0)}{m_i} (\Delta T)^2 + \sum_{k=1}^4 \frac{\mathbf{F}_i^{(k)}(T_0) (\Delta T)^{k+2}}{m_i (k+2)!} \quad (6)$$

$$\mathbf{v}_i(T) = \mathbf{v}_i(T_0) + \frac{\mathbf{F}_i(T_0)}{m_i} \Delta T + \sum_{k=1}^4 \frac{\mathbf{F}_i^{(k)}(T_0)}{m_i} \frac{(\Delta T)^{k+1}}{(k+1)!}, \quad \Delta T = T - T_0 \quad (7)$$

where \mathbf{F}_i is the total force on the particle and $\mathbf{F}_i^{(k)}$ is the k th derivative of the force with respect to time, evaluated at time T_0 . The forces and their derivatives consist of the sum of the regular and irregular parts

$$\mathbf{F}_i^{(k)} = \mathbf{K}_i^{(k)} + \mathbf{S}_i^{(k)} \quad k = 0, 1, 2, 3, 4. \quad (8)$$

In between the regular force calculations the divided difference polynomial method is used to extrapolate the force and its derivatives. The force polynomial is defined as

$$\begin{aligned}
 \mathbf{F}(T) = & \mathbf{F}(T_0) + \mathbf{D}[T_0, T_1](T - T_0) \\
 & + \mathbf{D}^2[T_0, T_2](T - T_0)(T - T_1) \\
 & + \mathbf{D}^3[T_0, T_3](T - T_0)(T - T_1)(T - T_2) \\
 & + \mathbf{D}^4[T, T_3](T - T_0)(T - T_1)(T - T_2)(T - T_3)
 \end{aligned} \tag{9}$$

where T_0, T_1, T_2, T_3 are the times, in decreasing order of the last four exact force calculations. The divided differences are defined as follows

$$\begin{aligned}
 \mathbf{D}[T_0, T_1] &= \frac{\mathbf{F}(T_0) - \mathbf{F}(T_1)}{T_0 - T_1}, \\
 \mathbf{D}^2[T_0, T_2] &= \frac{\mathbf{D}[T_0, T_1] - \mathbf{D}[T_1, T_2]}{T_0 - T_2}, \\
 \mathbf{D}^3[T_0, T_3] &= \frac{\mathbf{D}^2[T_0, T_2] - \mathbf{D}^2[T_1, T_3]}{T_0 - T_3}, \\
 \mathbf{D}^4[T, T_3] &= \frac{\mathbf{D}^3[T, T_2] - \mathbf{D}^3[T_0, T_3]}{T - T_3}.
 \end{aligned} \tag{10}$$

The advantage of using (9) and (10) is the ease in calculation. To obtain the force derivatives in terms of the divided differences we compare (9) with the Taylor expansion of the force

$$\mathbf{F}(T) = \mathbf{F}(T_0) + \sum_{k=1}^4 \mathbf{F}^{(k)} \frac{(T - T_0)^k}{k!} \tag{11}$$

to obtain

$$\begin{aligned}
 \mathbf{F}^{(1)} &= \mathbf{D}[T_0, T_1] + T_1' \mathbf{D}^2[T_0, T_2] + T_1' T_2' \mathbf{D}^3[T_0, T_3] + T_1' T_2' T_3' \mathbf{D}^4[T, T_3], \\
 \mathbf{F}^{(2)} &= 2! \{ \mathbf{D}^2[T_0, T_2] + (T_1' + T_2') \mathbf{D}^3[T_0, T_3] \\
 & \quad + (T' T_2' + T_2' T_3' + T_1' T_3') \mathbf{D}^4[T, T_3] \}, \\
 \mathbf{F}^{(3)} &= 3! \{ \mathbf{D}^3[T_0, T_3] + (T_1' + T_2' + T_3') \mathbf{D}^4[T, T_3] \}, \\
 \mathbf{F}^{(4)} &= 4! \mathbf{D}^4[T, T_3],
 \end{aligned} \tag{12}$$

where $T_k' = T_k - T_0$ for $k = 1, 2, 3$. It should be noted that to obtain the last term of (12) a semiiteration procedure is used. This is so since the calculation of

D^4 requires a knowledge of the force at time T , which of course is unknown at time T_0 , the time of the other $D^{(k)}$ calculations.

