

NUMERICAL EXPERIMENTS IN COLLISIONLESS SYSTEMS

R. H. MILLER

University of Chicago, Dept. of Astronomy and Astrophysics

Gravitational n -body calculations handling 10^5 particles have become routine in the last three or four years. The plasma physicists have also run calculations with comparable numbers of particles – you will hear of these from Dawson this afternoon. In the stellar dynamics area, two groups have been making these calculations – our own group with Kevin Prendergast and William Quirk, and the competition, the group that formed when Roger Hockney and Frank Hohl merged (and later parted ways). Our calculation is described in Miller and Prendergast (1968); theirs is best described by Hohl and Hockney (1969), and some very thorough discussions of computational speed of force calculations by Hockney (1970).

The interpretation of these calculations is beset by the same kinds of difficulties that trouble any numerical experiments – the astronomical value results from the questions asked and the kinds of experiments performed. And, all too often, the experimenter's prior prejudices come through unaffected by any experimental evidence. The astronomical interpretations of the results have been published elsewhere (Hockney and Hohl, 1969; Miller *et al.*, 1970; Quirk, 1970).

In this report, I want to take a different tack: this is supposed to be a meeting of experts in n -body calculations, so I want to concentrate on difficulties with the calculations. In particular, the emphasis will be on attempts to convince you that the bits running around inside those nice, big computers bear some relationship to the physics of stellar systems. It is not *ipso facto* evident that they do: the mere fact that the experimenter intends his calculation to relate to some kind of system in the sky does not assure any similarity.

These calculations deal with particles moving on a plane under $1/r^2$ forces. The ability to handle large numbers of particles comes from the fact that forces are computed only at a restricted set of points – if needed, forces are obtained for intermediate locations by interpolation rules. This avoids the need to compute forces between particle pairs, and makes the amount of computation necessary to obtain the forces independent of the number of particles. Of course, the main computational difficulty in gravitational n -body problems comes from the need to handle close encounters – or the possible formation of binaries. In these calculations, the forces are cut off for close encounters, thus sidestepping that problem. The justification ultimately lies in the observation that we know how to correct stellar dynamical calculations for the divergences at close encounters, but the long-range effects are the feature that defies theoretical treatment.

There is nothing magic about two dimensions. Three dimensional calculations have been made (Miller and Alton, 1968). Storage requirements and the amount of computation increase explosively in going from two to three dimensions, and the problem

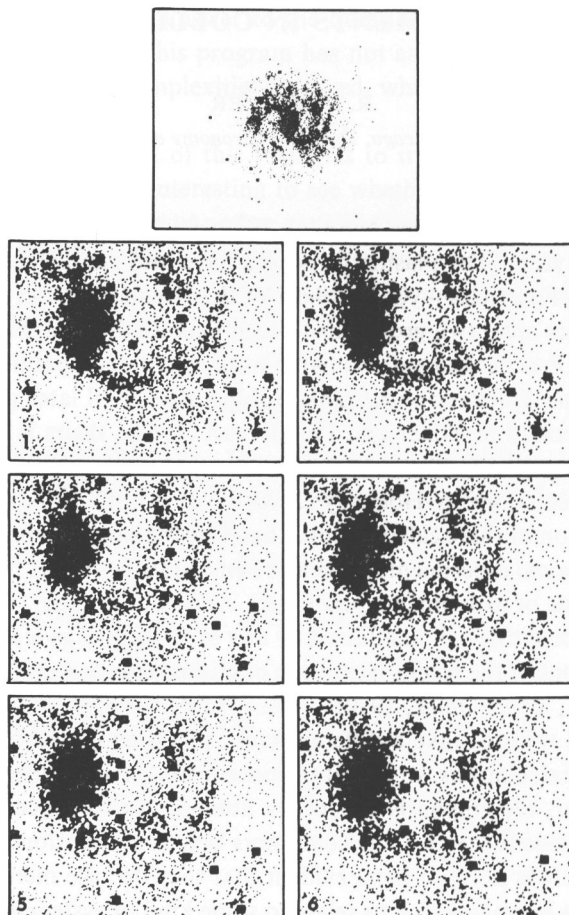


Fig. 1. Particles moving through spiral feature, from Miller *et al.* (1970). Complete pattern at beginning is shown. Time between frames is about 1/150 pattern rotation time.

of displaying the results in a comprehensible manner is much more difficult in three dimensions. There are so many challenging problems in two dimensions that there is little urge to go on to three. We just haven't gotten to three dimensions yet.

The early models used periodic configuration spaces, but both groups now have gotten rid of the periodic replications. A rather coarse grid for computing the forces remains.

The astronomically interesting problems so far attacked using these computer models are (1) the 'Jeans instability' – gravitational collapse of an initially uniform plane system, (2) attempts at static self-consistent models, and (3) spiral patterns. If you concentrate on the spiral patterns, you are led to the other two, and the spirals are most interesting, so we'll follow that route today.

The spiral patterns obtained do some of the things that you think spirals should do.

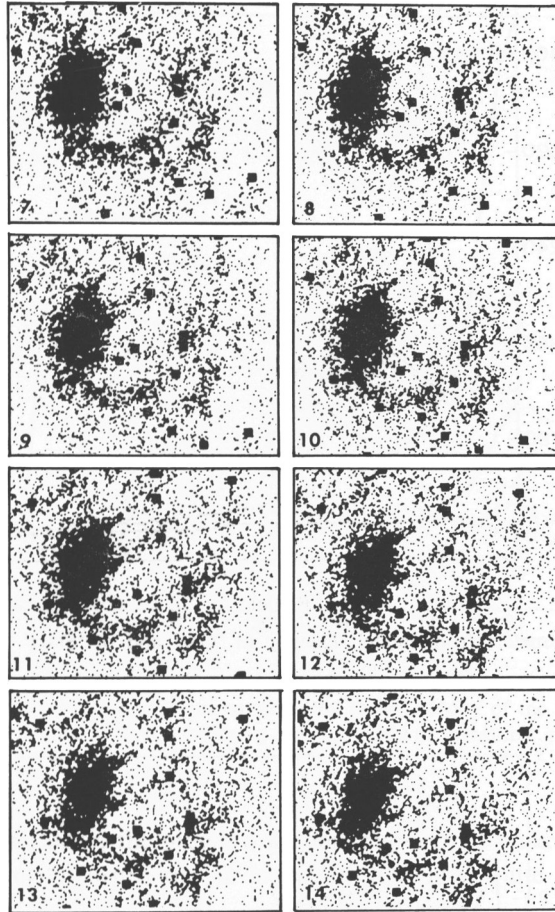


Fig. 2. Continuation of sequence of Figure 1.

