THE DYNAMICS OF SPHERICAL GALAXIES

Scott Tremaine
Canadian Institute for Theoretical Astrophysics
McLennan Labs, University of Toronto
60 St. George St., Toronto M5S 1A1
CANADA

ABSTRACT

The basic physics governing the structure of galaxies is described by the collisionless Boltzmann equation and the Jeans theorem. I review the derivation of these results and their application to spherical galaxies, and discuss a variety of techniques for constructing equilibrium models of spherical galaxies. I review general results on the stability of spherical galaxies and recent work on the radial-orbit instability. Finally, I discuss the statistical mechanics of violent relaxation and derive a heuristic distribution function that appears to describe the end-product of violent relaxation.

1. INTRODUCTION

One of the most important research areas of the last decade in galactic dynamics has been the study of triaxial galaxies, in which all three principal moments of inertia are different. These systems arise naturally in the collapse of non-equilibrium stellar systems and in mergers of stellar systems, and many, possibly most, real galaxies are likely to be triaxial to some extent. Thus it is proper to begin this lecture by describing why I think the study of spherical galaxies is still important. First, spherical galaxies are the simplest realistic stellar systems. Infinite slabs or cylinders are artificial because they do not approximate the global geometry of real galaxies, and disk systems have more complicated potential-density relations than spherical ones. Second, spherical systems are an important special case of triaxial systems, and it is difficult to understand the behavior of triaxial systems until we understand their limiting behavior in the spherical case. Finally, there are many major issues about the dynamics of spherical galaxies that are still unresolved.

The lecture will be divided into sections on equilibrium (§2), stability (§3), and formation (§4). As far as possible I will treat each topic independently. Much of the basic material is covered in a textbook that James Binney and I have recently completed[7)], and from which I have drawn heavily.

1.1 *Basic Equations*

The fundamental function that describes the distribution of stars in a galaxy is the distribution function (hereafter DF) $f(\mathbf{x},\mathbf{v},t)$, defined so that $f(\mathbf{x},\mathbf{v},t)d\mathbf{x}d\mathbf{v}$ is the mass contained at time t in the phase-space volume element $(\mathbf{x},\mathbf{v}) \to (\mathbf{x}+d\mathbf{x},\mathbf{v}+d\mathbf{v})$.[1] Of course, the DF can be defined for quantities other than mass, for example luminosity, number of stars, number of K giant stars, etc. Initially, I will define the DF in terms of luminosity since that is directly observable in external galaxies while mass and number of stars are not.

The crucial feature of galaxies that makes their dynamics relatively easy to understand is that there are no binary interactions between stars. The collision time for a typical star in the solar neighborhood is about 10^{17} years, far longer than the age of the Galaxy. Moreover, gravitational interactions between individual stars are so small that we can assume that each star's trajectory is determined solely by the large-scale potential field of the galaxy. The recognition of this fact makes it easy to derive the equation that describes the evolution of the DF.

The equation of continuity in fluid mechanics is

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_i}(\rho v_i) = 0, \tag{1}$$

[1] Notice that I define phase space to be position-velocity space rather than position-momentum space.

where I have used the summation convention. The DF represents the density in phase space and satisfies a continuity equation as well, since stars are neither created nor destroyed:

$$\frac{\partial f}{\partial t} + \frac{\partial}{\partial x_i}(f\dot{x}_i) + \frac{\partial}{\partial v_i}(f\dot{v}_i) = 0, \tag{2}$$

where the dot denotes a time derivative. We know that $\dot{x}_i = v_i$ and, since collisions and star-star gravitational interactions are neglected, $\dot{v}_i = -(\nabla\Phi)_i$, where $\Phi(\mathbf{x}, t)$ is the gravitational potential of the galaxy. Moreover $\partial v_i/\partial x_i = 0$ (since x_i and v_i are independent coordinates in phase space), and $\partial\Phi/\partial v_i = 0$ (since the gravitational potential depends only on position, not velocity). Hence

$$\frac{\partial f}{\partial t} + v_i\frac{\partial f}{\partial x_i} - \frac{\partial\Phi}{\partial x_i}\frac{\partial f}{\partial v_i} = 0. \tag{3}$$

This is the *Vlasov* or *collisionless Boltzmann* equation, and is the fundamental equation that governs the time evolution and structure of galaxies.

