

STATISTICAL MECHANICS OF VIOLENT RELAXATION IN STELLAR SYSTEMS

D. Lynden-Bell

(Communicated by the Astronomer Royal)

(Received 1966 December 19)

Summary

An explanation of the observed light distributions of elliptical galaxies is sought and found.

The violently changing gravitational field of a newly formed galaxy is effective in changing the statistics of stellar orbits. The equilibrium distribution under this encounterless relaxation is found by use of a fourth type of statistics related to both Fermi–Dirac statistics and equipartition of energy per unit mass. In the relevant limit this becomes Maxwell’s distribution but with temperature proportional to mass.

The predicted light distributions are those of the modified isothermal spheres developed by Michie from considerations of stellar relaxation in globular clusters. Both these and the special case further developed by King are known to give agreement with observations of spherical systems. Application to clusters of galaxies will remove Zwicky’s paradox.

The theory is also developed for rotating systems where allowance must be made for anisotropy of stellar motions if the outer parts are not to be much flatter than the inner parts.

The new statistics developed here should have important applications to collisionless plasmas and collisionless shocks.

Kelvin’s theorem is rederived for collisionless dynamics.

It is suggested that the typical ‘equilibrium’ state of a stellar system may be hierarchical.

1. *Introduction.* The remarkable regularity in the light distribution in elliptical galaxies suggests that they have reached some form of natural equilibrium. However, estimates of the normal star–star relaxation show that it is too weak to establish equilibrium in the time available. Equipartition of energy would lead to a marked segregation by mass with the lighter stars at the outside; and, as a result, to greater colour differences than are observed. No relaxation mechanism which leads to equipartition of energy can be primarily responsible.

This paper discusses the relaxation that occurs when the mean gravitational field of the system is not steady and derives the form of equilibrium towards which this relaxation proceeds. The importance of this form of relaxation has previously been stressed by a number of authors including Henon (1) and King (2). Numerical experiments on it were recently made by Henon (1) and Lecar (3).

Both recently formed and tidally deformed stellar systems possess a large scale gravitational field which changes in time. Due to these changes the stars follow complicated paths along which the individual stellar energies are not conserved. In fact

$$\frac{d\epsilon^*}{dt} = -m \frac{\partial\psi}{\partial t} \quad (1)$$

Where ϵ^* is the energy of a star, m is its mass and $\psi(x, y, z, t)$ is the gravitational potential of the whole stellar system measured from a zero at infinity.

$$\epsilon^* = m(\frac{1}{2}c^2 - \psi)$$

where c is the star's velocity; it is natural to define an energy per unit mass $\epsilon = \epsilon^*/m$ so that equation (1) becomes

$$\frac{d\epsilon}{dt} = -\frac{\partial\psi}{\partial t}. \quad (2)$$

The relaxation time, T_r , may be defined as

$$T_r = \left\langle \frac{\left(\frac{d\epsilon}{dt}\right)^2}{\epsilon^2} \right\rangle^{-1/2} = \left\langle \frac{\left(\frac{\partial\psi}{\partial t}\right)^2}{\epsilon^2} \right\rangle^{-1/2} \quad (3)$$

To find an estimate for T_r we must know how rapidly ψ changes. The galaxy will vibrate turning potential energy into kinetic energy and back again in accordance with the time dependent virial theorem,

$$\frac{1}{2}\dot{I} = 2T + V.$$

Here

$$I = \sum mr^2$$

where r is the position of the mass m with respect to the centre of mass of the whole galaxy and the summation is extended over all the masses.

T is the kinetic energy of the galaxy with respect to its centre of mass; and V is the potential energy of the galaxy. We also define the total energy of the galaxy $E = T + V$.

At equilibrium $\dot{I} = 0$ so $T = -E$, $V = 2E$.

Away from equilibrium T and V will vibrate about these values since E is constant. T is the sum of the kinetic energies of the individual stars but V is half the sum of their potential energies (because it is mutual). For a typical star the kinetic energy must therefore vibrate about one quarter of its potential energy so

$$\frac{1}{2}mc^2 \sim \frac{1}{4}m\psi$$

whence

$$\epsilon \sim -\frac{3}{4}\psi.$$

For the relaxation time we have from (3) approximately

$$T_r = \frac{3}{4} \left\langle \frac{\dot{\psi}^2}{\psi^2} \right\rangle^{-1/2} \quad (4)$$

T_r is thus closely related to the time in which $\log \psi$ changes which is typically the same time scale as for the vibration of the whole galaxy. To get the factors 2π etc. approximately correct we now derive this from the virial theorem.

Define R by

$$-\frac{GM^2}{R} = V \quad (5)$$

where M is the mass of the galaxy.

Then $I = \lambda^2 MR^2$ (6)

where λ^2 is a number of order unity which is approximately constant for the

fundamental mode of vibration. λ^2 is $\frac{1}{3}$ for a body whose density falls like r^{-2} within some boundary and is approximately $\frac{1}{3}$ for a uniform body. From the Virial theorem

$$\lambda^2 \ddot{R}^2 = 2E/M + GM/R$$

To find the period we assume small amplitude vibrations about the equilibrium radius $R_0 = GM^2/(-2E)$.

Then

$$2\lambda^2 R_0 \delta \dot{R} = -\frac{GM}{R_0^2} \delta R \quad (7)$$

which is simple harmonic with angular frequency

$$n = \left(\frac{GM}{2\lambda^2 R_0^3} \right)^{1/2} = \frac{2}{\lambda GM} \left(\frac{-E}{M} \right)^{3/2} \quad (8)$$

or putting $\lambda^2 = \frac{1}{3}$ and $\bar{\rho} = M/(\frac{4}{3}\pi R_0^3)$

$$n = (2\pi G \bar{\rho})^{1/2}. \quad (9)$$

We now assume that initially the variation of R is as large as the mean, R_0 itself. Then for all positions inside the main body of the galaxy the changes in ψ will be of the same order as ψ itself. So from equation (4) we obtain

$$T_r \simeq \frac{3}{4n} = \frac{3P^*}{8\pi} \quad (10)$$

where $P^* = 2\pi/n$ is the typical radial period of the orbit of a star in the galaxy.

Formula (10) illustrates the violence of this form of relaxation. Throughout the above discussion the mass of the star cancelled out showing that gain or loss of energy per unit mass by any star is not dependent on its mass. We may predict at this early stage that this form of relaxation will not lead to any segregation by mass.

The vibrations of a galaxy are heavily damped by Landau Damping so they will not persist for more than a few periods. Basically this is because even if all the stars start falling inwards (i.e. all in step) their different galactic periods will soon spoil the synchronism (4). However, from our relaxation estimate it seems quite possible that this is long enough for the galaxy to make significant progress towards a state of complete 'mean field relaxation'. We therefore turn our attention to the question 'where is this form of relaxation leading?' In particular, does it lead to some form of statistical equilibrium?

2. *Invariants.* Any final equilibrium that is attained must have the same total energy, E , as the initial state. We could say the same of the total mass M , the total angular momentum \mathbf{H} and the total linear momentum \mathbf{P} . Conserving these quantities only, statistical mechanics leads us to Maxwell's distribution in moving and rotating axes. However, on the time scale which we are now considering star-star encounters are quite negligible and the system is described by the time-dependent Boltzmann-Liouville equation with no encounter term. We shall show that there are many more invariants than the above mentioned six. If $f d^3r d^3c$ is the total mass of those stars in a volume d^3r about \mathbf{r} flowing with velocities in the range d^3c about \mathbf{c} the Boltzmann-Liouville equation reads

$$\frac{Df}{Dt} \equiv \frac{\partial f}{\partial t} + \mathbf{c} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{\partial \psi}{\partial \mathbf{r}} \cdot \frac{\partial f}{\partial \mathbf{c}} = 0 \quad (11)$$

where

$$\begin{aligned} \mathbf{c} &= (u, v, w), \quad d^3c = du \, dv \, dw \\ \mathbf{r} &= (x, y, z), \quad d^3r = dx \, dy \, dz \\ \frac{\partial}{\partial \mathbf{c}} &= \left(\frac{\partial}{\partial u}, \frac{\partial}{\partial v}, \frac{\partial}{\partial w} \right) \end{aligned}$$

and D/Dt is the convective derivative following the motion in phase space. ψ , the gravitational potential, arises from the mass density $\int f d^3c$ by way of Poisson's integral

$$\psi(\mathbf{r}, t) = G \int \frac{\int f(\mathbf{r}', \mathbf{c}', t) d^3c'}{|\mathbf{r}' - \mathbf{r}'|} d^3r'. \quad (12)$$

Since $Df/Dt = 0$ each element of density in phase space conserves its phase-density as it moves (Liouville's theorem but here in six dimensions rather than $6N$). Each element of phase is always made up of the same stars wherever it moves to so it always has the same mass. Hence $m(f) \delta f$, the total mass of all those phase elements with phase-densities between f and $f + \delta f$, is conserved. To put it differently $M(f)$ the total mass of all those elements with phase-density greater than f is conserved for each and every f .