For example, individual mass points ('gas' particles) can be seen to move through the spiral features (Figures 1 and 2). The usual galactic parameters can be computed – the density of 'stars' and of 'gas' separately (Figure 3), the epicyclic frequencies and the angular velocity of the pattern (Figure 4), and the mean circular and radial velocity components (Figure 5). So far everything looks very nice and reasonable – but a closer look at Figure 5 shows that the 'stars' have half the local circular velocity. Aside from complicating the definitions of epicyclic frequencies, forcing care to use gradients of the forces rather than 'Oort constants' determined from the mean velocities, this means that an equilibrium model must be pressure-supported, or must have very large velocity dispersions. This is the first clear sign of trouble. While some real galaxies have significant differences between gas velocities and star velocities (Code, 1967), our own galaxy, for example, shows very little difference between velocities determined from 21-cm observations and those determined optically from the stellar motions.

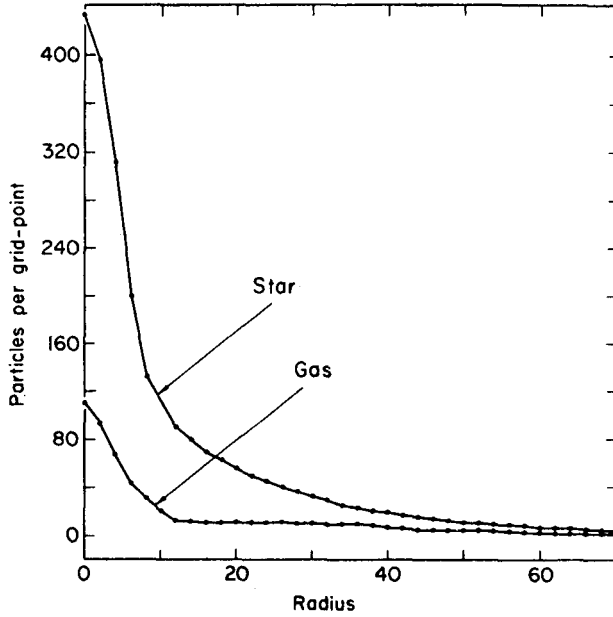


Fig. 3. Galactic parameters from model. Density of 'Gas' and of 'Stars'. From Quirk (1970).

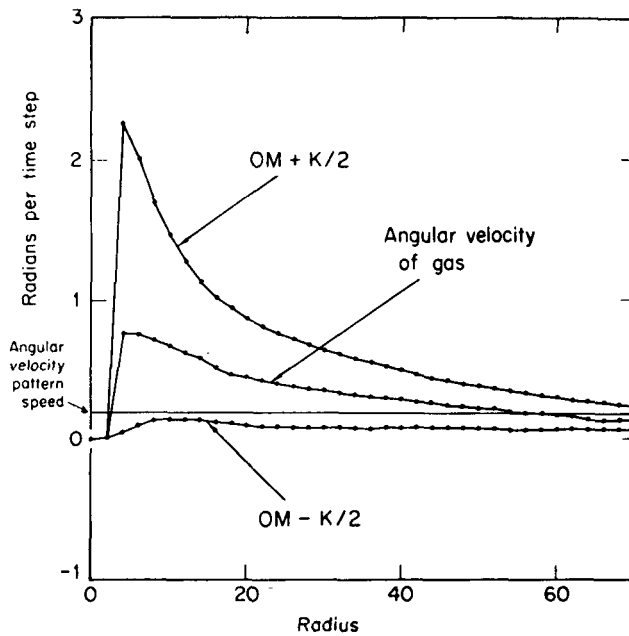


Fig. 4. Galactic parameters from model. Epicyclic frequencies and the angular velocity of gas. The pattern angular velocity is shown. From Quirk (1970). It is conceivable that the feature setting barred spirals apart from normal spirals might be the absence of an inner Lindblad resonance, as in this figure.

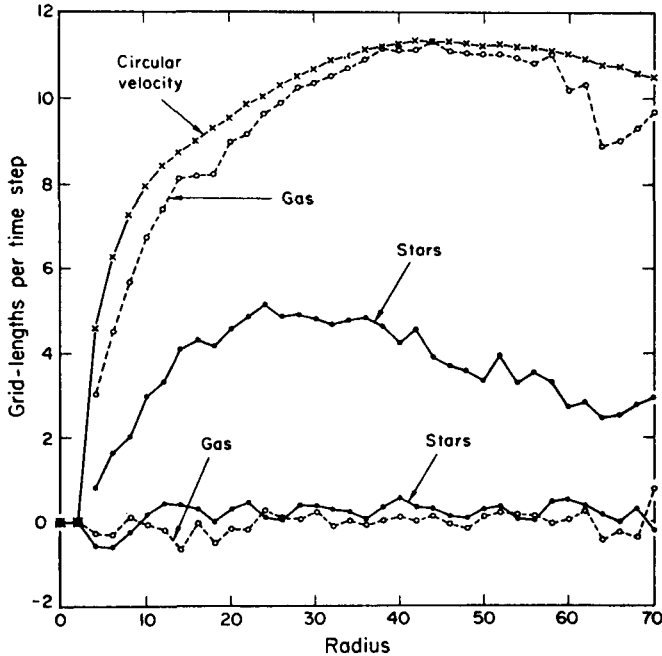


Fig. 5. Galactic parameters from model. Local circular velocity and local standards of rest for 'gas' and for 'stars'. From Quirk (1970). The circumferential mean velocity of the 'stars' is only half the local circular velocity, implying predominance of pressure-support.