The collisionless Boltzmann equation can be rewritten in other forms. Recall that in three dimensions the Eulerian and Lagrangian derivatives are respectively

$$\left.\frac{\partial}{\partial t}\right|_{\mathbf{x}}, \qquad \frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla = \left.\frac{\partial}{\partial t}\right|_{\mathbf{x}} + \dot{x}_i\left.\frac{\partial}{\partial x_i}\right|_{t}. \tag{4}$$

The Eulerian derivative is the rate of change measured by an observer at a fixed location, while the Lagrangian derivative is the rate of change measured by an observer travelling with a fixed particle or fluid element. Similarly, in six-dimensional phase space we may define the Lagrangian derivative as the rate of change measured by an observer travelling through phase space with a given star,

$$\frac{D}{Dt} \equiv \left.\frac{\partial}{\partial t}\right|_{\mathbf{x},\mathbf{v}} + \dot{x}_i\left.\frac{\partial}{\partial x_i}\right|_{t,\mathbf{v}} + \dot{v}_i\left.\frac{\partial}{\partial v_i}\right|_{t,\mathbf{x}} = \frac{\partial}{\partial t} + v_i\frac{\partial}{\partial x_i} - \frac{\partial\Phi}{\partial x_i}\frac{\partial}{\partial v_i}. \tag{5}$$

Comparison with equation (3) shows that the collisionless Boltzmann equation can be written in the simple form

$$\frac{Df}{Dt} = 0. \tag{6}$$

This equation says that if an observer is attached to a given star travelling through the galaxy, the local phase-space density of stars that he measures

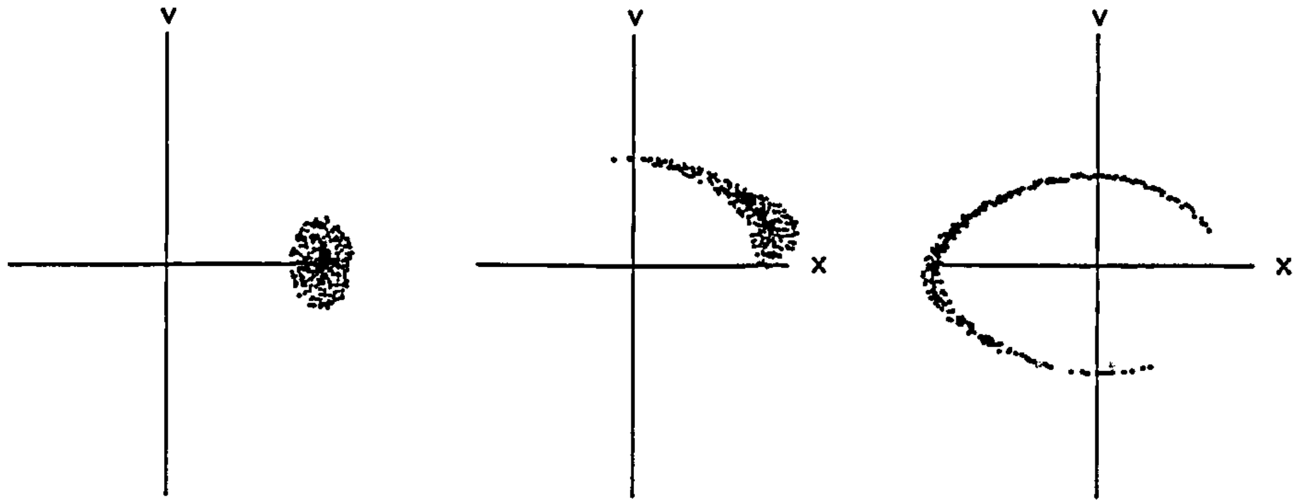

Figure 1. The evolution of the phase-space distribution of particles in the one-dimensional potential well $\Phi(x) = -(1+\sqrt{1+x^2})^{-1}$ (the *isochrone* potential). The initial DF is uniform in phase space in the region centered on $(x, v) = (1, 0)$ that is defined by $(x-1)^2 + v^2 < (0.2)^2$. The coordinate axes have unit length. The orbital period for a particle at $(x, v) = (1, 0)$ (the time required to return to its original position) is 16.67 time units. The three plots are at $t = 0$, $t = 15$, and $t = 60$.

around him remains constant. In other words, *the flow of stars in phase space is incompressible.*

A qualitative example of this kind of behavior is provided by a marathon race. If the runners travel at a constant speed, then the runners, like the stars in a galaxy, satisfy the collisionless Boltzmann equation. At the start of the race, the spatial density of runners is high, but they have a wide range of ability and hence travel with different speeds, so their phase-space density is not so high. At the finish line, the spatial density of runners is low, but at any particular time, the runners passing the finish line have nearly the same speed. Thus the phase-space density at the finish is the same as at the start, just as the collisionless Boltzmann equation predicts.