We have here a conserved function, an infinity of conserved quantities. It is natural to attempt the new problem in statistical mechanics which allows for this conservation. However, to do this correctly it is vital to know what is meant by an equilibrium and how it is attained. This has been neatly discussed by Gibbs (5) and a mathematical discussion of the process for a stellar system was given by Lynden-Bell (4). Here we shall only discuss a much simplified model which contains what is important for our present purposes.

3. *A model of approach to equilibrium.* Consider a set of many non-interacting particles released in a frictionless pig-trough as in Fig. 1.

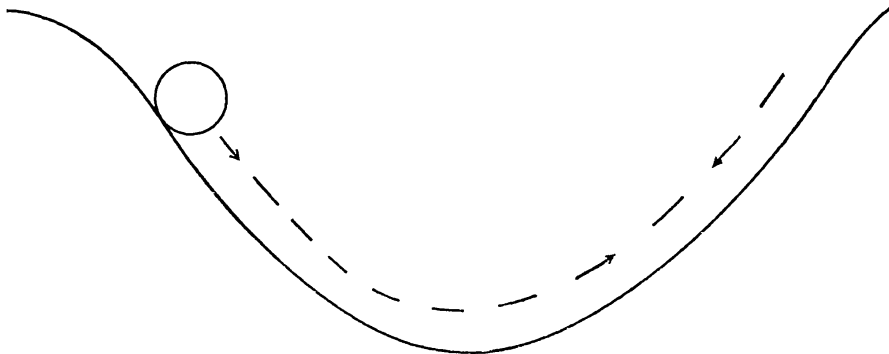


Fig. 1

Let the initial distribution function be $f(\epsilon, \phi)$ where ϵ is a single particle energy and ϕ is the phase of the oscillation across the pig-trough. Now plot phase space against ϕ and assume that like most dynamical systems the higher energy oscillations have longer periods. We plot contours of the distribution function $f(\epsilon, \phi, t)$. In fig. 2 $f(\epsilon, \phi, 0)$ is taken to be uniform inside a circle for illustration. The sequence of illustrations shows that

(i) The distribution function $f(\epsilon, \phi, t)$ never reaches an equilibrium if it is looked at on a fine enough scale. However, to see changes, progressively higher power must be used.

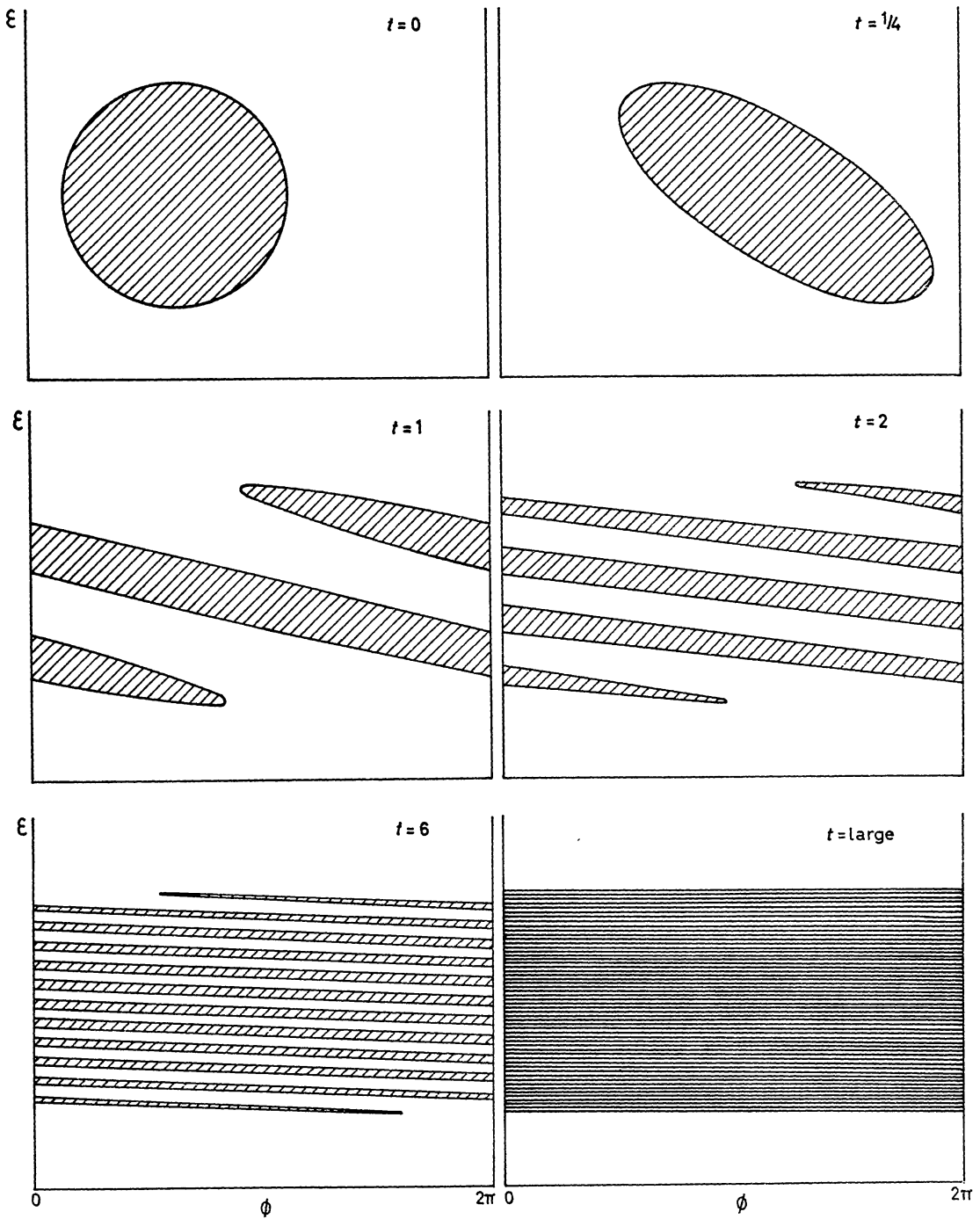


Fig. 2

(ii) If we use only a finite resolution then the distribution function appears to converge to what one would get by averaging $f(\epsilon, \phi, 0)$ over ϕ to obtain $\bar{f}(\epsilon)$.

(iii) In this process of looking with finite resolution we have averaged over regions of different phase densities. In this averaging process $M(f)$ is **not** conserved, so $M(\bar{f})$ is *not* the same function as $M(f)$. \bar{f} must, however, be the result of smoothing f thus there will be restrictions on $M(\bar{f})$. For instance the largest value \bar{f} attains can not be greater than the largest value f attained. Further it can be shown that the mathematical expression of (i) and (ii) is contained in the statement f converges in the mean to \bar{f} .

By our application of statistical mechanics we wish to predict the most probable distribution \bar{f} consistent with given conserved quantities. We do not wish $M(\bar{f})$ to be the same function as $M(f)$, however we shall require that \bar{f} is attained by smoothing f . The distinction between \bar{f} and f often occurs in statistical mechanics where \bar{f} is called the coarse-grained distribution and f the fine-grained distribution.

4. *The most probable state.* We re-write equations (11) and (12)

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + \mathbf{c} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{\partial}{\partial \mathbf{r}} \left\{ G \int \frac{f' d^3c'}{|\mathbf{r} - \mathbf{r}'|} d^3r' \right\} \cdot \frac{\partial f}{\partial \mathbf{c}} = 0 \quad (13)$$

where $f' = f(\mathbf{r}', \mathbf{c}', t)$.

