

THE MONTE CARLO METHOD

M. HÉNON

Observatoire de Nice

Abstract. We give here a detailed technical description of a Monte Carlo scheme for the dynamical evolution of spherical stellar systems. The philosophy of the method, as well as a few illustrative results, are given elsewhere (Hénon, 1971, hereafter called I).

1. Superstars

Each actual computation will be made with a set of n objects, with n having some definite value. On the other hand, we would like the results to be applicable to a system with an arbitrary number of stars N . We shall therefore make the convention that each object in the computation is a *superstar*, representing K actual stars, with: $K=N/n$. If the various units are suitably chosen (see below), K does not appear in the numerical computations, and is thus left as a freely adjustable parameter.

Let r be the distance of a star to the centre, v_r its radial velocity, v_t its transverse velocity, and m its mass. A superstar will be characterized at any given time by the four quantities, r , v_r , v_t , m , and will be understood as a collection of K stars having these coordinates. Thus the mass of the superstar is: Km . The two angular coordinates defining the position of each star on the sphere with radius r are not specified; the stars are assumed to be randomly distributed on the sphere. Similarly, the angular coordinate defining the orientation of the transverse velocity is assumed to be randomly distributed.

In this way, the introduction of the superstars has another important advantage: it makes the system exactly spherically symmetrical.

2. Initial Conditions

The computation is started by selecting the four quantities r , v_r , v_t , m for each of the n superstars at time $t=0$. This is done by prescribing an initial distribution function $f(\mathbf{r}, \mathbf{v}, m, 0)$, and then selecting the coordinates randomly in accordance with this distribution.

The choice of the quasi-random number generator requires some care since a large quantity of these numbers is used in the course of the computation (typically about 10^6 for one run with $n=1000$). We have used, to our entire satisfaction, one of the generators studied by Coveyou and Macpherson (1967):

$$n_{i+1} = \lambda n_i + 1 \pmod{2^{35}}, \quad (1)$$

with

$$\lambda = 273\,673\,163\,155 \text{ (octal)}. \quad (2)$$

Astrophysics and Space Science **14** (1971) 151–167. All Rights Reserved
Copyright © 1971 by D. Reidel Publishing Company, Dordrecht - Holland

The n_i are integers. Quasi-random numbers with a uniform distribution between 0 and 1 are obtained by: $X_i = 2^{-35} n_i$.

One restriction on the form of the initial distribution function is that it must be a steady state; the present scheme is not equipped to handle the initial phase of rapid collective motions, since it proceeds by time steps much larger than the crossing time t_c , and we assume that these motions have already died out at time $t=0$. In the case of spherical symmetry, this means that the distribution function has the form:

$$f = \Psi(E, A, m), \quad (3)$$

where E and A are the energy per unit mass and the angular momentum per unit mass:

$$\begin{aligned} E &= U(r) + \frac{1}{2}(v_r^2 + v_t^2), \\ A &= rv_t; \end{aligned} \quad (4)$$

$U(r)$ is the gravitational potential.

The function Ψ in (3) can be arbitrary. In practice, however, the computation of the initial conditions will be easier if Ψ has a simple analytical form; in particular it is preferable to choose a form for which the corresponding potential $U(r)$ can be obtained explicitly.

3. Potential

We need to know the mean, smoothed-out potential, in order to compute the unperturbed motion of the stars. Smoothing in the transverse directions is immediately achieved by forgetting the discrete nature of each superstar and assimilating it to a spherical shell of radius r , of mass Km . Smoothing in the radial direction would also be required in principle. This, however, raises some practical difficulties. One of them is the precise definition of the smoothed potential in that case. Another is that the kinetic energy of a star, computed from its total energy by (4a), may turn out to be negative in some cases if the actual potential is replaced by the smoothed-out potential. On the other hand, experience indicates that radial smoothing is not really necessary (see Figure 4). The reason is probably that a system of concentric shells is already smoothed to a considerable degree; in particular, since it has spherical symmetry, the two integrals (4) of the motion of a star are exactly conserved.

Therefore we take the mean potential as being that of the system of n concentric shells. To compute it, we sort first the n shells in order of increasing radius. Let r_k and Km_k be the radius and mass of the shell of rank k from the centre. We also define: $r_0=0$ and $r_{n+1}=\infty$. Then, in the interval $r_k \leq r \leq r_{k+1}$, the potential is:

$$U(r) = KG \left(-\frac{1}{r} \sum_{i=1}^k m_i - \sum_{i=k+1}^n \frac{m_i}{r_i} \right). \quad (5)$$

Thus U , plotted against $1/r$, is a succession of straight segments, as schematically shown by Figure 1 for the case $n=3$. In consequence, it is sufficient to store the values

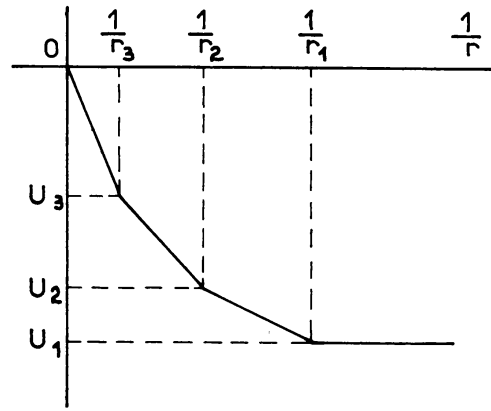


Fig. 1. Gravitational potential U as a function of $1/r$, for a system of three shells.