Another example is provided by a set of particles in a one-dimensional potential well (Figure 1). The particles are initially located in a small region with velocities near zero and phase-space density f_0, as in Figure 1(a). After their release, the particles fall down the potential well and oscillate back and forth. The particles with lower energy have shorter period, so the phase-space distribution begins to twist into a spiral, as in (b). As time goes on, the spiral becomes narrower and longer, as in (c). However, the phase-space density within the spiral remains constant, as the collisionless Boltzmann equation predicts.

After many orbits, the DF can become twisted into a very tightly wound spiral. In this case, if we measure the phase-space density with coarse resolution, we would find a lower phase-space density than f_0, because regions with $f = f_0$ and regions with $f = 0$ occupy the same resolution element. In other words, the so-called *coarse-grained* DF does *not* necessarily satisfy the collisionless Boltzmann equation. The coarse-grained DF $\bar{f}$ can be defined more precisely by dividing phase space into a set of *macrocells* and defining $\bar{f}$ to be the average value of the DF within each macrocell. The coarse-grained DF can be an important concept, because observations always have limited resolution and hence measure a coarse-grained DF rather than the "true" or *fine-grained* DF.

There is a simple application of these results to an interesting problem in cosmology. It has been suggested that neutrinos of non-zero rest mass are the particles that provide the dark mass in galaxies. In the early Universe, neutrinos are in thermal equilibrium and hence their DF has the Fermi-Dirac form

$$f_p = \frac{1}{h^3}\frac{1}{\exp(pc/kT)+1}. \tag{7}$$

Here h and k are Planck's constant and Boltzmann's constant, p is the momentum, and c is the speed of light.[2] After the Universe cools below about 1 MeV, the neutrinos are no longer in thermal equilibrium. Nevertheless, the DF (7) continues to apply, since both the momentum of the neutrinos and the temperature of the Universe fall at the same rate, proportional to $(1+z)^{-1}$, where z is the redshift[40)].

Once galaxies begin to form, the DF (7) is no longer valid. However, the neutrinos still satisfy the collisionless Boltzmann equation, so that $Df_p/Dt = 0$. Since the observable coarse-grained DF, which we may denote by $\bar{f}_p$, is the average over a macrocell of the fine-grained DF f_p, we may conclude that $\bar{f}_p$ can never exceed the maximum value of f_p, which is $\frac{1}{2}h^{-3}$ by equation (7). This argument can be used to set a lower limit to the neutrino rest mass if the neutrinos are indeed the main component of dark mass in galaxies[38)].

Professor Ruffini and I had a long discussion about the validity of this argument after his lecture. As I understand it, Professor Ruffini argues that since the collisionless Boltzmann equation is classical, it cannot be applied to situations in which the phase-space density of neutrinos is close to the limit set by the Pauli principle, h^{-3}. My contention, however, is that quantum

[2] In this discussion only, I will use a phase space based on position and momentum rather than position and velocity, and denote the corresponding DF by f_p. For non-relativistic particles of mass m, $f_p = f/m^3$.

mechanical effects are only important in this context if there are two-body interactions, and that since the neutrinos only interact with the mean gravitational field of the galaxy the classical collisionless Boltzmann equation provides an accurate description. Let me leave the resolution of this dispute as a homework exercise: Is the classical collisionless Boltzmann equation valid in galaxy-sized systems of non-interacting particles when the phase-space density is not small compared to the degenerate limit, h^{-3}? This problem has not been addressed in the literature, so correct answers to the homework problem should be publishable.

1.2 *Equations for Spherical Systems*

We now write the collisionless Boltzmann equation (3) in spherical coordinates (r, θ, ϕ). The simplest approach is to use the form $Df/Dt = 0$, equation (6), which can be written in spherical coordinates as

$$\frac{\partial f}{\partial t} + \dot{r}\frac{\partial f}{\partial r} + \dot{\theta}\frac{\partial f}{\partial \theta} + \dot{\phi}\frac{\partial f}{\partial \phi} + \dot{v}_r\frac{\partial f}{\partial v_r} + \dot{v}_\theta\frac{\partial f}{\partial v_\theta} + \dot{v}_\phi\frac{\partial f}{\partial v_\phi} = 0. \tag{8}$$

Here we have written the velocity in spherical coordinates as $\mathbf{v} = v_r\hat{\mathbf{e}}_r + v_\theta\hat{\mathbf{e}}_\theta + v_\phi\hat{\mathbf{e}}_\phi$. We eliminate $\dot{r}$, $\dot{\theta}$, and $\dot{\phi}$ using the relations $v_r = \dot{r}$, $v_\theta = r\dot{\theta}$, $v_\phi = r\sin\theta\dot{\phi}$, and use Newton's law in the form

$$\begin{aligned}
\dot{v}_r &= \frac{v_\theta^2 + v_\phi^2}{r} - \frac{\partial \Phi}{\partial r}, \\
\dot{v}_\theta &= -\frac{v_r v_\theta}{r} + \frac{v_\phi^2 \cot\theta}{r} - \frac{1}{r}\frac{\partial \Phi}{\partial \theta}, \\
\dot{v}_\phi &= -\frac{v_r v_\phi}{r} - \frac{v_\theta v_\phi \cot\theta}{r} - \frac{1}{r\sin\theta}\frac{\partial \Phi}{\partial \phi}.
\end{aligned} \tag{9}$$