Given f initially equation 13 determines f for all later times without mention of a stellar mass. It therefore does not matter how f is made up out of stars of different masses, it still obeys the same equation.

Evidently the equation describes the complicated motion in phase space of each element of phase. We shall assume that there are so many equivalent elements and such violent variations in the mean gravity field that a typical element of phase is equally likely to be found anywhere in phase space subject to the following restrictions:

(a) The total number of elements of phase which have any given phase density is the same as it was initially.

(b) The total energy is conserved.

(c) As a corollary of (a) no two elements of phase can overlap in phase space for then the phase space density would be different in the region of overlap.

We have ignored the integral of linear momentum because including it merely leads to the same system but in rectilinear motion.

The angular momentum integral will be included in Section 7. For a discussion of further conserved quantities which may or may not give isolating integrals see Appendix II.

Fixed phase density. We shall in this first calculation assume that all the elements of phase have the same density η . We show in Appendix I that a distribution of phase densities leads to similar results.

To apply statistics we turn our distribution into a set of numbers by dividing phase-space into a very large number of micro-cells each of volume $\tilde{\omega}$. These micro-cells will be so hyper-fine that even the fine grained distribution function f is adequately described by giving the mass of the phase element that occupies each cell. In the present instance these numbers will be 0 or $\eta\tilde{\omega}$. We shall group these micro-cells into coarse grained macro-cells each of which contains many micro-cells but is nevertheless so small that its spread in velocity and position space is infinitesimal compared that that of the whole galaxy. We call the number of micro-cells in each macro-cell ν so the volume of each large cell is $\nu\tilde{\omega}$.

In the present case all phase elements have volume $\tilde{\omega}$ and mass $\eta\tilde{\omega}$ so the total mass is $M = N\eta\tilde{\omega}$ where N is the number of occupied micro-cells.

Consider the configuration in which there are n_i phase elements in the i th macro-cell, each occupying one of the ν micro-cells with no cohabitation. The phase-elements are distinguishable so the number of ways of assigning a cell to the first element is ν , to the second $\nu - 1$ etc. The number of ways of assigning cells to all n_i elements is thus

$$\frac{\nu!}{(\nu - n_i)!} \quad (14)$$

Note that if we allowed cohabitation this number would be ν^{n_i} whereas if the elements were indistinguishable this number would be

$$\frac{\nu!}{(\nu - n_i)! n_i!} \quad (15)$$

as for Fermi–Dirac statistics.

To obtain the number of microstates corresponding to the configuration defined by the numbers n_i we must take the product of the terms such as (14) and multiply it by the number of ways of splitting our total of N elements into the groups n_i . Thus the number of microstates is

$$W = \frac{N!}{\prod_i (n_i)!} \times \prod_i \frac{\nu!}{(\nu - n_i)!} \quad (16)$$

which should be compared with the Maxwell–Boltzmann expression

$$W_{\text{MB}} = \frac{N!}{\prod_i n_i!} \times \prod_i \nu^{n_i} \quad (17)$$

and the Fermi–Dirac expression

$$W_{\text{FD}} = 1 \times \prod_i \frac{\nu!}{(n_i)! (\nu - n_i)!} \quad (18)$$

R. M. Lynden-Bell makes the interesting point that morphologically (6) there are four types of statistics as follows:

	Indistinguishable particles	Distinguishable particles
<i>No exclusion</i>	I Einstein–Bose	II Maxwell–Boltzmann
<i>Exclusion</i>	III Fermi–Dirac	IV

It is the fourth type of statistics that concerns us here. However, comparison of expressions (16) and (18) shows that we have arrived with a W which is the same as the Fermi–Dirac one but for a normalization. When all the particles are equivalent exclusion leaves the same number of full and empty microstates and the distinguishability or indistinguishability determines only whether we count each of these configurations $N!$ times or only once. Unlike particles, phase elements of *different* densities exclude one another so the new statistics IV of such elements is not equivalent to Fermi–Dirac statistics. This matter is explored in Appendix I.

Following normal procedures we now maximize $\log W$ subject to the constraints to obtain the most probable state

$$\log W = N(\log N - 1) - \sum_i [n_i(\log n_i - 1) + (\nu - n_i)\{\log(\nu - n_i) - 1\} - \nu(\log \nu - 1)].$$

It is convenient here to return to a representation in terms of a distribution function giving the average phase–density in the i th macro-cell

$$\bar{f}_i = f(\mathbf{r}_i, \mathbf{c}_i) = n_i \eta \tilde{\omega} / \nu \tilde{\omega} = \frac{n_i \eta}{\nu}$$

$$\log W = N(\log N - 1) -$$

$$\int \left[\frac{\nu}{\eta} \left\{ \bar{f} \left[\log \left(\frac{\nu \bar{f}}{\eta} \right) - 1 \right] + (\eta - \bar{f}) \left[\log \left\{ \frac{\nu}{\eta} (\eta - \bar{f}) \right\} - 1 \right] \right\} - \nu(\log \nu - 1) \right] \frac{d}{\nu}$$

where $d^6\tau$ indicates integration over all phase space. Our constraints are

$$\int f \frac{c^2}{2} d^6\tau - \frac{G}{2} \iint \frac{f f'}{|\mathbf{r} - \mathbf{r}'|} d^6\tau' d^6\tau = E$$

$$\int f d^6\tau = M$$

Using Lagrange multipliers $\alpha/\eta\tilde{\omega}$ and $\beta/\eta\tilde{\omega}$ to take account of our constraints we have

$$\delta(\log W) = 0 = - \int \frac{\delta f}{\eta\tilde{\omega}} \left\{ \log \frac{f}{\eta - f} + \alpha + \beta \left(\frac{c^2}{2} - \psi \right) \right\} d^6\tau \quad (19)$$

where

$$\psi(\mathbf{r}) = G \int \frac{f'}{|\mathbf{r} - \mathbf{r}'|} d^6\tau. \quad (20)$$

This last term arises from δE as follows

$$\delta E = 0 = \int \delta f \frac{c^2}{2} d^6\tau - \frac{G}{2} \iint \frac{\delta f f'}{|\mathbf{r} - \mathbf{r}'|} d^6\tau d^6\tau' - \frac{G}{2} \iint \frac{f \delta f'}{|\mathbf{r} - \mathbf{r}'|} d^6\tau d^6\tau'$$

by interchanging the dummy variables of integration τ and τ' the last terms are seen to be equal thus

$$\delta E = 0 = \int \delta f \left(\frac{c^2}{2} - G \int \frac{f'}{|\mathbf{r} - \mathbf{r}'|} d^6\tau' \right) d^6\tau.$$

Since the δf in equation (19) may now be considered to be independent the integrand must be zero.

$$\frac{f}{\eta - f} = \exp(-\alpha - \beta\epsilon)$$

where

$$\epsilon = \frac{c^2}{2} - \psi. \quad (21)$$

Hence

$$f = \eta \frac{\exp\{-\beta(\epsilon - \mu)\}}{1 + \exp\{-\beta(\epsilon - \mu)\}} \quad (22)$$

where μ , the chemical potential per unit mass, is defined as $-\alpha/\beta$. The equivalent result for a Fermi-Dirac gas is

$$f = mh^{-3} \frac{\exp\{-m\beta'(\epsilon - \mu)\}}{1 + \exp\{-m\beta'(\epsilon - \mu)\}}, \quad (23)$$

where h is Planck's constant and m is the mass of a particle and β' is $(kT)^{-1}$

Returning to the case in hand ψ is determined from equation (20). Differentiating that expression we obtain

$$\nabla^2\psi = -4\pi G \int f d^3c \quad (24)$$

writing in expression (22) for f the equation to be solved becomes

$$\nabla^2\psi = -16\pi^2\eta G\beta^{-3/2} \int_0^\infty \frac{\exp -x^2/2}{\exp\{-\beta(\psi + \mu)\} + \exp -x^2/2} x^2 dx \quad (25)$$

where $x^2 = \beta c^2$. This is, apart from a reinterpretation of the constants, the equation for the self-gravitating Fermi-Dirac gas. In the near fully degenerate case this has

been solved in connection with white dwarf stars. The form of the distribution function depends crucially on the degree of degeneracy. We devote the next section to estimating this.