The positions and velocities are updated in two stages. First, they are revised using (6) up to the F^3 term. The irregular force and its derivatives are then calculated using the new positions, allowing the computation of $D^4(T, T_3)$ for the irregular force. They are then corrected by including the F^4 term and using D^4 to

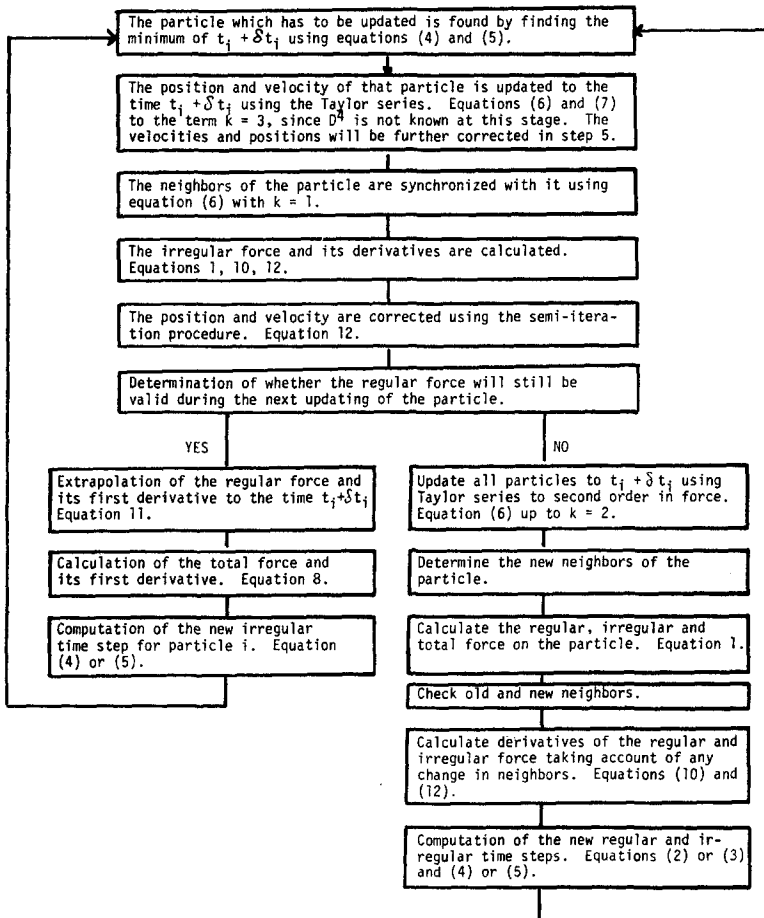


FIG. 1. Flow diagram of the integration. Stars are updated one at a time. At the beginning of the loop the position and velocity of the i th particle are known at time t_i but the positions and velocities of the other particles are not known at t_i . They must, therefore, be synchronized so that the new force on the particle being taken through the loop can be calculated. Also, at the beginning of the loop the regular and irregular forces and their first three divided differences are known.

correct the contributions of the lower order derivatives. If the regular force had to be revised then the same procedure is followed as for the irregular force.

The actual scheme then proceeds as follows. Figure 1 shows the flow of the program. Since the particles are not synchronized, we choose the particle whose position and velocity is to be revised first by finding the minimum of $t_i + \delta t_i$. That is, the particle whose irregular force needs recalculation before all others. In a previous stage a check has already been made to assure that the regular force is still good at this stage. The position and velocity are updated to the time $t_i + \delta t_i$ using the Taylor expansion (6) and (7) but carrying it out to only third order in force. The fourth-order term is taken into account at a later stage by the semi-iteration procedure described before.