$U_k = U(r_k)$ of the potential on the shells in order to be able to compute quickly the potential at any point. These values are computed recursively by:

$$\left. \begin{aligned} U_{n+1} &= 0, \\ M_n &= M = K \sum_{i=1}^n m_i, \\ U_k &= U_{k+1} - GM_k \left(\frac{1}{r_k} - \frac{1}{r_{k+1}} \right), \\ M_{k-1} &= M_k - Km_k. \end{aligned} \right\} \quad (k = n, \dots, 1) \quad (6)$$

M_k is the mass inside of and including the shell of rank k . M is the total mass of the system. The potential at any point is then given by:

$$U(r) = U_k + \frac{1/r_k - 1/r}{1/r_k - 1/r_{k+1}} (U_{k+1} - U_k), \quad \text{for } r_k \leq r \leq r_{k+1}. \quad (7)$$

4. Time Step

We let now the system evolve during a time Δt , which must be small compared to the relaxation time. The crossing time is given as a function of the total mass M and the total energy \mathcal{E} of the system by:

$$t_c = C_2 GM^{5/2} |\mathcal{E}|^{-3/2}, \quad (8)$$

where C_2 is a numerical constant (van Albada, 1968), and the relaxation time is related to the crossing time by:

$$t_r/t_c = C_1 N/\ln N, \quad (9)$$

where C_1 is another numerical constant (Chandrasekhar, 1942). Therefore we take for the time step:

$$\Delta t = b \frac{N}{\ln N} GM^{5/2} |\mathcal{E}|^{-3/2}, \quad (10)$$

where b is a small constant. Experience shows that $b=0.005$ gives reasonable results.

Δt as given by (10) is not constant but decreases markedly with time because N and M decrease as a result of the escape of stars.

5. Relaxation

We assume that the mean potential does not change during Δt . Then the unperturbed motion of each star would be a rosette motion, characterized by constant values of E and A . In fact, however, the perturbations will slightly change E and A during the time Δt . In order to simulate this effect in Monte Carlo fashion, we should first select at random the position of each star along its orbit. For the first time step, however, this is not necessary since the initial coordinates have been already chosen at random according to a given steady state distribution function. Thus we take each superstar where it is at $t=0$. We shall compute the effect of the perturbations during a small time δt , satisfying:

$$\delta t \ll t_c. \quad (11)$$

Later on we shall multiply the effect by the correcting factor $\Delta t/\delta t$, in order to compensate for the rest of the time interval.

Next, we must consider each star in turn as a test-star, and then select at random another star as a representative of the field stars. We call $(\mathbf{r}, \mathbf{v}, m)$ and $(\mathbf{r}', \mathbf{v}', m')$ the coordinates of the test star and the field star respectively. As is well known, the total perturbation from the field stars up to a distance l is approximately proportional to $\ln(Nl/R)$, where R is a typical radius of the system; thus, although all stars contribute to the perturbation, the largest part of it comes from relatively near stars. For example, if $N=10^5$, 80% of the perturbation comes from the field stars nearer than $R/10$. Therefore we shall make the simplifying assumption that the distribution of field stars with respect to mass and velocity is everywhere the same as where the test star is. In other words, we take the distribution function of field stars to be not $f(\mathbf{r}', \mathbf{v}', m')$, but $f(\mathbf{r}, \mathbf{v}', m')$. The advantage is that this last distribution is space-independent, so that we can now select separately the position of the field star on the one hand, its velocity and its mass on the other.

We select first the mass and the velocity, with a weight function proportional to the local distribution function. There is a very simple way to do this: we just select the mass and the velocity of the nearest star. A moment's reflexion will show that the probability to obtain given values m' and \mathbf{v}' is proportional to the local distribution function.

In fact, we do not know the positions of the stars, but only the radii of the superstars. Besides, it will be seen below that the choice of the nearest star leads to a divergence in the equations, which can be avoided by taking instead the p -th nearest star, with $p > 1$. Thus our practical rule will be: for a test star at r_k , take the mass and the velocity of the field star either at r_{k+p} or r_{k-p} . The probability to obtain

given values of m' and \mathbf{v}' is still proportional to the local distribution function, provided that p is not too large.

This procedure does not determine completely the velocity of the field star, because only v_r and v_t are specified for a superstar. More specifically, we do not know the angle φ between the orbital planes of the test star and of the field star, defined by (\mathbf{r}, \mathbf{v}) and $(\mathbf{r}, \mathbf{v}')$. However, because of the spherical symmetry, and the correlated assumption that the orbital planes of the stars belonging to a given superstar are randomly distributed, the angle φ will also be randomly distributed, and we simply select its value at random between 0 and 2π with a uniform probability.

We must now select the position of the field star. This position could be simply taken at random, with a uniform probability, inside the volume occupied by the system. However, a general rule in Monte Carlo computations says that whenever possible, a mean quantity should be computed exactly rather than estimated by sampling, in order to minimize the variance of the final results. It happens here that the mean effect of a field star with given mass and velocity can be computed rather simply. Therefore, instead of taking the position at random, we shall choose it in such a way that the effect of the field star will be equal to the computed mean effect.