As we explained in the introduction the distribution (22) is the form towards which mean field relaxation leads, rather than that necessarily attained. The mean field relaxation process is dependent on the strength of the variations in potential. As these die out the relaxation ceases and it is likely that the system may find a stable steady state before the relaxation process is completed. The distribution functions actually attained should be near (22) unless the stars were formed with a distribution function close to that of a steady state. In the latter case it would settle into the steady state with very little mean field relaxation. However, for the reasons given in the next section this case is very unlikely to occur in real galaxies. Incompleteness of the relaxation is further discussed in Section 6.

5. *Non-degeneracy of stellar systems.* Degeneracy becomes important when f attains phase-densities of order η . Now η is the typical phase-space density of a phase element, something which is conserved from birth. Thus we regard η as the phase space density at star formation. Leaving aside our belief that stars are made in clusters we must ask whether the stars could have been made at the low phase-space densities now observed in the field. If this density turns out to be too low for star formation then field stars will be far from degenerate.

In the following discussion we shall assume that the velocity dispersion of a set of newly formed stars is about the same as the velocity dispersion within the gas immediately prior to star-formation. From the condition that Jeans's criterion for instability must be satisfied for masses of stellar size we deduce that the phase-space densities at star formation must have been much higher than those now found in the field. Jeans's condition for instability towards fragmentation into masses as small as m^* is

$$\frac{c^3}{\rho^{1/2}} < m^* \left(\frac{G}{\pi} \right)^{3/2} \quad (26)$$

where c^2 is the velocity dispersion of elements of gas and ρ is the gas density. The distribution function phase density in the field is

$$f = \frac{\bar{\rho}}{c_g^3} \sim \frac{\bar{\rho}^{1/2}}{M} \left(\frac{\pi}{G} \right)^{3/2} \quad (27)$$

where $\bar{\rho}$ is the galaxy's density, c_g^2 its velocity dispersion and the last expression is derived by saying that the whole elliptical galaxy is the Jeans mass at its own density and dispersion. This last statement may be derived from the virial theorem. From equation 26 the phase density at star formation must satisfy

$$\eta = \frac{\rho}{c_g^3} > \frac{\rho^{1/2}}{m^*} \left(\frac{\pi}{G} \right)^{3/2} = \frac{M}{m^*} \left(\frac{\rho}{\bar{\rho}} \right)^{1/2} f. \quad (28)$$

Now ρ the density at star formation can hardly be less than $\bar{\rho}$ whereas M must be some $10^{10} m^*$. Thus f must be less than the phase space density at star formation, η , by a factor which is likely to be greater than 10^{10} .

Thus galaxies are in the non-degenerate limit and we may use the Maxwell-Boltzmann approximation to our statistics. That is $f \ll \eta$; hence from equation (22)

$$f = \eta \exp \{ -\beta(\epsilon - \mu) \} = A \exp(-\beta\epsilon) \quad (29)$$

where $A = \eta \exp(\beta\mu)$.

It should be emphasized again that ϵ is the energy per unit mass so this distribution function shows no segregation among stars of different masses. Our statistical analysis has led us back to the isothermal sphere with this equipartition of energy per unit mass rather than energy as the sole change. It should be noted that galactic nuclei may be an exception to the above argument and our type of degeneracy may be important there.

6. *Limitations of the relaxation.* In real systems relaxation is not complete. Variations in the radius of the galaxy as it falls about do not appreciably affect the potential at points well outside the main body of the galaxy. We therefore assume that the relaxation is spatially limited to the region within a sphere of radius R_1 about the centre of mass. To estimate R_1 suppose that the system were homogeneous with its present energy and oscillating about its equilibrium radius (given by $V = 2E$). Then it would fill a sphere of radius

$$R_1 = \frac{5}{3} \left(\frac{-2E}{GM^2} \right) = \frac{5}{3} R_0.$$

Furthermore we shall assume that considerable star-formation occurs only in the denser central regions inside such a sphere and that any stars now outside have only got there as a result of relaxation. All stars pass inside or close to R_1 and their orbits are at least partially relaxed. Orbits not satisfying these criteria are depopulated. A modification to the isothermal equilibrium that adequately allows for this depopulation is the replacement of

$$\bar{f} = A \exp(-\beta\epsilon) \quad (30)$$

by

$$f = A \exp[-\beta(\epsilon + \frac{1}{2}h^2 R_1^{-2})] \quad (31)$$

$$R_1 \sim \frac{5}{3} \frac{-2E}{GM^2}$$

where $h = |\mathbf{r} \times \mathbf{c}|$ the angular momentum per unit mass about the galactic centre. (Note we have dropped the bar on f .) For $r \geq R_1$ this provides a significant depopulation of stars moving transversely. For $r \ll R_1$ the distribution function is hardly affected. This modification has a long history in the works of Eddington (7), Oort & van Herk (8) and Michie (9) among others.

Even thus modified, the distribution (31) does not lead to a body of finite mass. The fundamental difficulty lies in the long range nature of gravity and the resulting divergence in the volume of phase space open to stars bound to the galaxy. Consider only the radial motions for bound stars $c_r^2 < 2\psi$ so the volume of phase space open to such radial motions is $\int 2\sqrt{2\psi} dr$. For any finite system $\psi \rightarrow GM/r$ as $r \rightarrow \infty$ so this integral diverges. This divergence becomes worse if non radial motions are allowed. Even theories that restrict tangential motions to zero and consider only phase space bound to the galaxy will still fail to give bodies of finite mass if equal weights are allotted in phase space. The infinite weight of phase space at infinity is overpowering. An example which exemplifies this is the distribution function

$$f = A' \delta(h^2) \exp(-\beta\epsilon) \quad \epsilon \leq 0 \quad (32)$$

(δ is Dirac's δ function)

to which (31) tends at great distances. Integration over velocities leads to

$$\rho \simeq 2A' \frac{\sqrt{2\psi}}{r^2} \quad \text{for } \psi \text{ small.} \quad (33)$$

This density distribution clearly leads to an infinite total mass. This difficulty has arisen because we have applied statistical mechanical arguments without due care to their domain of validity. Two important effects that we have so far omitted occur because

(i) relaxing conditions persist only while the galaxy is dynamically unsteady. Orbits which lie partly outside the relaxing region and have periods longer than the time for which the galaxy is unsteady will not acquire their full quotas of stars; and

(ii) in practice the galaxy will not be isolated but will be subject to the tides of other systems.

Both of these lead to similar effects so it may be difficult to decide which process is dominant except in a few obvious cases.

We now consider a Fokker-Planck equation to describe the encounterless relaxation process. To do this our first aim is to find how the relaxation time depends on velocity for the high velocity phase elements. Let such an element encounter fluctuations of potential of magnitude $\delta\psi$ of length scale L . Then the number of fluctuations undergone in time t is ct/L where c is the element's velocity. The change in energy in each fluctuation is proportional to $\delta\psi$, so

$$(\Delta\epsilon)^2 = \alpha(\delta\psi)^2.$$

Defining the relaxation time by

$$\sum (\Delta\epsilon)^2 = (\frac{1}{2}c^2)^2, \quad \text{the summation being over time } T_r$$

we find

$$\frac{cT_r}{L} \alpha(\delta\psi)^2 = \frac{1}{4}c^4. \quad (34)$$

$$T_r \propto c^3.$$

We must point out that relaxation by this mechanism is important only in the central parts where most of the stars have achieved an equilibrium distribution under its influence. The precise form of the Fokker-Planck diffusion terms is unimportant for them because Maxwell's distribution is unchanged. It is the high velocity tail of the distribution functions in the central regions of the cluster which both departs from the equilibrium distribution and is subject to the relaxation. For such elements (34) is a good approximation. Following Chandrasekhar (10) (11) we consider each phase element as diffusing in velocity space under the influence of this relaxation. In terms of our statistical mechanical model it is elements of \bar{f} that is the smooth distribution function that undergo this diffusion which would be impossible for elements of f . Pure diffusion would be represented by

$$\left(\frac{\partial \bar{f}}{\partial t}\right)_{\text{fluctuation}} = + \frac{\partial}{\partial \mathbf{c}} \left(q \frac{\partial \bar{f}}{\partial \mathbf{c}} \right)$$

however under the influence of pure diffusion \bar{f} would smooth itself to ever higher energies and the Maxwellian would not be stable against encounters. Chandrasekhar's dynamical friction allows for conservation of energy in the

presence of diffusion and gives the form

$$\left(\frac{\partial \bar{f}}{\partial t}\right)_{\text{fluctuation}} = \frac{\partial}{\partial \mathbf{c}} \cdot \left(q \frac{\partial \bar{f}}{\partial \mathbf{c}} + \eta \bar{f} \mathbf{c} \right) \quad (35)$$

where η the coefficient of dynamical friction is inversely proportional to the relaxation time. The invariance of the Maxwell distribution function (29) under the influence of these fluctuations yields the relation $q\beta = \eta$.