Combining these results, we obtain the collisionless Boltzmann equation in spherical coordinates,

$$\begin{aligned}
&\frac{\partial f}{\partial t} + v_r\frac{\partial f}{\partial r} + \frac{v_\theta}{r}\frac{\partial f}{\partial \theta} + \frac{v_\phi}{r\sin\theta}\frac{\partial f}{\partial \phi} + \left(\frac{v_\theta^2 + v_\phi^2}{r} - \frac{\partial \Phi}{\partial r}\right)\frac{\partial f}{\partial v_r} \\
&+ \frac{1}{r}\left(v_\phi^2\cot\theta - v_r v_\theta - \frac{\partial \Phi}{\partial \theta}\right)\frac{\partial f}{\partial v_\theta} - \frac{1}{r}\left[v_\phi(v_r + v_\theta\cot\theta) + \frac{1}{\sin\theta}\frac{\partial \Phi}{\partial \phi}\right]\frac{\partial f}{\partial v_\phi} = 0.
\end{aligned} \tag{10}$$

In the simple case of spherical symmetry, where $\partial/\partial\theta = \partial/\partial\phi = 0$ and v_ϕ and v_θ enter only in the combination $v_\theta^2 + v_\phi^2 \equiv v_t^2$, we have

$$\frac{\partial f}{\partial t} + v_r \frac{\partial f}{\partial r} + \left(\frac{v_t^2}{r} - \frac{d\Phi}{dr}\right)\frac{\partial f}{\partial v_r} - \frac{v_r v_t}{r}\frac{\partial f}{\partial v_t} = 0. \tag{11}$$

The most we can hope to obtain from observations of a spherical galaxy is the surface brightness and distribution of line-of-sight velocities as a function of projected radial distance from the center of the galaxy. The line-of-sight velocity distribution is measured from the broadening of spectral lines that arises because different stars have different Doppler shifts. In fact, with present-day techniques the detailed velocity distribution cannot be measured accurately; the best we can do is to determine the second moment of the distribution, i.e., the root-mean-square velocity along the line of sight, $\sigma_p(R)$.[3] Therefore it is useful to relate the DF to σ_p and other observables.

The luminosity density or emissivity is

$$j(\mathbf{x}) = \int f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}. \tag{12}$$

The mean velocity $\overline{\mathbf{v}}$ is given by

$$j\overline{v}_i = \int v_i f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}, \tag{13}$$

and the velocity-dispersion tensor $\boldsymbol{\sigma}$ is given by

$$j\sigma_{ij} = \int (v_i - \overline{v}_i)(v_j - \overline{v}_j) f(\mathbf{x}, \mathbf{v}, t) d\mathbf{v}. \tag{14}$$

Of course $\boldsymbol{\sigma}$ is a symmetric tensor. If the galaxy has spherical symmetry then the emissivity depends only on the radial coordinate r. Furthermore, I shall henceforth assume that *all* properties of the galaxy are independent of direction; thus the mean velocities in the two angular directions must vanish,

$$\overline{v}_\theta = \overline{v}_\phi = 0, \tag{15}$$

and the velocity-dispersion tensor must satisfy

$$\sigma_{\theta\theta} = \sigma_{\phi\phi}, \qquad \sigma_{\theta\phi} = \sigma_{r\phi} = \sigma_{r\theta} = 0. \tag{16}$$

[3] I shall use r and R, respectively, to denote distances from the center of the galaxy in three dimensions and in the plane of the sky.

We also assume that the galaxy is in a steady state, so $\overline{v}_r = 0$.