The neighbors of i are updated to $t_i + \delta t_i$ using a Taylor series only up to first order in the force. This synchronization allows the calculation of the irregular force and its derivatives. We have found that for this purpose going to higher order in the Taylor series is not as efficient as controlling accuracy via α . The positions and velocities are then corrected by the semiiteration procedure to include the fourth-contribution term in the force.

We now check whether the next time around the regular force would need revision. This is done by checking $t_i + 2\delta t_i$ against $T_i + \Delta T_i$. If the latter is greater, then the regular force polynomial and its derivative are extrapolated to $t_i + \delta t_i$. The new δt_i for that particle is determined and the loop is begun again by finding the new particle with the minimum $t_i + \delta t_i$.

If the regular force polynomial has to be revised then all the particles are synchronized to $t_i + \delta t_i$ using (6) only to first order in the force. The neighbors of i are determined and any change of the previous neighbors is noted. If there are no new neighbors and all of the old neighbors are still neighbors, one straight forwardly calculates forces and derivatives. If there has been a change in neighbors, the forces and derivatives are first calculated assuming there has been no change. Then these are corrected by using analytic formulas to calculate contributions from old neighbors which are no longer neighbors and/or the new neighbors. These are then subtracted and added, respectively. New ΔT_i and δT_i are calculated for the particle and the cycle is repeated.

As the foregoing method is not self-starting, a separate initial routine is used to start the process. At each stage the main program requires a knowledge of force and its derivatives. The quantities are not known at $t = 0$ and are, thus, obtained from the analytic expression for force and its derivatives. This involves a straight-forward differentiation of (1). The explicit expressions can be found in Ref. [6]. Having calculated the force and derivatives for each particle the time steps and neighbors are then found and the main program is begun with these quantities as if they had been obtained through the procedure of the main program.

3. NUMERICAL RESULTS

The scaling we have used is to choose the gravitational constant and the total mass of the system equal to one. The total energy of the system defined by

$$E = \frac{1}{2} \sum_{i=1}^N m_i v_i^2 - \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{G m_i m_j}{r_{ij}} \quad (13)$$

is taken to be $-\frac{1}{8}$. The advantage of scaling is to make most of the relevant physical quantities independent of the number of particles in the system.

The basic time unit is taken to be the crossing time which is defined as

$$T_c = \frac{R}{V}, \quad (14)$$

where R is the "radius" of the system and V is the root mean square velocity. Assuming the virial theorem, (14) is equal to

$$T_c = GM^{5/2}(-2E)^{-3/2}, \quad (15)$$

which equals 8 with our choice of units.

The initial conditions for the tests described below were chosen from the following equilibrium distribution of position and velocity

$$\begin{aligned} f(\epsilon) &= A & \epsilon < \epsilon_0 \\ &= 0 & \epsilon > \epsilon_0, \end{aligned} \quad (16)$$

where

$$\epsilon = \frac{1}{2}v^2 + \psi(r), \quad \epsilon_0 = -MG/R_0 \quad (17)$$

and $\psi(r)$ is the potential energy per unit mass. A is a constant chosen so that the normalization of f equals the total mass and R_0 is the radius

$$\begin{aligned} R_0 &= 24/7, \\ A &= 0.01217. \end{aligned} \quad (18)$$

The mass distribution given by

$$\rho(\mathbf{r}) = \int f(\epsilon) dv \quad (19)$$

is the Emden polytrope of index 3/2. Although ρ cannot be explicitly found the differential equation it satisfies (Poisson's equation coupled with (19)) is readily

solved numerically. The advantage of using a distribution which is a function of ϵ is that the macroscopic properties of the system do not change too violently. In practice the positions of the N stars are generated first by randomly distributing them from the Emden distribution. The velocity of each star is then chosen randomly from the distribution given by

$$\begin{aligned} \rho(\mathbf{v}) &= A & \frac{1}{2}v^2 + \psi(r) < \epsilon_0 \\ &= 0 & \frac{1}{2}v^2 + \psi(r) > \epsilon_0, \end{aligned} \tag{20}$$

where r is the already known position of the star. The positions and velocities are then linearly adjusted to conform to a total energy of $-\frac{1}{8}$.