The change in \bar{f} due to fluctuations occurs as we follow the same piece of phase through the average non-fluctuating field thus the full equation for the evolution of the system is

$$\frac{D\bar{f}}{Dt} = \frac{\partial}{\partial \mathbf{c}} \cdot \left[\frac{\eta}{\beta} \left(\frac{\partial \bar{f}}{\partial \mathbf{c}} + \beta \bar{f} \mathbf{c} \right) \right] \quad (36)$$

where D/Dt follows the motion in the non-fluctuating average potential ψ . Remembering that where this R.H.S. is important $T_r \propto c^3$ and that $\eta \propto 1/T_r$ we may write

$$\frac{D\bar{f}}{Dt} = \frac{\partial}{\partial \mathbf{c}} \cdot \left[\frac{K}{c^3} \left(\frac{\partial \bar{f}}{\partial \mathbf{c}} + \beta \bar{f} \mathbf{c} \right) \right] \quad (37)$$

Now in the central region of interest $\bar{f} = \bar{f}(c^2/2 - \psi) = \bar{f}(\epsilon)$ so we convert the R.H.S. to an equation in terms of energies

$$\begin{aligned} \frac{\partial}{\partial \mathbf{c}} \cdot \left[\frac{K\mathbf{c}}{c^3} \left(\frac{\partial \bar{f}}{\partial \epsilon} + \beta \bar{f} \right) \right] &= - \frac{\partial}{\partial \mathbf{c}} \cdot \left[\frac{\partial}{\partial \mathbf{c}} \left(\frac{1}{c} \right) K \left(\frac{\partial \bar{f}}{\partial \epsilon} + \beta \bar{f} \right) \right] \\ &= \frac{\mathbf{c}}{c^3} \cdot \frac{\partial}{\partial \mathbf{c}} \left[K \left(\frac{\partial \bar{f}}{\partial \epsilon} + \beta \bar{f} \right) \right] = \frac{1}{c} \frac{\partial}{\partial \epsilon} \left[K \left(\frac{\partial \bar{f}}{\partial \epsilon} + \beta \bar{f} \right) \right] \end{aligned}$$

where I have not worried about the singularity at $c = 0$ because this approximation to the Fokker-Planck coefficient is incorrect there anyway. Thus the equation true for the high energies is

$$\frac{D\bar{f}}{Dt} = \frac{1}{c} \frac{\partial}{\partial \epsilon} \left[K \left(\frac{\partial \bar{f}}{\partial \epsilon} + \beta \bar{f} \right) \right] \quad (38)$$

while at low energies \bar{f} is Maxwellian to a sufficient approximation. Michie's solution (9) of this equation is to put $K(\partial \bar{f} / \partial \epsilon + \beta \bar{f}) = \text{const} = K\beta B$ say which clearly gives a stationary solution in the outer parts. We then obtain

$$\begin{aligned} \bar{f} - B &= A \exp(-\beta\epsilon) \\ \bar{f} &= A \exp(-\beta\epsilon) - B. \end{aligned} \quad (39)$$

\bar{f} becomes zero at the energy $\epsilon = -1/\beta \log(B/A) = \epsilon_e$ which is identified as the energy of escape in the tidal case. For $\epsilon > \epsilon_e$, $\bar{f} = 0$ since stars there would be removed by the tidal field. In practical cases the escape energy is near zero so $B = A \exp(-\beta\epsilon_e)$ is small compared with $A \exp(-\beta\epsilon)$ and the Maxwellian is only significantly modified close to the escape energy. Similar considerations applied to our modified distribution function which is Maxwellian where the diffusion occurs but anisotropic in the outer parts, leads to the function

$$\bar{f} = (A \exp(-\beta\epsilon) - B) \exp -\frac{1}{2}\beta h^2 R_1^{-2} \quad (40)$$

In the general case it is instructive to assume that the central regions of the

galaxy are not changing rapidly so that β changes slowly if at all. We return to our non-stationary equation (38) and write it

$$\frac{Df}{Dt} = \frac{1}{c} \frac{\partial}{\partial \epsilon} \left[K \exp(-\beta\epsilon) \frac{\partial}{\partial \epsilon} (f \exp \beta\epsilon) \right] \quad (41)$$

Making a somewhat drastic assumption we shall replace c by $c_e = \sqrt{2\psi_0}$ the velocity of escape from the central regions on the grounds that the R.H.S. is significant only for high energy stars in the central regions. The diffusion coefficient K is independent of ϵ though it will depend on time rapidly becoming zero as the oscillations of the galaxy die out. We define a new time variable

$$\tau = \int_0^t (2\psi_0)^{-1/2} K dt \quad (42)$$

and note that τ will tend to some finite value τ_∞ as $t \rightarrow \infty$. Looking for solutions of equation (38) of the form $f = \bar{f}(\epsilon, t)$, the operator D/Dt reduces to $\partial/\partial t$ and when we use τ as independent variable we have

$$\frac{\partial \bar{f}}{\partial \tau} = \frac{\partial^2 \bar{f}}{\partial \epsilon^2} + \beta \frac{\partial \bar{f}}{\partial \epsilon} \quad (43)$$

This is a form of diffusion equation. We need solutions such that $\bar{f} = 0$ at ϵ_e the energy of escape and such that there is no flux of stars towards higher energies coming through the energy $\epsilon = -\psi_0$ (which corresponds to the energy of stationary stars at the cluster centre). Thus $\partial \bar{f} / \partial \epsilon + \beta \bar{f} = 0$ at $\epsilon = -\psi_0$. Noting that equation (43) is linear in \bar{f} with constant coefficients its solution is of the form

$$\bar{f} = \sum_i A_i \exp(-\sigma_i \tau - \lambda_i \epsilon)$$

where λ_i and σ_i are related by

$$-\sigma_i = \lambda_i^2 - \beta \lambda_i$$

Thus

$$\lambda_i = \frac{\beta}{2} \pm \sqrt{\frac{\beta^2}{4} - \sigma_i} = \begin{cases} \lambda_i^+ \\ \lambda_i^- \end{cases} \text{ say} \quad (44)$$

In order to satisfy the boundary conditions at the escape energy \bar{f} must take the form

$$\bar{f} = \sum_i A_i \exp -\sigma_i \tau \cdot [\exp -\lambda_i^+(\epsilon - \epsilon_e) - \exp -\lambda_i^-(\epsilon - \epsilon_e)]$$

while the other boundary condition leads to the relationship

$$(\lambda_i^+ - \beta) \exp \lambda_i^+(\psi_0 + \epsilon_e) = (\lambda_i^- - \beta) \exp \lambda_i^-(\psi_0 + \epsilon_e) \quad (45)$$

Putting $s = \sqrt{(\beta^2/4 - \sigma_i)}$ and using equation (44)

$$(s - \beta/2) \exp s(\psi_0 + \epsilon_e) = (-s - \beta/2) \exp -s(\psi_0 + \epsilon_e)$$

so

$$\frac{2s}{\beta} = \tanh [s(\psi_0 + \epsilon_e)] \quad (46)$$

It is normally the case that $\frac{1}{2}\beta(\psi_0 + \epsilon_e) > 1$ so equation (46) will have one real and many imaginary solutions for s . The imaginary solutions correspond to larger eigenvalues σ_i , that is to more rapidly decaying eigen modes of the diffusion equation, than the one real solution. In so far as different solutions will converge

as the diffusion proceeds they will do so to the most persistent mode. We therefore consider the situation when only this mode with real s has survived. Assuming $\frac{1}{2}\beta(\psi_0 + \epsilon_e)$ is not close to 1 the approximate solution of equation (46) is

$$s = \frac{\beta}{2} th \left[\frac{\beta}{2} (\psi_0 + \epsilon_e) \right] \simeq \frac{\beta}{2} \left(1 - 2 \exp \left[-\frac{\beta}{2} (\psi_0 + \epsilon_e) \right] \right)$$

so the solution for \bar{f} tends to

$$\bar{f} = A [\exp -\beta(1-Q)(\epsilon - \epsilon_e) - \exp -\beta Q(\epsilon - \epsilon_e)] \quad (47)$$

where $Q = \exp \left[-\frac{1}{2}\beta(\psi_0 + \epsilon_e) \right]$,

$$A = A_1 \exp \left[-\beta^2 Q \tau_\infty \right]$$

and I have written τ_∞ for τ .