We consider a moving frame of reference in which the centre of mass of the test star and the field star is at rest. In this frame, the modulus of the velocity of the test star does not change during the encounter; the direction of the velocity is deflected by an angle β , generally small, given by:

$$\tan \frac{\beta}{2} = \frac{G(m + m')}{w^2 l} = \frac{l_0}{l}, \quad (12)$$

where l is the impact parameter and w is the relative velocity:

$$w = |\mathbf{v}' - \mathbf{v}|; \quad (13)$$

thus l_0 defined by (12) is independent of the position of the field star. Let ν be the local number of field stars per unit volume. Then, since we sample one field star at random out of N , the probability to find it in a volume dV is: $\nu dV/N$; and the probability for the test star to 'meet' the field star during the time δt , with an impact parameter between l and dl , is:

$$2\pi l dl w \delta t \nu N^{-1}. \quad (14)$$

Therefore the mean squared deflection is:

$$\langle \beta^2 \rangle = 8\pi l_0^2 w \delta t \nu N^{-1} \int (dl/l); \quad (15)$$

we have made here the approximation: $\beta/2 \approx \tan \beta/2$. The appropriate cut-offs for the integral are respectively of the order of R/N and R (Hénon, 1958), so that the integral is approximately equal to $\ln N$.

We multiply this now by the correcting factors $\Delta t/\delta t$ for the rest of the motion,

and N for the other field stars, obtaining:

$$\langle \beta^2 \rangle = 8\pi l_0^2 w \Delta t v \ln N. \quad (16)$$

This value will also result if we take a field star with a fixed impact parameter equal to:

$$l = (2\pi w \Delta t v \ln N)^{-1/2}. \quad (17)$$

There remains only the problem of estimating v , the local number density of the field stars. It would be difficult to compute this density accurately, since there is only a limited number of superstars in the computation. Therefore we revert here to a sampling procedure: we shall estimate v from the radial distance of the p -th nearest superstar, the same which has already been used to provide the mass and the velocity of the field star (this is not necessary, but convenient). Let Δr be this distance:

$$\Delta r = |r_{k \pm p} - r_k|. \quad (18)$$

Clearly the estimated density, v^* , should be inversely proportional to Δr :

$$v^* = C/\Delta r. \quad (19)$$

Since v enters linearly in (16), the constant C must be adjusted so that

$$\langle v^* \rangle = v. \quad (20)$$

Let σ be the number of superstars per unit interval in r ; σ is related to v by:

$$K\sigma = 4\pi r^2 v. \quad (21)$$

The probability for the p -th nearest superstar to be at a radial distance Δr is given by the Poisson formula:

$$e^{-x} \frac{x^{p-1}}{(p-1)!} dx, \quad (22)$$

with: $x = \sigma \Delta r$. Therefore:

$$\langle v^* \rangle = C\sigma \int_0^{\infty} e^{-x} \frac{x^{p-2}}{(p-1)!} dx = \frac{C\sigma}{p-1}. \quad (23)$$

From (20), (21) and (23) we obtain:

$$C = K \frac{p-1}{4\pi r^2}. \quad (24)$$

Using (19) and substituting v^* for v in (17), we finally have for the impact distance:

$$l = \left[\frac{(p-1) K w \Delta t \ln N}{2r^2 \Delta r} \right]^{-1/2}. \quad (25)$$

Equation (23) shows why we cannot use $p=1$, i.e. the nearest superstar: $\langle v^* \rangle$ would

be infinite. (This fact had been overlooked in Hénon (1966) and Equation (3) in that paper is wrong). We have used the next value, $p=2$, in the computations. Superstars are associated in pairs, according to their rank from the centre, as follows: 1–3, 2–4, 5–7, 6–8, etc. If n is not a multiple of 4, one or two superstars are left out at the outside; the corresponding error is negligible since n is large, and also because relaxation effects are very small in the outer parts of the system. For each pair, a single encounter is computed, and the velocities of both superstars are perturbed accordingly; in other words, each member of the pair acts as a field star for the other. This symmetrical scheme saves a factor 2 in computing time, and has also the advantage that the total energy of the system is exactly conserved during this stage of the computation.

In order to specify completely the encounter, one must know, in addition to the impact parameter l , the angle ψ of the plane of relative motion ($\mathbf{r}' - \mathbf{r}$, $\mathbf{v}' - \mathbf{v}$) with some reference plane. Because of the assumed homogeneous distribution of the field stars, ψ is uniformly distributed, and therefore is chosen at random between 0 and 2π .

The computation of an encounter for two superstars (r, v_r, v_t, m) and (r', v'_r, v'_t, m') proceeds in practice as follows. We consider a system of rectangular axes, x, y, z with the z axis parallel to \mathbf{r} and the (x, z) plane parallel to \mathbf{v} . In this system the velocities of the two stars are:

$$\mathbf{v} = (v_r, 0, v_t), \quad \mathbf{v}' = (v'_t \cos \varphi, v'_t \sin \varphi, v'_r). \quad (26)$$

φ is computed by:

$$\varphi = 2\pi X, \quad (27)$$

where X is a random number between 0 and 1. We compute the relative velocity:

$$\mathbf{w} = (v'_t \cos \varphi - v_t, v'_t \sin \varphi, v'_r - v_r) = (w_x, w_y, w_z), \quad (28)$$

and its modulus w . We compute the impact parameter l from (25), with:

$$\Delta r = |r' - r|, \quad (29)$$

and r replaced by the mean value $(r+r')/2$ for the two superstars. We compute the deflection angle β from (12).