A. Computing Time as a Function of the Number of Particles

Due to the N^2 nature of the gravitational interaction, computing time for schemes based on individual time steps increase at least as fast as N^2 . Investigators have reported a dependence varying from N^2 to N^3 . But these estimates of computing time are not based on extensive studies, for only a few values of N are used and the number of cases integrated for each N is small. A careful study would require an enormous amount of computing time not commensurate with the interest in knowing the precise dependence. Figure 2 shows the N dependence of computing time for the method described above. Many cases were integrated for the low values of N but only a few for $N = 600$ and 1000. The logarithm of computing time per crossing time is plotted against the logarithm of the number of particles. The time for the 100 body case is taken to be 100 units of time. For the cases considered the energy error is constant within a factor of 2.

If we assume a power law for computing time, C_T , versus the number of particle, then approximately

$$C_T = \gamma \left(\frac{N}{100} \right)^{1.6} \tag{21}$$

where γ is about 40 sec for the CDC 6600 per crossing time and relative energy error of about 10^{-4} .

B. Computing Time as a Function of the Number of Neighbors

As the operational definition of irregular force is the force due to the particles within a sphere of radius R surrounding each particle, there obviously must be an optimum value of R depending on the average density. Choosing a too small value for R would result in the regular force changing too rapidly and, hence, a too frequent recalculation of the regular force. This is due to particles which are near the particles in question but, nonetheless, being included as part of the regular force. Choosing an R too large increases the number of neighbors to the point

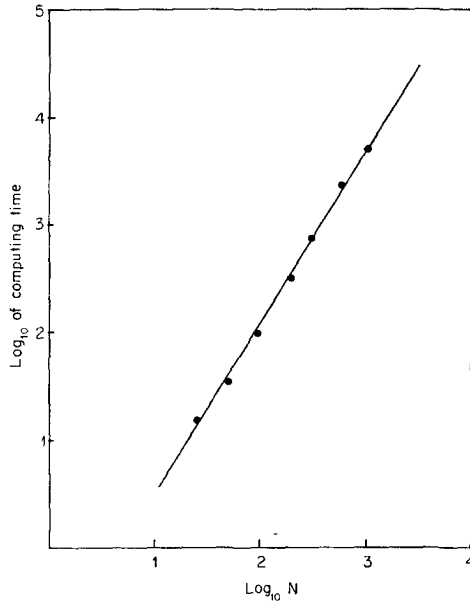


FIG. 2. Computing time as a function of the number of particles being integrated. The abscissa points represent the cases where $N = 25, 50, 100, 200, 300, 600,$ and $1000,$ respectively. The computing times are relative with 100 units of time for the 100 body case. In each case, the system was integrated for one crossing time.

that the time spent in updating the irregular force overwhelms the time saved by separation of the total force into two parts. Figure 3 shows a set of integrations for $N = 50$ and 200 where all the parameters of the program were kept the same except for maximum allowed number of neighbors. The number of neighbors is controlled by choosing different values of R^3 . In each case the figures show, as expected, that there is an optimum value of R which minimizes computing time.