In practical cases the system is Maxwellian over quite a range of energies and $\beta(\psi_0 + \epsilon_e)$ is of order 6 to 10. Q is therefore small. It will be noted that in the limit $Q=0$ the form of \bar{f} reduces to equation (39). For small Q the differences are negligible so we shall use Michie's solution (39), modified by him to include the anisotropic velocities equation (40), in all cases.

Michie & Bodenheimer (12) have already computed the detailed density distributions corresponding to such models. King (2) using the isotropic special case of these models has computed all the quantities required and has given a thorough discussion of the observations. Further support comes from a detailed discussion of Baum's Observations of M 87.

7. *Rotating elliptical systems.* To allow for conservation of angular momentum we must maximize $\log W$ as we did in equation (19) but subject to the extra constraint that

$$\int \bar{f} \mathbf{r} \times \mathbf{c} d^3c = \mathbf{H}$$

Introducing a vector Lagrange multiplier $\boldsymbol{\gamma}/\eta W$ we obtain

$$\delta(\log W) = 0 = - \int \frac{\delta \bar{f}}{\eta \bar{\omega}} \left\{ \log \frac{\bar{f}}{\eta - \bar{f}} + \alpha + \beta \left(\frac{c^2}{2} - \psi \right) + \boldsymbol{\gamma} \cdot (\mathbf{r} \times \mathbf{c}) \right\} d^6\tau$$

Hence

$$\bar{f} = \eta \frac{\exp \{ -\beta[\epsilon - \mu - (\boldsymbol{\Omega} \times \mathbf{r}) \cdot \mathbf{c}] \}}{1 + \exp \{ -\beta[\epsilon - \mu - (\boldsymbol{\Omega} \times \mathbf{r}) \cdot \mathbf{c}] \}} \quad (48)$$

where $\boldsymbol{\Omega} = -\boldsymbol{\gamma}/\beta$.

If we again use the non-degeneracy condition we obtain the uniformly rotating Maxwellian

$$\bar{f} = A \exp \{ -\beta[\epsilon - (\boldsymbol{\Omega} \times \mathbf{r}) \cdot \mathbf{c}] \} \quad (49)$$

Modification of this law is necessary to take account of the spatial limitation of the relaxation. However we can no longer account for this by a depopulation factor

$$\exp \left\{ -\beta \frac{h^2}{2R_1^2} \right\} \quad (50)$$

because the total angular momentum per unit mass of a star is no longer an integral of the motion. To replace h in this formula by the angular momentum about the galactic axis is insufficient to ensure that stars reach the relaxing region, and leads to pancake shaped velocity dispersion ellipsoids with the galactic centre in the plane

of the pancake. To put in the restriction that the stars pass through the relaxing region we must invoke a third integral. In general these integrals are not in closed form but to resort to a computer at this early stage before we even know what potential interests us seems unpromising. We note that our depopulation factor must be important only for stars of relatively high energy. In the region covered by such orbits the potential may probably be approximated by one of Eddington's type (13)–(16) viz.

$$\psi = \frac{\zeta(\lambda) - \eta(\mu)}{\lambda - \mu} \quad (51)$$

where λ and μ are spheroidal coordinates and ζ and η are arbitrary functions. [To avoid singularities at the foci ζ and η should be taken as the same functional form. λ and μ have ranges with but one value in common.] For large r where our depopulation factor will be important this reduces to the simpler form

$$\psi = F(r) - \frac{g(\theta)}{r^2} \quad (52)$$

where F and g are arbitrary and g without loss of generality may be taken to be zero on the galactic axis. Eddington's potentials are separable and yield exact quadratic third integrals. Indeed in the potentials (52) the third integral is (16)

$$I = \frac{h^2}{2} - g(\theta) \quad (53)$$

Although real systems will not show this exact separation, nevertheless this integral is simple enough to be used and is presumably a good approximation to the complicated adelpic (17) or third integrals that appear to exist for most orbits in smooth galaxy-like potentials (17–20). We therefore take our depopulation factor to be

$$\exp \left\{ -\beta \frac{h}{R_1^2} \right\} \quad (54)$$

which leads to distribution functions of the form

$$f = A \exp \left\{ -\beta \left[\epsilon - (\boldsymbol{\Omega} \times \mathbf{r}) \cdot \mathbf{c} + \frac{1}{2} \frac{h^2}{R_1^2} - \frac{g(\theta)}{R_1^2} \right] \right\} \quad (55)$$

Orienting the polar axis along $\boldsymbol{\Omega}$ and writing \mathbf{c} in spherical polar coordinates $\mathbf{c} = (c_r, c_\theta, c_\phi)$ and R for distance from the axis, this becomes

$$f = A \exp \left\{ -\beta \left[\frac{c_r^2}{2} + \frac{1}{2} \left(1 + \frac{r^2}{R_1^2} \right) \left(c_\phi - \frac{\Omega R}{1 + r^2/R_1^2} \right)^2 + \frac{1}{2} \left(1 + \frac{r^2}{R_1^2} \right) c_\theta^2 - \frac{\Omega^2 R^2}{2(1 + r^2/R_1^2)} - \psi - \frac{g(\theta)}{R_1^2} \right] \right\}$$

which shows a velocity law of approximately

$$\bar{c}_\phi = \frac{\Omega R}{1 + r^2/R_1^2} \quad (56)$$

Note that this velocity decreases as one rises from the galactic plane. It should be noted however that our assumptions are not suitable for a disk-like galaxy for which there are certainly stars moving in circular orbit which never pass through any relaxation region. The above f suitably modified for escape by subtracting a constant might well represent a halo population but the disk would have to be added. On the galactic axis $R=0$, $g(\theta)=0$ and the distribution function reduces to

the form considered for non-rotating systems. The density distribution up the galactic axis will again follow the observed law at large distances from the centre. More generally in the limit of large r the density distribution unmodified for escape is

$$\rho \propto r^{-2} \left(\psi + \frac{g(\theta)}{R_1^2} + \frac{1}{2} \Omega^2 R_1^2 \sin^2 \theta \right)^{1/2} \quad (57)$$

It is interesting to compare this with the functional forms obtained with no velocity anisotropy viz.

$$\rho = \rho(\psi + \frac{1}{2} \Omega^2 r^2 \sin^2 \theta). \quad (58)$$

The flattening of the equidensity contours of this formula is very pronounced in the outer parts and increases rapidly with r . Formula (57) gives a much more nearly constant eccentricity to these contours in agreement with the observation that the isophotes of E0–E5 elliptical galaxies have an ellipticity almost independent of projected radius (21).

To explore these matters further it is necessary to compute the self-gravitating models corresponding to this distribution function, a very considerable project in its own right and outside the scope of the present discussion.

Disk systems. In systems of high angular momentum the outer parts are held far from the central regions and have no chance to share in any overall equilibrium. The shearing motions keep the system far from equilibrium and there is evidence that the invariant function $M(h)$ which gives the total mass with angular momentum per unit mass greater than h is similar to that of the uniformly rotating uniform sphere or spheroid. Statistical mechanics even when conservation of $M(h)$ is supplied as a constraint gives no difference between the velocity dispersion in the direction of the galactic centre and in the direction perpendicular to the galactic plane whereas this difference is a major feature of observed dispersions. If significant relaxation has occurred it must have been driven by anisotropic forces possibly connected with spiral arm formation.