To obtain an equation involving σ we simply multiply the collisionless Boltzmann equation (11) by v_r and integrate over velocity space, $\int d\mathbf{v} = 2\pi \int_{-\infty}^{\infty} dv_r \int_0^{\infty} v_t dv_t$. Since we assume a steady state, $\partial f/\partial t = 0$ and we have

$$\frac{d}{dr}\int d\mathbf{v} f v_r^2 + \int d\mathbf{v} \left(\frac{v_t^2}{r} - \frac{d\Phi}{dr}\right) v_r \frac{\partial f}{\partial v_r} - \frac{1}{r}\int d\mathbf{v} v_r^2 v_t \frac{\partial f}{\partial v_t} = 0. \tag{17}$$

The first integral is just $j\sigma_{rr}$. To evaluate the second we first carry out the integral over v_r, $\int dv_r v_r (\partial f/\partial v_r) = -\int dv_r f$, where we assume that $f \to 0$ as $|v_r| \to \infty$. To evaluate the third we first carry out the integral over v_t, $\int 2\pi v_t dv_t v_t (\partial f/\partial v_t) = -2\int 2\pi v_t dv_t f$, where we assume $f \to 0$ as $v_t \to \infty$. Thus equation (17) becomes

$$\frac{d}{dr}(j\sigma_{rr}) + \frac{2j}{r}(\sigma_{rr} - \sigma_{\theta\theta}) = -j\frac{d\Phi}{dr}. \tag{18}$$

Here we have used the relation $\int d\mathbf{v} v_t^2 f = 2j\sigma_{\theta\theta} = 2j\sigma_{\phi\phi}$. Equation (18) is sometimes called the equation of hydrostatic equilibrium since it is similar to the fluid equation,

$$\frac{dp}{dr} = -\rho\frac{d\Phi}{dr}, \tag{19}$$

where p is the pressure, equal to the mass density times the one-dimensional velocity dispersion. The main difference between equation (19) and equation (18) is that in the stellar system, in contrast to the fluid, the mean-square velocities in different directions need not be the same.

A stellar system in which $\sigma_{rr} = \sigma_{\theta\theta} = \sigma_{\phi\phi}$ is said to be *isotropic* since its velocity-dispersion tensor is isotropic, $\sigma_{ij} \equiv \sigma^2\delta_{ij}$. In this case equations (18) and (19) are identical if we replace j by ρ and $j\sigma^2$ by p. Another special case is the *shell model*, in which $\sigma_{rr} = 0$, that is, all the stars move on circular orbits. Shell models are mainly useful because they are easy to construct and analyze; it seems unlikely that real galaxies are well-approximated by shell models. A final special case is the *radial-orbit model*, in which $\sigma_{\theta\theta} = 0$; one disadvantage of radial-orbit models is that they always imply a singular density at the center ($j(r)$ diverges at least as fast as r^{-2} as $r \to 0$).

One general and useful result is the following: in any spherical, steady-state stellar system that has no singularities or divergences in the potential or DF, the ratio $\sigma_{rr}/\sigma_{\theta\theta} \to 1$ as $r \to 0$, i.e., the velocity-dispersion tensor is isotropic near the center. Doug Richstone and I convinced ourselves of this several years ago, but the argument was never published and the exact conditions on the potential and DF needed for central isotropy have never been worked out. Once again, I leave this as a (probably publishable) homework exercise.

In general, however, there is no reason to suppose that stellar systems are isotropic, except near their centers. In fact, most plausible models of galaxy formation suggest that the velocity-dispersion tensor is nearly radial ($\sigma_{rr} \gg \sigma_{\theta\theta}$) in the outer parts of the galaxy.

1.3 *The Jeans Theorem*

The most efficient way to solve the collisionless Boltzmann equation for spherical systems is through the use of integrals of motion. An *integral of motion* is any function $I(\mathbf{x}, \mathbf{v})$ of the phase-space coordinates such that

$$\frac{d}{dt} I[\mathbf{x}(t), \mathbf{v}(t)] = 0 \tag{20}$$

along the trajectory of a star $\mathbf{x}(t)$, $\mathbf{v}(t) = \dot{\mathbf{x}}(t)$. The simplest example of an integral is the energy per unit mass in a steady-state potential $\Phi(\mathbf{x})$, $E = \frac{1}{2}v^2 + \Phi(\mathbf{x})$. In a spherical potential the three components of the angular momentum per unit mass $\mathbf{J} = \mathbf{x} \times \mathbf{v}$ are also integrals. (From now on, for the sake of brevity I will drop the phrase "per unit mass" in using these expressions.) Notice that the expression on the left side of equation (20) is just DI/Dt where D/Dt denotes the Lagrangian derivative in phase space defined in equation (5).

The concept of an integral of motion leads immediately to the *Jeans theorem*, one of the central results in the study of galaxy dynamics: *Any* DF *that is a function only of integrals of motion is a solution of the steady-state collisionless Boltzmann equation.* The proof is very easy. Consider a DF of the form $f[I_1(\mathbf{x}, \mathbf{v}), I_2(\mathbf{x}, \mathbf{v}), \ldots, I_k(\mathbf{x}, \mathbf{v})]$, where $I_1, \ldots, I_k$ are k integrals of motion. We have

$$\frac{Df}{Dt} = \sum_{m=1}^{k} \frac{\partial f}{\partial I_m} \frac{DI_m}{Dt}. \tag{21}$$

But $DI_m/Dt = 0$ since I_m is an integral of motion; hence $Df/Dt = 0$, which means that f satisfies the collisionless Boltzmann equation in the form (6).[4]

[4] There is also a converse of the Jeans theorem; however it is more complicated to formulate and to prove and it will not be needed for our purposes[7].