We compute the quantity:

$$w_p = (w_x^2 + w_y^2)^{1/2}. \quad (30)$$

Let us now consider the two vectors:

$$\begin{aligned} \mathbf{w}_1 &= (w_y w / w_p, -w_x w / w_p, 0), \\ \mathbf{w}_2 &= (-w_x w_z / w_p, -w_y w_z / w_p, w_p). \end{aligned} \quad (31)$$

It is readily verified that \mathbf{w}_1 and \mathbf{w}_2 have the same modulus as w and that the three vectors \mathbf{w} , \mathbf{w}_1 , \mathbf{w}_2 are mutually perpendicular. We take the plane $(\mathbf{w}, \mathbf{w}_1)$ as origin for the angle ψ , which is computed by:

$$\psi = 2\pi X, \quad (32)$$

where X is a random number between 0 and 1. The new relative velocity of the two stars after the encounter is then:

$$\mathbf{w}^* = \mathbf{w} \cos \beta + \mathbf{w}_1 \sin \beta \cos \psi + \mathbf{w}_2 \sin \beta \sin \psi. \quad (33)$$

Therefore the new velocities of the two stars are:

$$\begin{aligned} \mathbf{v}^* &= \mathbf{v} - \frac{m'}{m + m'} (\mathbf{w}^* - \mathbf{w}), \\ \mathbf{v}'^* &= \mathbf{v}' + \frac{m}{m + m'} (\mathbf{w}^* - \mathbf{w}). \end{aligned} \quad (34)$$

Using (33) and (34), we compute the three components of \mathbf{v}^* and \mathbf{v}'^* . Finally, we compute the new radial and transverse velocities of the first star:

$$v_r^* = v_z^*, \quad v_t^* = (v_x^{*2} + v_y^{*2})^{1/2}, \quad (35)$$

the new energy and angular momentum of the first star:

$$E^* = U(r) + \frac{1}{2} (v_r^{*2} + v_t^{*2}), \quad A^* = r v_t^*, \quad (36)$$

and similar quantities for the second star.

6. Escape

We shall consider two cases:

A. ISOLATED SYSTEM

It may happen that the new energy of a star given by (36) is positive. In that case the star will escape from the system in a time of the order of t_c ; on the relaxation time scale t_r , it can be considered as immediately lost. Therefore the superstar is removed from the computations. n is decreased by one unit.

B. NON-ISOLATED SYSTEM

If the system is subjected to an external gravitational field, escape is facilitated: in some directions the tidal force is away from the centre and exceeds the attraction of the system past a given distance. The tidal field is not spherically symmetrical in general; therefore it can only be represented in an approximate way in the present scheme which assumes spherical symmetry. We shall assume that all stars which go beyond a certain distance r_e from the centre escape from the system. r_e should be proportional to $M^{1/3}$ (von Hoerner, 1958), and we define it by:

$$r_e = r_{e0} (M/M_0)^{1/3}, \quad (37)$$

where M_0 is the initial mass of the system, and r_{e0} is a given constant, depending on the assumed strength of the tidal field. Then a star escapes iff:

$$2E - 2U(r_e) - A^2/r_e^2 > 0 \quad (38)$$

(see next paragraph). Since the whole system is contained inside r_e , this can also be written:

$$2E + 2GM/r_e - A^2/r_e^2 > 0. \quad (39)$$

7. New Positions

We consider now a new time step Δt , and this time we must begin by selecting a new position for each star.

Consider a particular star, which has values of E and A just computed by (36) (we drop the asterisks now). We assume that it moves in the smoothed-out potential $U(r)$ which has been computed above. The star describes a rosette orbit, with r oscillating between two extreme values r_{\min} and r_{\max} , which are the roots of:

$$Q(r) = 2E - 2U(r) - A^2/r^2 = 0. \quad (40)$$

We begin by computing r_{\min} and r_{\max} . If we introduce a new variable $z = 1/r$, then it is apparent from (40) and Figure 1 that d^2Q/dz^2 is always negative. On the other hand, Q is positive at the old position according to (36); for $z=0$, $Q=2E < 0$; and for $z = \infty$, $Q = -\infty$. It follows that Equation (40) has always exactly two roots (cf. Contopoulos, 1954).