8. *Comparison with observations.* For distribution functions that are approximately Maxwellian the inner parts of the light distribution are adequately fitted by scaling the model. King has shown convincingly that only one further parameter—the tidal cut-off—gives a good fit to the outer parts. No further easily available observational parameters exist to test the anisotropy of the velocity distributions in spherical galaxies. However when we consider rotating galaxies the variation of the eccentricity with radius provides such a parameter. Dickens & Woolley (22) found their rotating truncated Maxwellian was too flat in the outer parts to fit the observations of ω Centauri. Anisotropy will lead to the much more nearly constant eccentricities reported for elliptical galaxies (21). It will therefore be important to compute rotating models and compare them with observations. An edge-on Elliptical with accurately determined light distribution and detailed knowledge of rotation and velocity dispersion over the photographic image is within reach with an image tube and could do much when combined with theory to give an accurate mass and mass to light ratio. These in turn would help to determine the dominant stellar type of which these galaxies are made.

It is not yet really clear from observations that anisotropy is important at all. It could be that projection factors have led to galaxies with Roche model shapes (23) appearing far rounder than they actually are in their outer parts.

9. *Conclusions doubts and speculations.* I hope this paper has exploded the paradox of stellar dynamics that a system with a long relaxation time can nevertheless look like a Maxwellian. It provides a basis for using Maxwell's distribution with temperature proportional to mass and its modifications to galaxies. The work of Michie and King is thus vindicated. It is pointed out that measurement of the isophotal eccentricity as a function of distance from the centre provides information on the anisotropy of the stellar orbits in the outer parts.

Applied to clusters of galaxies such as Coma these ideas remove Zwicky's paradoxical age of 10^{16} – 10^{18} years (24).

However this work suffers from the following defects

(i) The size of the relaxation region was determined crudely, equation (31). It is important to determine this observationally as discussed in Section 8.

(ii) The general theory of the new statistics predicts a superposition of Maxwellians; it is not clear how far our use of just one is justified.

(iii) A detailed theory of the relaxation process including anisotropy should predict a relation between τ_{∞} and R_1 .

(iv) The theory of rotating systems should be worked out properly including integrations for the density.

(v) There is a fundamental weakness in the whole theory as presented and in stellar dynamics generally which may cause violent and exciting phenomena. This has been ignored in our tacit assumption of the existence of an equilibrium state. I enlarge on this below.

(vi) We have ignored knowledge of the other conserved quantities discussed in Appendix II.

Consider two stellar systems nearly but not quite identical and suppose that each is approximately Maxwellian at its centre so that we may talk of its temperature. Owing to the Virial theorem they will, like stars, grow hotter if energy is lost and cooler if energy is gained. We imagine some form of energy transfer between the two similar to thermal conduction in the sense that energy flows from the hotter cluster to the cooler one. Far from approaching equilibrium the hotter cluster gets hotter still while the cooler gains energy and cools down. There is no tendency here to energy sharing—quite the reverse. This basic tendency towards disequilibrium is further illustrated by Antonov's discovery that it can all be done on one cluster the centre becoming hotter and denser while the outside expands. His discussion shows that even if the system is enclosed in a large perfectly elastic sphere off which stars bounce with impunity there is no equilibrium for N stars of mass M unless the binding energy is small enough (i.e. close enough to zero). His result may be put in a different way. Let the binding energy be $-E$, and the volume of the sphere V . If E is unchanged but V is increased we reach a stage at which there is no relaxed equilibrium, the entropy is not even a local maximum at Maxwell's distribution the system can gain entropy by condensing its central parts and growing hotter while giving out the excess energy to the outer parts.

In many ways this behaviour is like a phase transition—it leads us to expect very concentrated systems immersed in or in contact with very diffuse ones, it also greatly stimulates the imagination particularly with a view to explaining violent events in old elliptical galaxies. However it should be noticed that the basic mechanism involved is an exchange of 'heat energy'. It is not clear that there is a relaxation mechanism in such systems that can lead to this exchange. I shall report elsewhere my investigations of Jeans' gravitational instability in such a cluster

when encounters are completely neglected but I have shown that Antonov's critical volume is not associated with an encounterless spherical instability of that type although it coincides with the critical instability radius for an isothermal gas sphere surrounded by a box. I consider the basic urge for disequilibrium in self-gravitating systems to sound the death knell of Boltzmann's hypothesis of the universe as an improbable deviation from equilibrium; it is probable that no equilibrium can exist even in theory!

It is the phenomena mentioned in the last paragraph which provide the possibility of new and startling results in stellar dynamics. Much more work should be done on these topics even if they are intrinsically difficult to investigate; in particular high order correlations corresponding to hierarchies of dynamical clusterings may be the typical 'equilibrium' of a stellar system rather than the smooth systems assumed by all theorists to date. Heat exchange could then occur and excessive densities could be attained stars would have glancing collisions stripping off parts of their atmospheres which would expand to the size of the whole condensed region due to the energy of collision. Following on these lines one is led to a system embedded in hot gas clouds with the cores of stars hurtling through. The low average density of gas would make 1-particle phenomena dominate, so electron scattering would be important and the spectrum should show wide emission lines. All these phenomena are so reminiscent of the nuclei of Seyfert galaxies that I am tempted to say that it happens but I think it is wrong to accept such speculations until it has been demonstrated that the typical equilibrium of a stellar system is hierarchical rather than smooth and the theory of hierarchical systems has been developed.

Acknowledgment. I would like to thank Professor Oort for the hospitality of the Leiden Observatory where this work began. I would also like to thank Dr Ostriker for stimulating discussions of fluctuations in stellar dynamics.

Royal Greenwich Observatory.
1966 March.

APPENDIX I

General case of the new statistics. Let there be a total number N_J of elements of phase of density η_J . Let their total mass $N_J \eta_J \tilde{\omega} = M_J$. Furthermore let the number of these elements in the i th macrocell be n_{iJ} . Then

$$\sum_i n_{iJ} = N_J \quad \text{all } J.$$

Consider the configuration of n_{iJ} phase-elements of class J in the i th macrocell each occupying one of the ν microcells with no cohabitation even among elements of different classes. The number of ways of assigning a cell to the first element is ν to the second $\nu - 1$ etc. Thus the number of ways of assigning micro-cells to all $\sum_J n_{iJ}$ elements in the i th macro-cell is

$$\frac{\nu!}{(\nu - \sum_J n_{iJ})!}$$

The total number of microstates W corresponding to the single macrostate defined by the numbers n_{iJ} is the product of the above numbers with the total

number of ways of splitting our N_J elements into the groups n_{iJ} . So,

$$W = \prod_J \frac{N_J!}{\prod_i (n_{iJ})!} \times \prod_i \frac{\nu!}{(\nu - \sum_J n_{iJ})!}$$

Hence

$$\begin{aligned} \log W = & \sum_J N_J (\log N_J - 1) - \sum_J \sum_i n_{iJ} (\log n_{iJ} - 1) + \\ & + \sum_i \{ \nu (\log \nu - 1) - (\nu - \sum_J n_{iJ}) [\log (\nu - \sum_J n_{iJ}) - 1] \} \end{aligned}$$

We now convert to phase-space densities by writing

$$\bar{f}_J(\mathbf{r}_i \mathbf{c}_i) = \frac{n_{iJ} \eta_J \tilde{\omega}}{\nu \tilde{\omega}} = \frac{n_{iJ} \eta_J}{\nu}$$

The coarse grained distribution function itself is

$$\bar{f}(\mathbf{r}_i \mathbf{c}_i) = \frac{\sum_J n_{iJ} \tilde{\omega}}{\nu} = \sum_J \bar{f}_J(\mathbf{r}_i, \mathbf{c}_i)$$

In this notation

$$\begin{aligned} \log W = \text{constant} - \int \sum_J \frac{\nu}{\eta_J} \bar{f}_J \left[\log \left(\frac{\nu \bar{f}_J}{\eta_J} \right) - 1 \right] + \nu \left[1 - \sum_K \frac{f_K}{\eta_K} \right] \\ \left\{ \log \left[\nu \left(1 - \sum_K \frac{f_K}{\eta_K} \right) \right] - 1 \right\} \frac{d^6 \tau}{\nu \tilde{\omega}} \end{aligned}$$

and our constraints are

$$\begin{aligned} \int \frac{\bar{f}^2 c^2}{2} d^6 \tau - \frac{G}{2} \iint \frac{\bar{f} \bar{f}'}{|\mathbf{r} - \mathbf{r}'|} d^6 \tau d^6 \tau = E \\ \int \bar{f}_J d^6 \tau = M_J \quad \text{all } J. \end{aligned}$$