According to the Jeans theorem, in a spherical system any DF of the form $f(E, \mathbf{J})$ is automatically a solution of the steady-state collisionless Boltzmann equation. If we assume once again that *all* properties of the galaxy are independent of direction, then f can depend on $\mathbf{J}$ only through its magnitude, and we may write

$$f(\mathbf{x}, \mathbf{v}) = f(E, J^2) = f(\tfrac{1}{2}v^2 + \Phi(r), r^2 v_t^2), \tag{22}$$

where as usual v_t is the tangential velocity component. It can also be shown that *all* solutions to the steady-state collisionless Boltzmann equation that are spherically symmetric in all their properties have the form (22). A particularly simple set of DFs are those that depend only on energy, $f = f(E)$. These are often called isotropic systems since the velocity-dispersion tensor is isotropic at all radii; however, this is somewhat misleading since there are isotropic systems in which the DF depends on both E and J^2 (see reference 10).

2. CONSTRUCTING EQUILIBRIUM SPHERICAL GALAXIES

The problem that I address in this section is how to construct a model galaxy, that is, how to find a DF $f(\mathbf{x}, \mathbf{v})$ that is consistent with all the observable properties of a given galaxy.

Of course, we do not know that any given galaxy is spherically symmetric; at most, we observe that the surface brightness distribution on the plane of the sky $I(\mathbf{x})$ is circularly symmetric and assume that the galaxy is spherical rather than, for example, oblate with its pole pointing towards us. A worthwhile exercise is to estimate how good this assumption is: for example, if the apparent axial ratio determined from the surface brightness is, say, in the range 0.9 to 1, what is the probability that the true axis ratio is, say, less than 0.5 (in which case the assumption that the galaxy was spherical would be badly in error)? This requires knowing the distribution of apparent axial ratios of galaxies and assuming that all galaxies are either oblate or prolate (see reference 28, §5-2).

With the assumption of spherical symmetry, the surface brightness is related to the emissivity by the integral equation[7)]

$$I(R) = 2\int_R^\infty \frac{j(r) r\, dr}{\sqrt{r^2 - R^2}}. \tag{23}$$

This an Abel-type integral equation which can be solved to yield

$$j(r) = -\frac{1}{\pi}\int_r^\infty \frac{dI}{dR}\frac{dR}{\sqrt{R^2 - r^2}}. \tag{24}$$

As I mentioned in §1.2, we can obtain a limited amount of information about the distribution of stars in velocity space from spectroscopic observations. About the best that can be done with present-day techniques is to determine the root-mean-square velocity along the line of sight, $\sigma_p(R)$. The line-of-sight dispersion velocity σ_p is related to the components of the velocity-dispersion tensor σ_{rr}, $\sigma_{\phi\phi} = \sigma_{\theta\theta}$ by

$$I(R)\sigma_p^2(R) = 2\int_R^\infty \frac{\sigma_{rr}(r^2 - R^2) + \sigma_{\theta\theta}R^2}{r^2}\frac{j(r)r\,dr}{\sqrt{r^2 - R^2}}. \tag{25}$$

In order to construct a dynamical model of the galaxy we must know the potential $\Phi(r)$. The potential can only be directly determined in rare cases, for example when the galaxy contains a circular disk of gas whose rotation curve can be measured. In general we are forced to make some plausible but arbitrary assumption about $\Phi(r)$. Perhaps the most common is that the mass-to-light ratio $\rho(r)/j(r) = \Upsilon(r)$ is independent of position. Then the potential can be obtained to within a constant multiplicative factor by solving Poisson's equation

$$\nabla^2\Phi = \frac{1}{r^2}\frac{d}{dr}r^2\frac{d\Phi}{dr} = 4\pi G\Upsilon j(r). \tag{26}$$

However, it should be remembered that there is very little direct evidence that this assumption is valid, especially in the outer parts of galaxies.

Thus, in the context of spherical galaxies, our goal may be stated as follows: for an assumed potential $\Phi(r)$, find the DF $f(E, J^2)$ that yields the observed surface brightness distribution $I(R)$ and the line-of-sight dispersion measurements $\sigma_p(R)$.