Scaled to our own galaxy, the pressure support required to make up this difference would require velocity dispersions in excess of 100 km/s, quite different from the 30–40 observed.

Both groups have obtained spiral patterns. Neither has obtained spirals in a pure stellar case, and the artifices used are interesting in themselves. Hockney and Hohl have placed their system in a background potential – essentially a Schmidt model. They then get spirals that live as long as the dollars last to explore them. Our group used a two-component system in which the component that looked like 'gas' showed a spiral pattern. With both groups, the reason that spirals could not be obtained in a pure stellar case was that the pure stellar cases got too 'hot' – the velocity dispersions or pressure became very large. Figure 6 shows a curve of mean velocities determined for a purely stellar system back in the days when we were first trying to produce spirals in stellar systems by cooling them to make them move with the circular velocity. Figure 6 gave the condition just before that kind of 'cooling' was attempted. Again, the local standard of rest has half the local circular velocity. This condition of 'hot' populations is present in Hohl's calculations, too. Figure 7, which appeared in an unpublished report Hohl was kind enough to send, shows Q , the ratio of the actual velocity dispersion in the computer model to Toomre's (1964) velocity dispersion required to stabilize the system against axisymmetric modes. Figure 7 is one of Hohl's

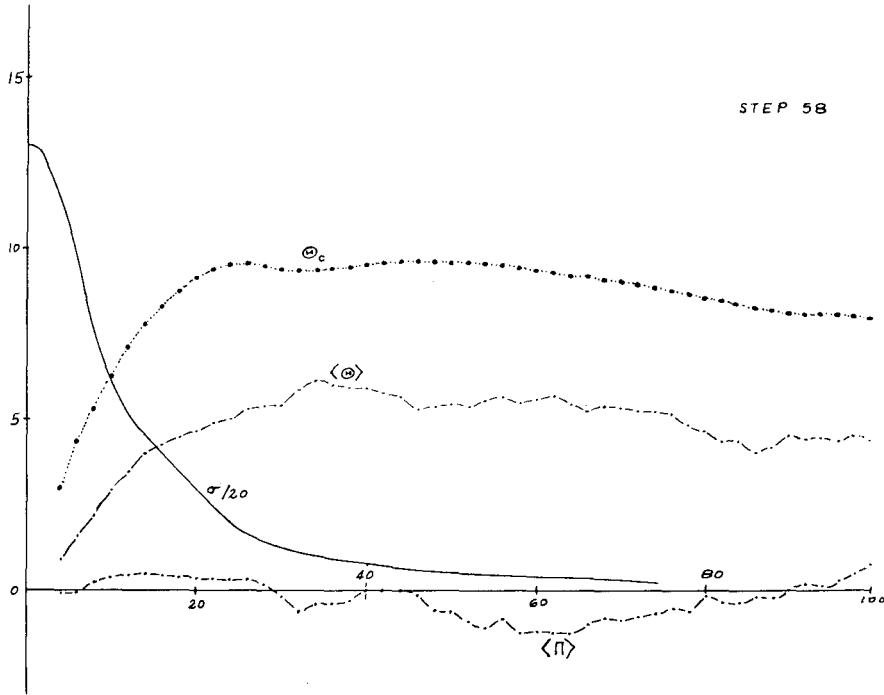


Fig. 6. Galactic parameters from a different model. 'Star' velocities compared to local circular velocity from a pure 'star' case run some time ago.

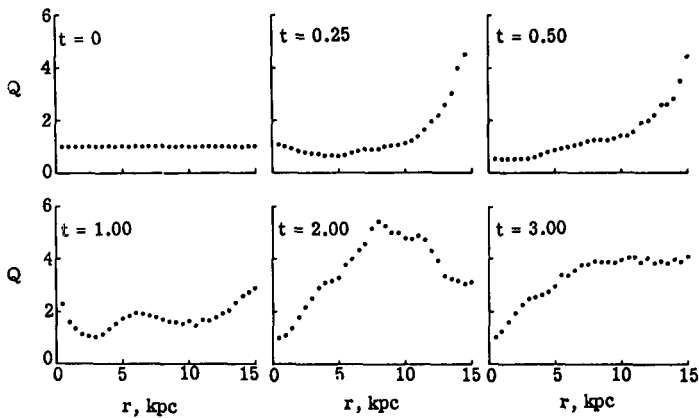


Fig. 7. Q , the ratio of the actual velocity dispersion in the model to that required for stability by the Toomre (1964) criterion. Hohl's data, courtesy of Frank Hohl.

more extreme cases, but it shows that these models can get very 'hot'. Toomre (1964) calculated that the velocity dispersions observed in our galaxy are, if anything, just a little bit too small for comfort under the stability criterion.

It is tempting to speculate that we may have constructed models that were a little bit too 'cool' and underwent a gravitational collapse. But is the price a stellar system

must pay for having violated the stability criterion (somewhere, not even everywhere) a subsequent heating to Q of a least 2, and more likely 3? If so, it is difficult to see how any galaxy could be as cool as our own.

Here, then, is the major problem that I want to stress. Why are these systems so hot? No one has yet succeeded in building a 'cool' stellar model. And no one has seen a reasonable looking spiral in a 'hot' population. We have gotten around this by introducing a second population that is strongly 'cooled' all the way through the calculation; a spiral pattern appeared in one of these experiments. Hohl has placed a stellar system in a background potential – a Schmidt (1965) model of the galaxy – and then finds that spirals can form and persist. But even then, Hohl finds Q to be $1\frac{1}{2}$ to 2 or more – so even these models are 'hot'. Hohl is now undertaking experiments at 'weak cooling' in which the velocity dispersion is very gently reduced taking care not to get too close to the Toomre stability criterion.