The interval in which r_{\min} falls is found by looking up the tables of the ordered radii r_k and of the corresponding potentials U_k ; that is, one determines k such that:

$$Q(r_k) < 0 < Q(r_{k+1}). \quad (41)$$

The potential is then given by (7), which can be written:

$$U = az + b, \quad (42)$$

with:

$$a = \frac{U_{k+1} - U_k}{z_{k+1} - z_k}, \quad b = \frac{U_k z_{k+1} - U_{k+1} z_k}{z_{k+1} - z_k}; \quad (43)$$

substituting into (40), we obtain a quadratic equation for z . Moreover, dQ/dr must be positive for $r = r_{\min}$, and therefore the largest root for z must be taken, i.e.:

$$\frac{1}{r_{\min}} = \frac{-a + [a^2 - 2A^2(b - E)]^{1/2}}{A^2}. \quad (44)$$

A similar computation gives r_{\max} ; the inequality signs in (41) are then reversed and the plus sign in (44) is replaced by a minus sign. For reasons of computational accuracy it is in fact preferable to write r_{\max} in the equivalent form:

$$r_{\max} = \frac{-a + [a^2 - 2A^2(b - E)]^{1/2}}{2(b - E)}. \quad (45)$$

Now we must select a position of the star between r_{\min} and r_{\max} . The probability to take it in an interval dr should be equal to the fraction of time spent by the star

in dr , i.e.:

$$\frac{dt}{T} = \frac{dr/|v_r|}{\int_{r_{\min}}^{r_{\max}} dr/|v_r|}, \quad (46)$$

with the radial velocity v_r given by:

$$|v_r| = [2E - 2U(r) - A^2/r^2]^{1/2} = [Q(r)]^{1/2}. \quad (47)$$

The computation of the half-period T would be time-consuming; fortunately it can be avoided by the classical von Neumann rejection technique (Hammersley and Handscomb, 1964). We want a probability distribution proportional to a known function $f(r)$ (in the present case: $f(r) = 1/|v_r|$), without knowing the constant of

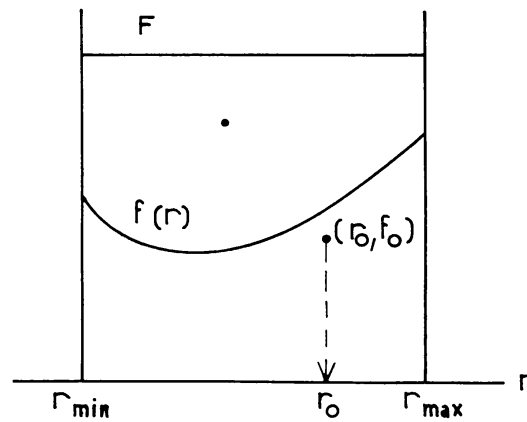


Fig. 2. The rejection technique.

proportionality. We take a number F which is everywhere larger than $f(r)$ (Figure 2). We select a point (r_0, f_0) at random in the rectangle $r_{\min} < r_0 < r_{\max}$, $0 < f_0 < F$, with a uniform distribution, i.e. we compute:

$$r_0 = r_{\min} + (r_{\max} - r_{\min}) X, \quad f_0 = F X', \quad (48)$$

where X and X' are a pair of normalized random numbers. If the point is below the curve: $f_0 < f(r_0)$, we take $r = r_0$ as the selected value. In the opposite case, this value is rejected, a new point in the rectangle is tried with a fresh pair of random numbers, and so on until a point below the curve is obtained.

In the present case, however, this procedure cannot be applied directly because $f(r) = 1/|v_r|$ becomes infinite at both ends of the interval. This difficulty is eliminated by the introduction of a new variable s , defined by an appropriate relation $r = r(s)$. The probability to sample a value in ds should be proportional to:

$$\frac{1}{|v_r|} \frac{dr}{ds} = g(s); \quad (49)$$

therefore $r(s)$ should be such that (49) remains finite in the whole interval. Since dQ/dr has a finite slope, $|v_r|$ is proportional to $(r-r_{\min})^{1/2}$ for $r \rightarrow r_{\min}$; therefore there should be: $r-r_{\min} \propto (s-s_{\min})^2$ for $r \rightarrow r_{\min}$, and similarly for $r \rightarrow r_{\max}$. Also r should be an easily computable function of s . A function $r(s)$ satisfying these requirements is:

$$r = \alpha_0 + \alpha_1 s + \alpha_2 s^2 + \alpha_3 s^3, \quad (50)$$

with a minimum equal to r_{\min} and a maximum equal to r_{\max} (Figure 3). We normalize

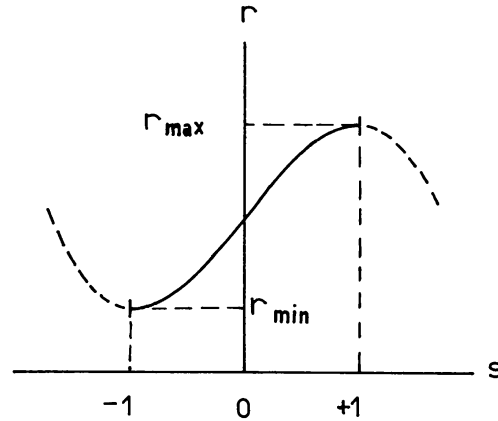


Fig. 3. Relation between r and s .

s by the conditions:

$$s_{\min} = -1, \quad s_{\max} = 1; \quad (51)$$

the coefficients in (50) are then determined and the relation becomes:

$$r = \frac{1}{2}(r_{\min} + r_{\max}) + \frac{1}{4}(r_{\max} - r_{\min})(3s - s^3). \quad (52)$$