Using Lagrange multipliers $\alpha_J / \tilde{\omega} \eta_J$ and $\beta / \tilde{\omega} \eta$ where $\eta = \sum_J \eta_J M_J / \sum_J M_J$ we obtain

$$\delta \log W = - \sum_J \int \frac{\nu}{\eta_J} \delta \bar{f}_J \left\{ \log \left(\frac{\nu \bar{f}_J}{\eta_J} \right) - \log \left[\nu \left(1 - \sum_K \frac{f_K}{\eta_K} \right) \right] + \alpha_J + \beta \frac{\eta_J}{\eta} \epsilon \right\} \frac{d^6 \tau}{\nu \tilde{\omega}}$$

Letting this equal to zero for all $\delta \bar{f}_J$ which may now be considered independent we obtain

$$\log \left\{ \frac{\bar{f}_J}{\eta_J \left(1 - \sum_K \frac{f_K}{\eta_K} \right)} \right\} = -\alpha_J - \beta \frac{\eta_J}{\eta} \epsilon$$

or

$$\bar{f}_J = \eta_J \left(1 - \sum \right) \exp -\beta_J (\epsilon - \mu_J)$$

where

$$\sum = \sum_K \frac{f_K}{\eta_K} \quad \text{and} \quad \beta_J = \frac{\eta_J}{\eta} \beta, \quad \mu_J = -\frac{\alpha_J}{\beta_J}$$

dividing by η_J and summing over J

$$\sum = \left(1 - \sum \right) \sum_J \exp -\beta_J (\epsilon - \mu_J)$$

hence

$$1 - \sum = [1 + \sum_J \exp -\beta_J(\epsilon - \mu_J)]^{-1}$$

$$f_J = \eta_J \frac{\exp -\beta_J(\epsilon - \mu_J)}{1 + \sum_J \exp -\beta_J(\epsilon - \mu_J)}$$

so our coarse grained distribution function f is

$$f = \sum_J \eta_J \frac{\exp -\beta_J(\epsilon - \mu_J)}{1 + \sum_J \exp -\beta_J(\epsilon - \mu_J)}$$

Note the interesting differences from Fermi-Dirac statistics.

By the argument of paragraph 5 we expect to be in the extreme non-degenerate limit $f_J \ll \eta_J$ for all J . Then each factor

$$\exp -\beta_J(\epsilon - \mu_J)$$

has to be small and we obtain the non-degenerate approximation

$$f_J = \sum A_J \exp -\beta_J \epsilon$$

where $A_J = \eta_J \exp + \beta_J \mu_J$ is determined from the condition $\int f_J d^6\tau = M_J$

$$\beta_J = \beta \frac{\eta_J}{\bar{\eta}} = \beta \frac{\eta_J M}{\sum_J \eta_J M_J}$$

and β is in turn determined from the total energy condition.

Our result shows the correct distribution function to be a superposition of Maxwellian components whose velocity dispersions are inversely proportional to the phase space density of the component at star formation. It should be remarked that if the luminosity function at birth is biased towards heavy stars whenever birth conditions have large phase space densities then such stars will have that same bias towards the central regions of the equilibrium object. This effect will be much smaller than the segregation by mass when equipartition of energy occurs and should not be confused with the latter.

APPENDIX II

Kelvin's Theorem generalized to phase space. If we were to apply the strictest dictates of statistical mechanics we should keep all conserved isolating integrals to definite values. We have not done this because there are many infinities of them. In particular we have the generalization of Kelvin's Theorem given below.

Consider any closed phase space path l_6 occupied by phase elements. Its projection defines a path l_3 in position space. We measure path length s around l_3 and associated with each point $P(s)$ on l_3 will be a point $Q(s)$ on l_6 which projects into P . Let the coordinates of $Q(s)$ be $(\mathbf{r}(s), \mathbf{c}(s))$. We prove that $\oint \mathbf{f}_i \cdot \mathbf{c}(s) \cdot ds$ is constant where l_3 is taken to move with the phase elements which compose it. The proof follows one of the normal proofs of Kelvin's theorem. Let D/Dt be

convective derivative following the phase space flow then

$$\begin{aligned} \frac{D}{Dt} \oint_{I_i} \mathbf{c}(s) \cdot d\mathbf{s} &= \oint \frac{D\mathbf{c}}{Dt} \cdot d\mathbf{s} + \oint \mathbf{c} \cdot \frac{D(d\mathbf{s})}{Dt} \\ &= \oint \frac{\partial \psi}{\partial \mathbf{r}} \cdot d\mathbf{s} + \oint \mathbf{c} \cdot \left(d\mathbf{s} \frac{\partial}{\partial s} \right) \mathbf{c} \\ &= 0 + \oint \frac{d}{ds} \left(\frac{c^2}{2} \right) ds = 0 \end{aligned}$$

The equivalent theorem for collisionless plasmas in magnetic fields may be shown to be

$$\oint_I \mathbf{c} \cdot d\mathbf{s} + \frac{q}{mc} \int_S \mathbf{B} \cdot d\mathbf{S} = \text{constant}$$

where \mathbf{S} is any surface spanning I , q is the charge of the particle species concerned and m its mass and c is the velocity of light. This result is already known in plasma physics.

References

- (1) M. Henon, 1964. *Anns Astrophys.*, **27**, 83.
- (2) I. R. King, 1966. *Astr. J., N.Y.*, **71**, 64.
- (3) M. Lecar, 1966. *I.A.U. Symposium No. 25*, p. 46. Academic Press.
- (4) D. Lynden-Bell, 1962. *Mon. Not. R. astr. Soc.*, **124**, 279.
- (5) J. W. Gibbs, 1900. *Elementary Principles of Statistical Mechanics*, p. 144. Dover.
- (6) F. Zwicky, 1957. *Morphological Astronomy*. Springer, Berlin.
- (7) A. S. Eddington, 1915. *Mon. Not. R. astr. Soc.*, **75**, 366.
- (8) J. H. Oort & G. Van Herk, 1959. *Bull. astr. Insts Neth.*, **14**, 299.
- (9) R. W. Michie, 1963. *Mon. Not. R. astr. Soc.*, **125**, 127.
- (10) S. Chandrasekhar, 1942. *Principles of Stellar Dynamics*, p. 251. Dover.
- (11) S. Chandrasekhar, 1943. *Rev. mod. Phys.*, **15**, 1.
- (12) R. W. Michie & R. H. Bodenheimer, 1963. *Mon. Not. R. astr. Soc.*, **126**, 267.
- (13) A. S. Eddington, 1915. *Mon. Not. R. astr. Soc.*, **76**, 37.
- (14) G. L. Clark, 1937. *Mon. Not. R. astr. Soc.*, **97**, 182.
- (15) G. G. Kuzmin, 1956. *Tartu Ülik. Öppe- ja Katsem. Teat.*, No. 2.
- (16) D. Lynden-Bell, 1962. *Mon. Not. R. astr. Soc.*, **124**, 95.
- (17) E. T. Whittaker, 1937. *Treatise on Analytical Dynamics* (4th edn), pp. 433-449. Cambridge.
- (18) G. Contopoulos, 1957. *Stokh. Obs. Annlr*, Vol. 19, No. 10. 1958. *Stokh. Obs. Annlr*, Vol. 20, No. 5.
- (19) A. Ollengren, 1965. *A. Rev. Astr. Astrophys.*, **3**, 113.
- (20) G. Contopoulos & L. Woltjer, 1964. *Astrophys. J.*, **140**, 1106.
- (21) C. J. Van Houten, 1962. *Bull. astr. Insts Neth.*, **16**, 1.
- (22) R. J. Dickens & R. v. d. R. Woolley, 1967. *R. Obs. Bull.*, (in press).
- (23) J. H. Jeans, 1919. *Cosmogony and Stellar Dynamics*, p. 149. Cambridge.
- (24) F. Zwicky, 1939. *Proc. natn. Acad. Sci. U.S.A.*, **25**, 604. 1960. *Publs astr. Soc. Pacif.*, **72**, 365.