Hohl's experiments using a fixed background potential, and some of our own along the same lines, leave the impression that the systems do not build up this high velocity dispersion merely to have a way of providing a potential field that will not collapse under its own weight. Toomre's models (1963, 1964) can be stabilized without such a large velocity dispersion, and should provide a noncollapsing potential. But are nonaxisymmetric models harder to stabilize? Ng's (1967) axisymmetric models are just as hot, but they were constructed in such a way that it was practically assured at the outset that they would turn out to be hot.

The 'gas' population seems to do more than merely to provide a population with small velocity dispersion that could feel irregularities in the potential to show up a spiral pattern. There is a cooperation between the two populations to make the spiral pattern (Miller *et al.*, 1970; Quirk, 1970). So the dissipative character may be necessary; but Hohl's model in a background is not dissipative.

This rather aimless discussion is included to point out the difficulty of really coming to grips with the essential problem: Why are the stars so hot? Almost any discussion of these models comes back to that fundamental question. Sometimes the point comes up in discussion that we shouldn't worry about needing a two-population system of 'gas' and 'stars' in order to build a model that shows a spiral pattern, since real galaxies show their spiral patterns in gas and in the hot young stars that have been created out of the gas but cannot live long enough to leave the neighborhood of the gas. That is fine for rationalizing a need for two populations; but real galaxies do not have stellar populations that are anywhere nearly this hot. So, even with two populations, we don't have a good model of a real galaxy. With one of the populations replaced by an 'equivalent gravitational potential', the remaining population is uncomfortably hot.

It is not clear where the difficulty lies – but as experimentalists, we must adopt the viewpoint that it most likely lies in our experiments. Unfortunately, computer experiments cannot be used to prove that systems must reach this 'hot' condition – the result need not follow from all initial conditions. If the argument is made probabilistically, that we have somehow sampled the parameter space of possible initial states, a

careful look will convince you that very few initial conditions have actually been tried. And even then, the search stopped once spirals were found. It was much more exciting to explore the properties of those spirals. The only way to prove that 'hot' systems are not necessarily the only final state for a computer model is to build a 'cool' model. But no one has yet built a 'cool' static self-consistent model.

The answer may lie in studies of the stability of purely 'stellar' systems. There are, of course, some features of the computer models that make the experiments seem strained or artificial. In a sense, neither the usual theoretical model nor the computer model properly represents a real stellar system. Each is an approximation that overlooks certain essential features. The approximations are complementary. The theoretical (continuum) models overlook the grainy structure of a stellar system, while the computer models are much grainier than actual stellar systems. It appears that a 'grainy' fluctuation in the force field tends to destabilize a stellar system, so continuum models may underestimate the difficulty of achieving stable systems, while computer models may exaggerate it. Computer models have other features that stellar systems lack. The discrete number representation in a computer makes it impossible to define an infinitesimal perturbation – any change is finite (although it may be quite small). This restricts the methods available for studying stability. Distinctions (such as between instability and failure to be static-self-consistent), which are clearly defined for the continuum theoretical models may not be meaningful in computer experiments. The distinction might be meaningful for a stellar system and still be difficult or impossible to apply operationally because an instability cannot be recognized.

The importance of the question, quite aside from a natural desire to make our models be as close to nature as possible, lies in the interpretations that are otherwise made of things like Toomre's stability criterion and the observed velocity dispersions. They figure heavily in the arguments of Lin *et al.* (1969), for example, in discussions of whether spiral galaxies that we observe are stable configurations.

Some attempts to construct static self-consistent models will be taken up next, then a discussion of some numerical properties of the Miller-Prendergast (1968) model as a preliminary to a different way of undertaking the construction of self-consistent models. Along the way, the advertised features of our model as providing an exact handling of the collision-free Boltzmann, or Vlasov, equation will be pointed out.

1. Self-Consistent Model

The models that we have used as initial conditions for various calculations have not been designed to be static self-consistent models. Invariably, the models have evolved through spectacular collapses into systems that were largely pressure-supported rather than being 'cool' in the sense of having small velocity dispersions.

The 'hot' condition does not seem to be necessary for disk galaxies, but the velocity dispersions of our evolved computer models are substantially greater (perhaps twice or more) than required for stability, and are disconcertingly large for comparison with real stellar systems. A static self-consistent model that can be maintained in the com-

puter with smaller velocity dispersions is essential for studying actual galaxies, particularly if the galaxies are to be built without resorting to several populations.

This work was undertaken to see whether a ‘cool’ static self-consistent model could be maintained in the computer. A successful attempt would show that there is nothing intrinsic to plane stellar systems that requires unusually large velocity dispersions in order to remain stable. The model was tested for its stability limits and to see what kinds of instabilities might develop. A complication in such a study is that the onset and early growth of an instability cannot be readily detected. Rather, the instability must grow to finite amplitude to be detected – unless we can devise a more sensitive indicator. A genuine instability is difficult to distinguish from mere failure to construct a truly static self-consistent model. Three-dimensional modes, such as bending, are not allowed.

A. THE MODEL

The model chosen was essentially ‘Model I’ of Toomre (1963). All of the standard galactic parameters can be easily worked out for this model, and the stability criterion of Toomre (1964) can be applied as well. This model has, in obvious notation:

$$\rho(\varpi, z) = \mu(\varpi) \delta(z) = \frac{Mb}{2\pi} (\varpi^2 + b^2)^{-3/2} \delta(z), \tag{1}$$

for which the corresponding potential is

$$V(\varpi, z) = -GM [\varpi^2 + (b + |z|)^2]^{-1/2}. \tag{2}$$

Here b is a scale parameter that sets the dimensions of the system. If all particles have unit mass, M is just the number of particles.