For $s \rightarrow -1$ it is easily computed from (47), (49) and (52) that:

$$g(s) \rightarrow g(-1) = [3(r_{\max} - r_{\min}) / (dQ/dr)_{r=r_{\min}}]^{1/2}, \quad (53)$$

and for $s \rightarrow +1$:

$$g(s) \rightarrow g(+1) = [-3(r_{\max} - r_{\min}) / (dQ/dr)_{r=r_{\max}}]^{1/2}. \quad (54)$$

dQ/dr is computed from (40) and (42):

$$\frac{dQ}{dr} = \frac{2A^2}{r^3} + \frac{2a}{r^2}. \quad (55)$$

An upper boundary for $g(s)$ is computed by the empirical formula:

$$F = 1.2 \max[g(-1), g(+1)]. \quad (56)$$

A check is kept on all computed values of $g(s)$ to verify that they do not exceed F . Experience shows that formula (56) is entirely satisfactory in that respect.

In practice, the computation proceeds as follows. First r_{\min} and r_{\max} are computed. Next, F is computed from (53), (54), (55) and (56). Then the rejection procedure is applied: a pair of random values (s_0, g_0) is computed by:

$$s_0 = -1 + 2X, \quad g_0 = FX'; \quad (57)$$

corresponding values of r , dr/ds , $|v_r|$, and $g(s_0)$ are computed by (52), (47), and (49); and as soon as a pair is found for which there is: $g_0 < g(s_0)$, the corresponding value of r is taken as the new position of the star. The new radial velocity is: $v_r = \pm |v_r|$, since the motion can be outwards or inwards; the sign is chosen at random with a probability $\frac{1}{2}$ for each. Finally, the new transverse velocity is computed by:

$$v_t = A/r. \quad (58)$$

This procedure is repeated for all superstars. The same potential $U(r)$ is used throughout; it is not updated after the displacement of each superstar. The reason is that such an updating would mean changing part of the tables r_k and U_k after each displacement, and thus would involve a total number of operations proportional to n^2 ; this is to be avoided since in all other parts of the program the number of operations is proportional only to n or to $\ln n$. The effect is that the potential $U(r)$ used lags slightly behind the real potential, by a time equal to one half of the time step Δt in the mean. This effect probably has no serious consequences, since Δt is small compared to the relaxation time t_r , which is the time scale for the changes in the potential.

8. Continuation

Now that we have new positions and velocities for the superstars, the loop is closed by going back to the computation of the potential described above, and a new cycle of operations begins. The cycle is repeated again and again (typically a few hundred times), until the desired evolution of the system has been achieved.

9. Units

All symbols appearing in the equations so far are to be understood, as usual, as physical quantities. They will be related to the numbers appearing in the numerical computations by relations such as:

$$r = U_l[r], \quad (59)$$

where r is the physical quantity, $[r]$ is the corresponding number, and U_l is the adopted unit of length.

The units should be chosen in such a way that the numbers computed are the same whatever the actual mass of the system considered, its dimensions, and the number N of its stars. The first two requirements are easily met by an appropriate choice of the units of mass and length. The third one is a little more tricky, because

N appears explicitly in some equations, namely (6b), (6d) (remembering that $K = N/n$), (10) and (25); it can be met only by using a non-coherent system of units.

First, we define *basic units* of mass, length and time by the condition that the gravitational constant G , the total mass M_0 and the total energy \mathcal{E}_0 of the system at time $t=0$, expressed in these units, should be respectively:

$$[G] = 1, \quad [M_0] = 1, \quad [\mathcal{E}_0] = -\frac{1}{4}. \quad (60)$$

This gives the following expressions for the basic units:

$$U_m = M_0, \quad U_l = GM_0^2 (-4\mathcal{E}_0)^{-1}, \quad U_t = GM_0^{5/2} (-4\mathcal{E}_0)^{-3/2}. \quad (61)$$

U_l is of the order of the dimensions of the system, and U_t is of the order of the crossing time (see Equation (8)).

We shall use, for the various physical quantities appearing in the equations, the units given in Table I; N_0 is the number of stars at time $t=0$. The choice has been

TABLE I

mass of the system, M ;	}	U_m
mass of a superstar;		
partial mass, M_k (Equation (6))		
mass of a star, m		$U_m N_0^{-1}$
distance to centre, r ;	}	U_l
distance between shells, Δr		
impact parameter, l		$U_l N_0^{-1}$
time step, Δt ;	}	$U_t N_0 / \ln N_0$
evolution time, t		
velocities v_r, v_t, w		$U_l U_t^{-1}$
energy per unit mass, E ;	}	$U_l^2 U_t^{-2}$
potential, U		
angular momentum per unit mass, A		$U_l^2 U_t^{-1}$
gravitational constant, G		$U_m^{-1} U_l^3 U_t^{-2}$

guided by physical considerations, and it can be readily verified that it eliminates N (and also G) from all equations.