C. Error Analysis

The analysis of the factors which contribute to errors in the numerical integration of the N -body problem is extremely difficult. As Miller [7, 8] has shown, two systems with slightly different initial conditions or two systems with identical initial conditions but integrated slightly differently will become increasingly divergent in phase space. The basic reason being that the instability is principally due to the errors caused by close encounters of two bodies. This was shown by Sandish [9] who repeated Miller's experiments with a modified force, the Aarseth potential, which has the effect of weakening the force law at close distances. The effect of a weaker force was to decrease the divergence in phase space. The

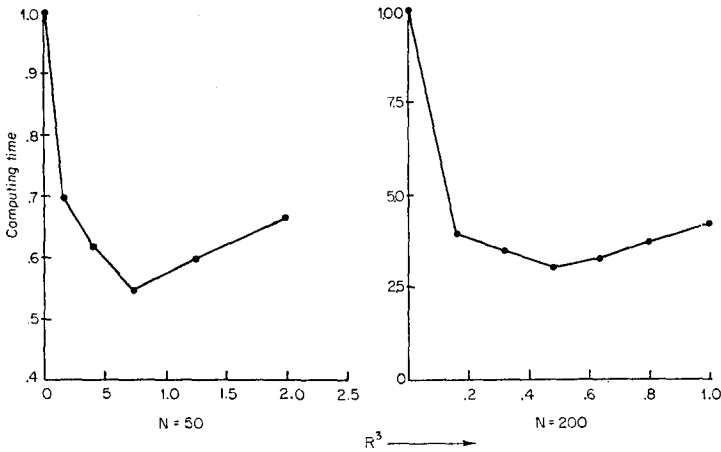


FIG. 3. Computing time as a function of R^3 . Computing times are relative.

physical reason for the instability is that the positions and velocities of the particles after the encounter are very sensitive to the coordinates before the encounter. Thus, as a particle suffers subsequent encounters the error gets progressively amplified. The obvious measure of error, microscopic reversibility, is, hence, much too stringent a test and the impossibility of it in actual practice makes it a useless criterion. It is generally accepted that the microscopic instability is not reflected in the macroscopic properties of the solution. Although this may be more a belief than a proven fact there are indications [2, 9] that this may be so.

As an alternative to using microscopic reversibility as a criterion for error, the total energy of the system given by (13) has been adopted by most investigators for measuring the error of integrations. But the use of energy as a measure is not totally satisfactory either. This is again due to the microscopic instabilities although the effect is not so great on the total energy as it is on positions and velocities. As an example of the situations which may arise we cite the following situation. Two cases with the same initial conditions may be integrated on the same computer, one in single precision and the other in double precision, but, nonetheless, the case done with double precision may show a larger error! This is usually due to the fact that single precision integration may have missed a close encounter, and, hence, its energy remained relatively stable. Due to the above facts one must therefore not expect exact relationships between the controls of the program and resulting errors. At best one would hope to find a general trend which may be a useful guide in practice. A detailed error analysis was made by Wielen for his method which uses single time steps. In our scheme we have two primary error control parameters, α and β , appearing in the regular and irregular time

steps. To study the dependence of the error parameter we kept one of the two parameters constant while integrating different cases of varying the other. In each instance the integrations were done for one crossing time and the energy was sampled eight times. We have used two different criteria for the energy error

$$\Delta E = \frac{1}{n} \sum_{i=1}^n |E_{i+1} - E_i| \tag{22}$$

$$\sigma_E = \frac{1}{n} \left[\sum_{i=1}^n (E_i - \bar{E})^2 \right]^{1/2}, \tag{23}$$