It is immediately evident that this problem is *underconstrained*: even if we are given the potential, and even if the line-of-sight dispersion is known at all radii, we cannot determine a function of two variables (the DF, which is a function of E and J^2) from two functions of one variable (the surface brightness and dispersion, which are functions of the radius). Thus we must decide on some prescription for choosing which of the many possible DFs that solve the problem is likely to provide the most realistic representation of the galaxy.

2.1 *The Eddington-Osipkov-Merritt Method*

The first approach to determining the DFs of spherical stellar systems was invented by Eddington[12)] and was generalized and greatly improved by

Osipkov[29] and Merritt[23]. The basic assumption of the Eddington-Osipkov-Merritt (hereafter EOM) method is that the DF is a function of a single variable

$$Q = E + \frac{J^2}{2r_a^2}, \tag{27}$$

with a free parameter r_a which is called the *anisotropy radius*. (Models with $r_a^2 < 0$ can also be constructed but these are generally less interesting.)

Using the definition of the emissivity (12) we have

$$\begin{aligned} j(r) &= \int f(E, J^2) d\mathbf{v} \\ &= 2\pi \int_{-\infty}^{\infty} dv_r \int_0^{\infty} v_t dv_t f\left(\tfrac{1}{2}v_r^2 + \tfrac{1}{2}v_t^2 + \Phi, J^2\right) \\ &= \frac{2\pi}{r^2} \int_{\Phi}^{\infty} dE \int_0^{2r^2(E-\Phi)} \frac{dJ^2 f(E, J^2)}{\sqrt{2(E-\Phi) - J^2/r^2}}. \end{aligned} \tag{28}$$

According to the EOM assumption, $f(E, J^2) = f(Q)$. Substituting Q for J^2 and changing the order of integration we obtain

$$j(r) = \frac{4\pi\sqrt{2}}{1 + \frac{r^2}{r_a^2}} \int_{\Phi}^{\infty} dQ f(Q) \sqrt{Q - \Phi}. \tag{29}$$

This is an Abel-type integral equation, which can be solved analytically to yield $f(Q)$.

The EOM method yields a solution for the DF for each assumed value of the anisotropy radius r_a. Of course, it may be that the resulting DF is negative in some region of phase space, in which case it is not a physically acceptable solution (stars emit light but do not absorb it!). For each physically acceptable anisotropy radius, we can compute the line-of-sight dispersion $\sigma_p(R)$ from the DF, and then choose the anisotropy radius that provides the best fit to the observed dispersions.

The physical meaning of the anisotropy radius is that for $r \lesssim r_a$ the velocity-dispersion tensor is approximately isotropic, while for $r \gg r_a$ the orbits are mainly radial. [In fact, it is straightforward to verify that for any EOM model $\sigma_{rr}(r) = (1 + r^2/r_a^2)\sigma_{\theta\theta}(r)$.]

The EOM method has many virtues. It is simple to apply and yields a straightforward one-parameter family of models for comparison with the data.

Moreover, it yields models that have an isotropic velocity-dispersion tensor in the central regions and nearly radial orbits in the outer parts, features that I argued in §1.2 were common to many plausible models of galaxy formation.

The main disadvantage of the EOM method is that it is based on a specific hypothesis [$f = f(Q)$] that may not apply to most real galaxies. It is also difficult to generalize to triaxial models.

The search for black holes in the centers of galaxies described by John Kormendy is one example of a problem for which the EOM method is ill-suited. The argument for the presence of a black hole in the center of a galaxy like M31 is that no physically plausible DF can match the observations of the emissivity distribution and the dispersion profile unless the potential includes a contribution from a central dark mass. This argument could not be made using the EOM method alone, since one could always argue that a more general form for the DF, without a central dark mass, would permit models to be constructed that agreed with the observations.

2.2 *Linear Programming*

The use of linear programming for constructing DFs of stellar systems was introduced in a seminal paper by Schwarzschild[33]. Schwarzschild was concerned with the much more difficult problem of constructing models of triaxial galaxies, where approaches like the EOM method are inapplicable since in general we do not even have explicit expressions for the integrals of motion. However, linear programming can also be applied to the simpler task of constructing spherical models[31].