In the initial conditions generated for a machine calculation, a number of particles was chosen (125000), and a scale parameter b . With $b=13$, the number of particles per location drops below $\frac{1}{2}$ at a radius $\varpi=80$. Velocities were chosen for the particles at each configuration space location to make a Gaussian (Schwarzschild) velocity distribution whose mean is the local circular velocity and with a disperion whose axis ratios conformed to the usual ones computed from the Oort constants (see e.g. Chandrasekhar, 1960, Sec. 4.3) with the actual root-mean-square velocities being a multiple, T , of a convenient form that is large at the center and decreases outward. The actual velocity of each particle was generated by a pseudorandom number generator, to meet these conditions.

The velocity dispersion that results with Toomre’s stability criterion is large enough to require allowance for pressure-support. This is done by reducing the local circular velocity according to the usual hydrodynamical equations of stellar dynamics (see, e.g. Chandrasekhar, 1960. Sec. 4.8 (iv)). It can happen (at T about 1.5 that of the Toomre condition) that the system becomes wholly pressure supported in the center – there is then no rotation in those parts of the system that are pressure-supported.

B. MODIFICATION TO THE ACTUAL FORCE LAW

The force between pairs of particles, used in the calculation, is derivable from a potential

$$\varphi(x, y) = \text{const.} [(x - x')^2 + (y - y')^2 + a^2]^{-1/2}. \quad (3)$$

This modification was used to avoid computational troubles from near encounters. It is convenient to think of this potential in a 3-dimensional space. The particles move in a plane at $z = \text{constant}$, and the force on a test particle is measured a units above that plane, at $z = a$. Only those force components that lie in the plane $z = a$ enter the problem – the test particle is not permitted to leave the plane. For most calculations $a = 3$. This change, which seems small, produces a surprisingly large change in the actual force field near the center of a configuration like that of Equation (1). Corrected expressions for the parameters involving the force (rotational velocity, Oort constants, epicyclic frequency, and so on) can be obtained from the usual expressions by substituting $c = b + a$ in place of b wherever it appears. Thus all the model parameters can be exactly corrected for the revised force law.

A remarkable feature of this force law is that the same notion can be used to make the appropriate modification to the stability calculation. Toomre has also noticed that this force law permits the stability calculation to be carried through. The result is interesting by itself because it separates long-range from short-range contributions to the predicted stability conditions. Toomre's calculation proceeds from a linearized collision-free Boltzmann equation (Equation (43) of Toomre, 1964) in which the only term affected by the modified force law (assuming that the model representing the unperturbed system is properly self-consistent with the revised force-law) is that containing the force due to the perturbed density distribution. The force for each Fourier mode is obtainable from the usual solution to the Poisson equation,

$$V'(x, z) = \text{const.} \frac{2\pi G}{k} e^{ikx} e^{-k|z|}, \quad (4)$$

which indicates that the replacement $G \rightarrow Ge^{-ka}$ wherever G appears will properly account for the revised force law. Toomre's calculation now follows, yielding a modified dispersion relation

$$\frac{\kappa^2 a}{G\mu} = 2\pi\beta \frac{1 - e^{-y} I_0(y)}{\sqrt{y}} e^{-\beta y}, \quad (5)$$

where

$$y = \frac{\sigma_u^2 k_n^2}{\kappa^2}, \quad \beta = \frac{\kappa a}{\sigma_u}, \quad (6)$$

k_n is the wavenumber of the neutrally stable mode, σ_u is the velocity dispersion in the radial direction, and κ is the epicyclic frequency. The system is stable if $\kappa^2 a / G\mu$ is greater than the righthand side of Equation (5). Let $K(\beta)$ be the maximum value

attained by the righthand side of Equation (5) for any value of $\gamma(\geq 0)$ and the given value of β . Since the lefthand side contains (local) quantities that are completely determined by the model, this relation can be solved for β anywhere in the system. An interesting feature is that $K(\beta)$ attains an asymptotic value $2\pi/e=2.31145\dots$ as $\beta \rightarrow \infty$ and any value of $\kappa^2 a/G\mu$ greater than this can be stabilized with any value of σ_u including $\sigma_u=0$, by taking β infinite. Thus, with this modification to the force law (near cutoff in the forces), there are situations in which rotation alone is sufficient to stabilize the system – a result that appears quite different from Toomre's. (The distinction from the results of Lynden-Bell (1962) and Lee (1967) concerning rotational stabilization should be noted: these both refer to three-dimensional mass-distributions.)

In the model designed for the machine calculation, with $(GM)^{1/2}=3.0155 c$ (a value that sets the maximum force to 3.5), the rotation should have been sufficient to stabilize the system inside $\varpi \sim 26$. The actual velocity dispersion (in the radial direction) that is required for stability is shown, in units appropriate to the model, in Figure 8.

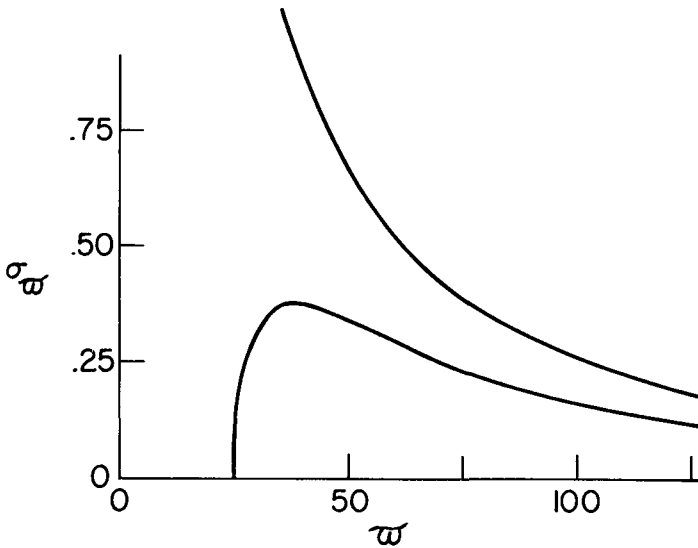


Fig. 8. Velocity dispersion in the radial direction required to stabilize configuration against axisymmetric modes. The lower curve is for the modified force law, according to Equation (5); the upper curve is for the same configuration with $1/r^2$ forces. The upper curve would reach 2.91 at the center.