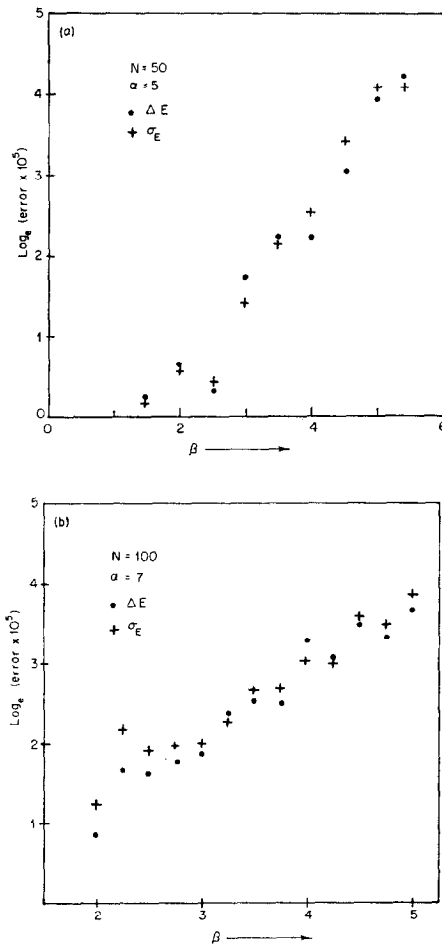


FIG. 4a, b. Error of integration as a function of the regular time step control parameter.

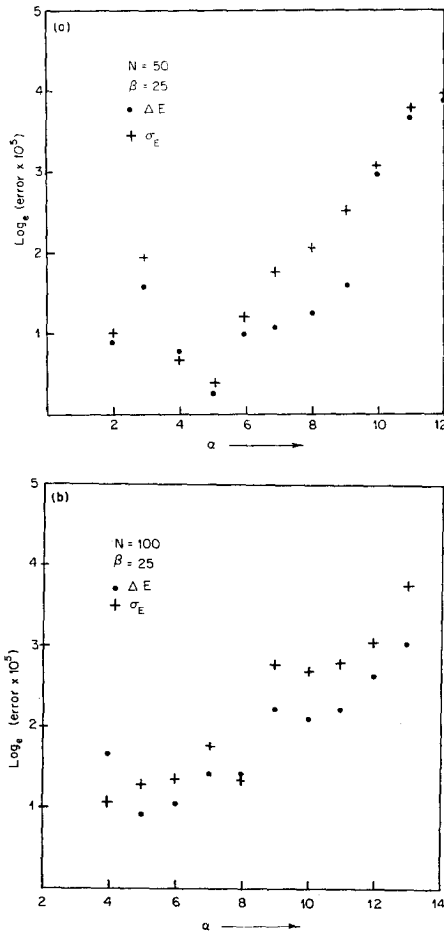


FIG. 5a, b. Error of integration as a function of the irregular time step control parameter.

where E_i is the energy at each of the samplings, \bar{E} the average of the E_i s and n is the number of samples taken. In (22), n equaled 8 but in (23), $n = 9$ as the initial energy, -0.1250 in our system of units, is used as one of the samples.

Figures 4a, b and 5a, b show the relative error as a function of the irregular and regular time step control parameters appearing in (2) and (4). The figures show a general trend even though there is quite a high scatter. The trend is reliable and has appeared in most cases that we have seen. Of course, there may be times when the relative error becomes very large but after some time it returns to its average relative error. This is a well known [6] phenomenon in *N*-body integrations and is usually due to a very close encounter.

ACKNOWLEDGMENTS

It is a pleasure to thank S. Aarseth, M. Lecar, and R. Wielen for numerous discussions relating to the above work. The help of V. Gavrilov on various aspects of the programming is gratefully acknowledged.

REFERENCES

1. S. J. AARSETH, *M.N.R.A.S.*, **123** (1963), 223.
2. R. WIELEN, *Veroeffentlichungen Astr. Rechen-Institut Heidelberg Nr. 19* (1967).
3. S. J. AARSETH AND F. HOYLE, *Astrophysica Norwegica* **29** (1964), 313.
4. A. HAYLI, *Bulletin Astronomique* **2** (1967), 67.
5. S. CHANDRASEKHAR, *Rev. Mod. Phys.* **15** (1943), 1.
6. M. LECAR AND C. C. GONZALEZ, *Bulletin Astronomique* **3** (1968), 209.
7. R. H. MILLER, *Astrophys. J.* **140** (1964), 250.
8. R. H. MILLER, *J. Computational Phys.* **8** (1971), 449.
9. E. M. STANDISH, thesis, Yale University, 1968.