In fact, Equations (6), (10), (12), (25) become:

$$\begin{aligned}
 [U_{n+1}] &= 0, \\
 [M_n] &= [M] = \frac{1}{n_0} \sum_{i=1}^n [m_i], \\
 [U_k] &= [U_{k+1}] - [M_k] \left(\frac{1}{[r_k]} - \frac{1}{[r_{k+1}]} \right), \\
 [M_{k-1}] &= [M_k] - \frac{1}{n_0} [m_k];
 \end{aligned}
 \left. \vphantom{\begin{aligned} [U_k] \\ [M_{k-1}] \end{aligned}} \right\} (k = n, \dots, 1) \quad (62)$$

$$[\Delta t] = b \frac{n}{n_0} [M]^{5/2} [|\mathcal{E}|]^{-3/2}; \quad (63)$$

$$\tan \frac{\beta}{2} = \frac{[m] + [m']}{[w]^2 [l]}; \quad (64)$$

$$[l] = \left\{ \frac{(p-1)[w][\Delta t]}{2[r]^2 n_0 [\Delta r]} \right\}^{-1/2}. \quad (65)$$

In (63) and (65) we have made the approximation: $\ln N = \ln N_0$, which is acceptable since N is large. The other equations used in the computation are unaffected, i.e. one has simply to put each quantity into brackets.

The mean relaxation time is (Hénon, 1958):

$$t_r = \frac{1}{4\pi \bar{\rho}} \frac{(\bar{v}^2)^{3/2} N}{G^2 M \ln N}, \quad (66)$$

where $\bar{\rho}$ is the mean density of the system. We estimate it by:

$$\bar{\rho} = \frac{M/2}{4\pi R^3/3}, \quad (67)$$

where R is the mean radius of the system, and from the virial theorem we have:

$$R \approx \frac{GM^2}{-4\mathcal{E}}, \quad \bar{v}^2 \approx \frac{-2\mathcal{E}}{M}. \quad (68)$$

Hence,

$$t_r = \frac{1}{24\sqrt{2}} \frac{N}{\ln N} GM^{5/2} (-\mathcal{E})^{-3/2}. \quad (69)$$

Thus, in our units and at time $t=0$, the mean relaxation time is:

$$[t_r(0)] = \frac{1}{3\sqrt{2}} = 0.236\dots \quad (70)$$

From (10) and (69) we derive also:

$$\frac{\Delta t}{t_r} = 24\sqrt{2}b. \quad (71)$$

Thus for $b=0.005$, the time step is equal to about $\frac{1}{8}$ of the mean relaxation time.

With the help of Equations (61) and Table I, the results of a Monte Carlo computation can be compared to any real system, or to any other theoretical model. If one considers for example a typical globular cluster with $N_0 = 10^5$, $\bar{m} = 0.5 m_\odot$ and $(\bar{v}^2)^{1/2} = 5$ km/s, there is: $U_m = M_0 = 5 \times 10^4 m_\odot$, $-4 \mathcal{E}_0 = 2.5 \times 10^6 m_\odot \text{ km}^2 \text{ s}^{-2}$, $U_l = 4.5$ pc, $U_t = 6.4 \times 10^5$ yr, $U_t N_0 / \ln N_0 = 5.5 \times 10^9$ yr.

10. Technical Problems

We describe now a few practical problems which have not been yet entirely solved.

A. DENSE CORE

The relaxation time at a given point varies as the inverse of the density. Thus, as a model evolves and begins to form a very dense core (see I), the relaxation time near the centre decreases sharply and eventually becomes shorter than the time step Δt . The computation then no longer represents the real evolution of the core. The main effect is to slow down artificially the evolution of the core.

The situation is aggravated, in the case of stars of equal masses, by the following subsidiary effect. For $t_r \ll \Delta t$, the computed impact distance l will be in general much smaller than l_0 , and the deflection angle β computed from (12) will be close to 180° . Thus the two stars merely exchange their velocities, and from the point of view of the distribution function nothing happens. For stars of unequal masses, the effect is not so radical, but it still tends to slow down the computed relaxation.

This effect can be eliminated by limiting the deflection angle to 90° . Experience shows that this device effectively increases the rate of evolution of the core, and it has been incorporated into our procedure.

Still, the main problem remains. One could, of course, take a time step Δt equal to a fraction of the central relaxation time, rather than a fraction of the mean relaxation time. However, this would practically bring the computation to a standstill as the central relaxation time becomes very small.

A similar difficulty exists in the exact N -body computations, where it has been partially remedied by the introduction of individual time steps: stars are recomputed much more frequently in the core than in the halo. Perhaps a similar device could be introduced in the Monte Carlo scheme.

B. SPURIOUS RELAXATION

As already mentioned, the mean potential used to compute the orbital motion of the stars is not completely smooth: it contains random fluctuations in the radial direction. These fluctuations are time-dependent since at every new time step the potential is recomputed from the new positions of the stars. They produce a spurious relaxation, in addition to the normal relaxation which is already incorporated into the program. In order to estimate the importance of this effect, trial computations have been run with the normal relaxation suppressed: the computation of the encounters is simply by-passed. The initial steady state should then in principle maintain itself indefinitely. Figure 4 shows the results for two cases with 100 and 1000 superstars respectively. The initial state is a polytrope of index 1, with an isotropic distribution of velocities and with equal masses. Each curve represents the radius of the sphere containing a given fraction of the total mass. For $n=1000$, the steady state is very well conserved; comparing with I, Figure 2, which represents the same case with encounters, one can see that the spurious relaxation is entirely