With reasonable values for the velocity dispersions, the greatest instability should occur around $\varpi=70$ in these models – fairly far out where the density is already quite low. Furthermore, the velocity dispersions required to stabilize are reduced by at least a factor of two below those expected without allowing for the near cutoff of the forces.

C. EXPERIMENTAL RESULTS

In the machine representation, all physical quantities are integers. To avoid sharp discontinuities in the initial density and velocity fields, we rounded by adding a

pseudorandom number, uniform in $(0, 1)$ to the 'real' value before truncating. The rounding is unbiased, but has twice the variance of the usual rounding rule. The variance acts like a velocity dispersion (with a circular velocity ellipse), and can contribute to the stabilization of the system through random motions. In fact, this term alone should be sufficient to stabilize the systems with the modified force law.

Roundoff 'noise' in the force-values is expected to be destabilizing, in contrast to the stabilizing influence of velocity dispersions. This can be seen by multiplying the linearized Boltzmann equations appropriate to different phase space points, and assuming that cross terms correlating forces to velocities, velocities to positions, and forces to coordinates can be neglected. The term with force covariances enters with the opposite sign to those with velocity variances. In physical terms, fluctuations in the force seem directly to induce fluctuations in the part of the distribution function that depends upon the velocities, broadens the velocity dependence. The effect of a velocity spread behaves differently, since it is already included in the distribution function. Quantitative estimates of the effect are complicated by correlations in the 'fluctuations' of force-values at different spatial locations. The force rounding remained with the 'add one-half' rule, to preserve the reversibility that we feel is an important feature of our model.

Two series of experiments were run with these configurations. The experiments started from an initial load produced by the methods just described, then followed the evolution of the system for a few integration steps. Normally, ten integration steps were sufficient to tell how the system was going to behave.

Each series of experiments was a sequence of runs with different values of T , the parameter that controls the velocity dispersion in the initial load. The two series of experiments were run with different numbers of configuration space locations per periodic cell length. The number of lattice points per periodic length was increased by the square of the factor by which the timestep is reduced, so force-values were unaltered.

The results of each experiment were observed in step-by-step density plots of the configuration and through values of the force tabulated for certain control points in the system. Some runs, duplicated with different sequences of pseudorandom numbers used in the initial loading routine, gave results that were qualitatively similar. The extent of quantitative difference indicates the reliability of the numerical results. The similarities of two systems starting with different pseudorandom number sequences were comparable with the retention of gross symmetry within a given system. Thus, these results are not peculiar to the particular set of pseudorandom numbers used.

The experimental results may be qualitatively described as follows. First, the systems run with our usual 256×256 grid appeared not to be static self-consistent models. They tended to rearrange themselves rather quickly, although those with larger velocity dispersions did so less rapidly. This probably means that discrete models are not adequately approximated by models designed according to continuous density distributions, although the departure is less troublesome for 'hotter' systems. There were no instabilities that were recognizable as such.

Second, the experiments with ‘reduced time-step’ (more points per configuration space lattice), were more nearly static. Runs were made with $\frac{1}{3}$ to $\frac{1}{3T}$ of the normal time-step, with T about 1.3 times the value required for stabilization in the absence of ‘quantization noise’. All these runs behaved about the same way, so $\frac{1}{3}$ was used thereafter. A run with $T=0$ showed a clumping similar to that seen earlier with the plane gravitational instability – very irregular patterns leading to large density variations characterized by fairly short length scales. These seem to be genuine instabilities, as distinguished from mere failure to have achieved a static self-consistent model, although the distinction is not clear-cut. The instability was strongest near the center, where the system should have been rotationally stabilized.

A sequence of runs was made with $T=0, 0.1, 0.2, 0.3, 0.4$ and 0.5 . Stabilization should be achieved (without ‘quantization noise’) at $T=0.267$ for the form of velocity dispersion used. The velocity dispersion was largest at the center and fell off uniformly toward the outside; the precise analytic form is not important. The system becomes completely pressure-supported near the center at $T=0.457$. In general, larger values of T lead to systems that are more nearly static, self-consistent, and stable, as expected. At $T=0.1$, an instability is still present but it grows less rapidly. Values of T greater than 0.2 do not produce recognizable instabilities, although the systems still rearrange themselves in search of an equilibrium configuration. The rearrangement is slowed with larger values of T . There does not seem to be a recognizable boundary between genuine instabilities (if that is what they are) and rearrangements. There is rather a uniform trend toward slower rearrangement and longer characteristic lengths with larger values of T (larger velocity dispersions).

Hockney and Hohl (1969) have carried out similar experiments with a different model. Their low-velocity-dispersion configurations form into patterns characterized by short length scales; configurations with increased velocity dispersion change more slowly and form into patterns with larger length scales. Their treatment of the forces effectively introduces a near-cutoff that must influence the analytic stability criterion in some way like that of subsection *b*; their treatment of the forces also introduced local departures from the analytic form that are equivalent to the stochastic contribution noted above. It is difficult to distinguish an instability from a failure of self-consistency in their models, just as it is in ours. The picture that emerges from their published results seems similar.

D. DISCUSSION

These experimental results seem to disagree with expectations on two counts. (1) Something happens, even though ‘quantization noise’ resulting from the discrete allowed velocities alone should be sufficient to stabilize the system. As noted earlier, ‘quantization noise’ in the forces tends to destabilize the configuration; it works against the effects of the velocity dispersion, and could make the system unstable. There is not a clear-cut stability threshold to compare with theory. (2) That ‘something’ is strongest near the center, a place where rotation alone should be adequate for stabilization.

There are several points at which these experiments fail to meet the conditions of the stability calculations. First, the system is discrete, so the continuous model is an approximation. It is difficult to estimate the effect of the approximation and even more difficult to do the calculation for a discrete system (but see Section 3.) Second, the onset of the instability is not detected. The instability (if indeed that is what it is) may well have started where theory indicated and later have spread to the center, where we were able to observe it. Third, the configuration is no longer axisymmetric at the time that it is clear that something has happened. We cannot say whether the unstable mode(s) that started the whole process were axisymmetric; the final state that we observe is not. Fourth, the initial state is not strictly axisymmetric. Here again, we do not know how important the ‘perturbations’ of the computer model are.

Our earlier experiences in trying to produce ‘cool’ models led to the conjecture that the ‘hot’ condition resulted from a rearrangement following a gravitational instability which developed at some place that was too ‘cool’. The response to this instability could be a local ‘heating’. The fact that our self-consistent equilibrium systems were so far beyond the stability threshold is disturbing. It suggests that the penalty for violating the stability criterion might be rather severe; that it could lead to ‘heating’ far in excess of that just sufficient to lift the instability, and that this ‘heating’ could extend over the entire system even if it arose in response to a localized instability.

The experiments reported here neither confirm nor disprove this conjecture. The cause cannot easily be tied in with instabilities because it is operationally difficult to identify instabilities as such. The picture evidently can be tied to a failure to be self-consistent as well as to instabilities. Systems that started with low velocity dispersions (small values of T) rather quickly developed ‘clumps’ of ‘hot’ stars. Those with higher velocity dispersions did not change so rapidly, but did rearrange themselves, doubtless with some resultant ‘heating’. None of these models was run long enough to tell how ‘hot’ the system would have looked if it had reached an equilibrium state, but it seems likely that any one of them would have become as ‘hot’ as our other typical models. A ‘cool’ stable equilibrium was not reached, but this does not prove that none exists.

2. Numerical Considerations

The tendencies of computed collisionless systems to get hotter than the real galaxies that we see in the sky might result from the approximations necessary to represent the systems in a computer. The entire set of numerical problems is of utmost importance in n -body calculations in general. Because some of these matters, especially as applied to the large n -body calculations, have been discussed in detail elsewhere, we merely refer to these discussions (Miller, 1970).

3. A Thought-Experiment

Static self-consistent models should be straightforward to construct in the discrete

phase space if they are undertaken directly rather than by analogy to continuous models as described in Section 1. These discrete static self-consistent models are instructive, although they may not be very interesting to construct in the computer.

In the discrete phase space, a point moves over an integer lattice according to the rules of the 'game' described by Miller and Prendergast (1968). If the model were truly static and self-consistent, the force at any allowed configuration space point would be the same at one integration step as at any other.

A realistic 'galaxy' model should contain one isolated region of nonzero forces (periodicity is unimportant for this discussion) in an infinite expanse where the forces are zero. Orbits that are restricted to the region in configuration space where the forces are nonzero occupy a bounded portion of the phase space – there is a maximum velocity that a particle may have and still be in a bound orbit. Any allowed phase space point is either on a periodic orbit or on an escape orbit (it is on a unique orbit, and the number of points in the region containing bounded orbits is finite). This is the discrete analogue of Hopf's first theorem (see, e.g., Contopoulos, 1966). The static self-consistent model can be constructed of these periodic orbits.

Consider a periodic orbit containing one particle at each phase space point along the orbit. At any integration step, the particle occupying any one of these points will move to another point along the orbit, but the original point will be filled by a particle that has moved up from the previous point. The integration process looks like a huge game of 'musical chairs', with just as many 'players' as 'chairs'. After the integration step, the entire system looks exactly as it did before the step.

The problem of constructing static self-consistent models consists in: (1) Design a 'reasonable' force-field; (2) Find all the periodic orbits in that force field; (3) Find a subset of the periodic orbits that, when projected to configuration space, will produce the original force field. If there is no such subset, pick a new force field and start over. The whole process sounds trivial – but it is not at all trivial except in some particularly simple cases.

Such a system, once constructed would not be very interesting. It would sit forever and never change. As an example, suppose there were a ring of particles – the discrete analogue of a set of particles at the vertices of a regular polygon, rotating in the plane. In the discrete phase space, this could be mimicked by two-dimensional harmonic oscillator orbits, chosen to be periodic in one rotation around the origin (same periodicity in x and in y ; two-dimensional harmonic oscillator orbits tend to make Lissajous figures with periods not 1:1 in the discrete phase space). The forces produced by the ring of particles, in the neighborhood of the ring itself, will look like harmonic oscillator forces. Adjust the force constant to balance things.

This ring would, if not disturbed, circulate forever. Yet it represents a system that is known to be unstable in continuous space. This system is stable although 'cold' in the sense of having no random velocities. It is not a very interesting galaxy model. The interest attaches to the response of the system to a perturbation – but this system cannot be perturbed by arbitrarily small amounts. The least perturbation that can

be given to it is to remove one of the particles to another orbit. If this were done, the ring would probably disintegrate to a chaotic motion very quickly.

The same thought-experiment can be carried out with the kind of model attempted earlier. Suppose we had succeeded in constructing a static self-consistent model. It would circulate forever, each particle on its own periodic orbit. But then remove one particle somewhere, either leaving a hole or replacing the particle elsewhere in some other orbit – preferably periodic so that it will stay around for a while. The two anomalies will move around the phase space until they happen to reach locations at which they conspire to make the force value round differently than it would in the absence of the perturbation. More particles will be disturbed by this change. Thereafter, the system is no longer static and self-consistent; it will take off and can change drastically. The argument is exactly like the reversibility argument of Miller (1970).