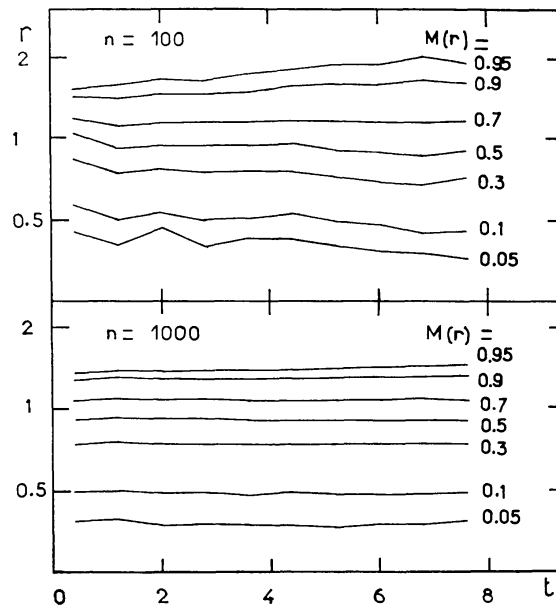


Fig. 4. Evolution of systems of 100 and 1000 superstars when the relaxation is suppressed.

negligible, at least during the first stages of the computation. For $n=100$, the spurious relaxation becomes visible (Figure 4). One can define a 'spurious relaxation time', which, from experimental results as well as from theoretical arguments, appears to be of the order of $n\Delta t$. Therefore the effect will be small compared to real relaxation if:

$$n\Delta t \gg t_r. \quad (72)$$

Since Δt is of the order of $t_r/6$, this criterion is easily satisfied. One can also require that the cumulative effect remains small to the end of the computation; this leads to the more stringent criterion:

$$n\Delta t \gg t_{\max}, \quad (73)$$

where t_{\max} is the length of the computed evolution. Equation (73) means: the number of superstars should be large compared to the total number of time steps. A typical computation corresponds to 200 to 300 time steps; therefore the number of superstars should be at least 1000, and preferably more.

A particular consequence of the spurious relaxation is a slow but systematic drift of the total energy upwards, superimposed on random fluctuations. The explanation is simple: each star interacts with a fluctuating field and tends in the mean to be accelerated, just as with Fermi's mechanism for the acceleration of the cosmic rays. This effect becomes more serious as a dense core develops, because the fluctuations of the potential near the centre become quite wild. In order to neutralize it, we have sometimes adopted the procedure of renormalizing the system after each time step, so as to keep the total energy constant. This procedure, however, does not really eliminate the spurious relaxation, which is not uniform throughout the system; in

fact, because the spurious relaxation is much more active near the centre, the re-normalization produces an artificial contraction of the outer parts of the system, which becomes visible at the end of the evolution (I, Figures 2 and 6).

The total energy could also be made strictly constant by recomputing the potential after the displacement of each superstar. However, this would require a number of operations proportional to n^2 , as already mentioned; and again the spurious relaxation would probably not be really eliminated.

C. SPURIOUS ESCAPE

Stars can escape from a cluster by two different mechanisms: either a slow, gradual variation of the energy resulting from the cumulative effect of many encounters, or a sudden variation of the energy resulting from a single close encounter (Hénon, 1960; Hayli, 1970). The second effect is normally smaller by a factor of the order of $\ln N$, and therefore negligible. In the case of an isolated system, however, the first effect is inoperative (Hénon, 1960), so that the second effect becomes important.

The first effect depends only on the mean squared value of the deflections, and it is correctly reproduced in our procedure where we compute the impact distance l so as to obtain the same mean squared deflection as in reality. The second effect, however, depends on the shape of the tail of the distribution of deflections, or roughly speaking, on the number of large deflections. The Monte Carlo procedure as described above does not reproduce exactly the tail of the distribution of deflections; therefore it does not represent correctly the mechanism of escape by a single close encounter.

Perhaps it would be possible to modify the procedure so as to reproduce the tail exactly. But, since the effect is proportional to $1/\ln N$, it would then be necessary to specify a value of N , thus restricting the generality of the results.

Note added in proof. A similar scheme has been recently developed by Spitzer and his collaborators (Spitzer and Hart, 1971a, b; Spitzer and Shapiro, 1971; Spitzer and Thuan, 1971).

References

- Chandrasekhar, S.: 1942, *Principles of Stellar Dynamics*, Dover Publications.
 Contopoulos, G.: 1954, *Z. Astrophys.* **35**, 67.
 Coveyou, R. R. and Macpherson, R. D.: 1967, *J. Ass. Comput. Mach.* **14**, 100.
 Hammersley, J. M. and Handscomb, D. C.: 1964, *Monte Carlo Methods*, Methuen, London.
 Hayli, A.: 1970, *Astron. Astrophys.* **7**, 17.
 Hénon, M.: 1958, *Ann. Astrophys.* **21**, 186.
 Hénon, M.: 1960, *Ann. Astrophys.* **23**, 668.
 Hénon, M.: 1966, *Compt. Rend. Acad. Sci. Paris* **262**, 666.
 Hénon, M.: 1971, *Astrophys. Space Sci.* **13**, 284.
 van Albada, T. S.: 1968, *Bull. Astron. Inst. Neth.* **19**, 479.
 von Hoerner, S.: 1958, *Z. Astrophys.* **44**, 221.