Without actually doing the experiment, we cannot say whether the system that would eventually result would be ‘cool’ or ‘hot’, or whether it would be approximately axisymmetric. Some points are clear, however. First, it is simple to estimate how long, on the average, a particle or a hole would have to move around the system before it would, in a probabilistic sense, cause a computed force value to round differently. For the model tested, this is 10–15 integration steps. It is this long because the force near a single particle is much smaller in magnitude than the interval between successive allowed values of the force, and because the force field around a particle effectively extends only over a very small spatial region. The system might not ever notice that it had been perturbed. This, of course, is just what is meant, operationally, by an ‘infinitesimal’ perturbation. But the system response is nonlinear, and it is the nonlinear response that causes the trouble. A second point is that the model described in Section 2 is hard to make self-consistent because of a design error – it had particles in regions of zero force. These must be on escape orbits. This is not serious with periodically replicated systems.

This discussion makes it clear that there is a smallest perturbation that can be given to the system and a smallest perturbation to which the system can respond. Because of the discrete number representation in computers, any model in a computer shares these attributes. But with the very fine discrete phase space allowed with full-precision computer numbers, the number of particles in the self-consistent model must be very large. It is straightforward to estimate how the number of particles scales with the number of points allowed in each coordinate direction and in each velocity component as the spatial lattice and time-steps are separately refined. The present calculation is just crude enough to build somewhat realistic static self-consistent models out of the number of particles that it can handle. But we have not yet developed an algorithm that will generate such models at will. Initial steps in this direction have produced some very interesting results, showing that a phase space point may be on a periodic orbit while its neighbors are on escape orbits – some regions have small fractions of periodic orbits.

There are no nice continuity properties to the phase space, of course; there is not a closed surface inside which all orbits are periodic and outside which all orbits are of

the escape variety. Some orbits require several thousand steps to repeat periodically; some have extraordinary shapes. The bookkeeping required to keep track of even these few periodic orbits seems prohibitive.

The notion of such self-consistent models can be extended to models that move in regular ways. A pattern that translates itself across the plane without change of shape is one such possibility – patterns that appear to rotate seem more difficult.

A system containing many ‘holes’ should be able to respond to smaller perturbations. This might provide a useful route to attack the self-consistency problem posed in this report.

A truly static self-consistent model seems difficult to build. An attempt to do so is very instructive, and provides considerable insight into the way the computer models behave.

4. Conclusions

The search for reasons why the computer models might be so hot has led through some interesting investigations. At the end of all these investigations, we still don’t know whether it is possible to build a static self-consistent model that is cool. However, along the lines of the ‘thought-experiment’ of Section 3, it seems quite likely that one can be constructed. Whether it would be stable (whatever that means) to small disturbances is unclear, but the suspicion is strong that, if the system felt the disturbance at all, it might well be quite unstable. That means that its response to as small a disturbance as it could feel would be a major rearrangement that might result in its being as hot as all the other models.

A related question is why spiral patterns do not last any longer than they do. Hohl’s last quite long – for reasons that are as mysterious as why ours don’t last very long. The obvious reasons might be (1) that there aren’t enough particles to build a long-lived spiral, or (2) that the spirals dissipate because of numerical effects. After all this, there is no clear-cut basis to choose between the two alternatives, but there is no *a priori* reason why numerical effects should work against the maintenance of spirals instead of helping to maintain them.

Acknowledgements

It is a pleasure to acknowledge the assistance of many people who have participated in this program. The bulk of the calculations were carried out at the Goddard Institute for Space Studies in New York through the courtesy of Dr Robert Jastrow, Director. My colleagues on this project, Kevin Prendergast and William J. Quirk of Columbia University have been full partners in the work at all stages. Many thanks are due to Frank Hohl for generously taking time to give me rather complete descriptions of what he and Roger Hockney have been doing – I hope that the descriptions of his work in this report do him justice. Hohl has essentially gotten as far in this work as we have, a fact that is obscured by the heavy emphasis on our work in this report.

References

- Chandrasekhar, S.: 1960, *Principles of Stellar Dynamics*, Dover Publications, Inc., New York.
- Code, A. D.: 1967, *Astron. J.* **72**, 789–90 (A).
- Contopoulos, G.: 1966, 'Problems of Stellar Dynamics', pp. 169–257 in *Relativity Theory and Astrophysics* (J. Ehlers, ed.) Vol.1 American Mathematical Society, Providence, R.I.
- Hockney, R. W.: 1970, *Methods in Computational Physics*, Vol. 9, pp. 136–212, (B. Adler, S. Fernbach, and M. Rotenberg, eds.), Academic Press New York.
- Hockney, R. W. and Hohl, F.: 1969, *Astron. J.* **74**, 1102.
- Hohl, F. and Hockney, R. W.: 1969, *J. Comp. Phys.* **4**, 306.
- Lee, E. P.: 1967, *Astrophys. J.* **148**, 185.
- Lin, C. C., Yuan, C., and Shu, F. H.: 1969, *Astrophys. J.* **155**, 721.
- Lynden-Bell, D.: 1962, *Monthly Notices Roy. Astron. Soc.* **144**, 279.
- Miller, R. H.: 1970, *J. Comp. Phys.* **6**, 449.
- Miller, R. H. and Alton, N.: 1968, 'Three Dimensional N-Body Calculation', ICR Quarterly Report, No. 18 (August 1, 1968), Institute for Computer Research, University of Chicago (Unpublished).
- Miller, R. H. and Prendergast, K. H.: 1968, *Astrophys. J.* **151**, 699–709.
- Miller, R. H., Prendergast, K. H., and Quirk, W. J.: 1970, *Astrophys. J.* (to be published).
- Ng, E.: 1967, *Astrophys. J.* **150**, 787.
- Quirk, William J.: 1970, Ph.D. Thesis, Columbia University, (Unpublished).
- Schmidt, M.: 1965, *Galactic Structure*, pp. 513–30. (A. Blaauw and M. Schmidt, eds.) University of Chicago Press, Chicago, Vol. V, of Stars and Stellar Systems, (B. M. Middlehurst and G. Kuiper, eds.).
- Toomre, A.: 1963, *Astrophys. J.* **138**, 385.
- Toomre, A.: 1964, *Astrophys. J.* **139**, 1217.