In the context of spherical galaxies, Schwarzschild's method begins by dividing energy-angular momentum space up into K bins, centered at (E_i, J_i^2), $i = 1, \ldots, K$ and occupying the area $E_i - \frac{1}{2}\Delta E_i < E < E_i + \frac{1}{2}\Delta E_i$, $J_i^2 - \frac{1}{2}\Delta J_i^2 < J^2 < J_i^2 + \frac{1}{2}\Delta J_i^2$. The six-dimensional phase-space volume associated with bin i can be shown to be

$$\Delta V_i \equiv \int_{\text{bin } i} d\mathbf{x} d\mathbf{v} = 4\pi^2 T_r \Delta E_i \Delta J_i^2, \tag{30}$$

where $T_r(E, J^2)$ is the *radial period*, that is, the time required for a star to travel from apocenter to apocenter (see eq. 44). It is assumed that all the stars in bin i have energy E_i and angular momentum J_i; in other words, the stars are assumed to lie on only K distinct orbits (strictly, on K distinct sets of orbits, with random orientations but the same energy and angular momentum within each set). This is equivalent to the assumption that the DF has the form $f(E, J^2) = \sum_{i=1}^{K} f_i \delta(E - E_i)\delta(J^2 - J_i^2)$, where δ denotes the Dirac delta function. The set of values (E_i, J_i^2) is called the ***orbit library***. I will denote the mass of stars in orbit i by w_i, the ***orbit weight***; the relation between w_i and f_i is

$$w_i = \Upsilon f_i \Delta V_i, \tag{31}$$

where Υ is the mass-to-light ratio.

The next step is to divide the galaxy into discrete bins in radius, with bin j containing the radii $r_{j-1} < r < r_j$, $j = 1, \ldots, N$. Here $r_0 \equiv 0$. The luminosity contained in bin j is

$$\Delta L_j = 4\pi \int_{r_{j-1}}^{r_j} r^2 dr\, j(r). \tag{32}$$

Instead of trying to construct a DF that reproduces the emissivity distribution $j(r)$ exactly, we only attempt to reproduce the luminosities ΔL_j.

Assuming that the potential $\Phi(r)$ is known, it is straightforward to compute the quantity m_{ij}, the fractional time that orbit i spends in bin j. A galaxy model consists of the K orbit weights w_i since these specify the distribution of stars in phase space. The equation that we must solve to determine the $\{w_i\}$ is

$$\Delta L_j = \sum_{i=1}^{K} m_{ij} w_i, \qquad j = 1, \ldots, N. \tag{33}$$

There are several features of this equation that deserve to be explicitly stated:

(i) Since the DF must be non-negative, we must have $w_i \geq 0$ in any physically acceptable model.

(ii) Measurements of the line-of-sight dispersion $\sigma_p(R)$ also provide constraints that are linear functions of the w_i; hence fitting the dispersion profile is no different in principle from fitting the surface brightness profile.

(iii) The number of orbits in the orbit library, K, must generally be much larger than the number of radial bins, N, since two-dimensional (E, J^2) space must be sampled by more points than one-dimensional r space. Thus the solutions $\{w_i\}$ to the set of linear equations (33) are not unique; this corresponds to a property mentioned earlier, that the DF of a spherical galaxy is underconstrained.

(iv) Any solution $\{w_i\}$ represents an *exact* solution of the collisionless Boltzmann equation, since the DF is a set of Dirac delta functions at specific values of E and J^2 and hence satisfies the collisionless Boltzmann equation by the Jeans theorem. The only approximation involved is that the solution only reproduces the luminosity emitted in each radial bin, ΔL_j, rather than the exact luminosity density $j(r)$. This is not a serious limitation, since the bin sizes can easily be made much smaller than the resolution of our telescopes. However, this solution for the DF, though exact, is unrealistic, since it consists of a set of spikes in (E, J^2) space whereas the real DF probably varies smoothly. A better solution could be obtained by dividing (E, J^2) space into cells and replacing single orbits by bundles of orbits that generate a constant DF within each cell, but I will neglect this improvement in this lecture for the sake of simplicity.

We now use linear programming to choose one solution $\{w_i\}$ from the many possible solutions of equation (33). Linear programming is a mathematical technique for determining the maximum of a linear function of K variables (the *profit function*), when the range of the variables is restricted by linear equalities or inequalities. In our case the variables are the w_i's, and the range of the w_i's is restricted by the N equalities of equation (33) and the K inequalities $w_i \geq 0$, $i = 1, \ldots, K$.

The term "profit function" arises, of course, because linear programming is most commonly used in business, economics, and other areas where limited (and non-negative!) resources must be allocated in such a way as to maximize profit. One remarkable and very appealing feature of linear programming is that by using an iterative technique known as the *simplex method* the solution can usually be found in a number of iterations that is only of order K.

To construct a galaxy model using linear programming, we must choose some linear function of the orbit weights to be the profit function, P. There is no unique "best" profit function, and the choice of profit function must reflect our preconceptions about galactic structure. For example, if we believe that galaxies consist of stars on near-radial orbits, we may choose

$$P = -\sum_{i=1}^{K} J_i^2 w_i, \qquad (34)$$

which minimizes the mean-square angular momentum of the stars. On the other hand, to construct a model in which the stars are on nearly circular orbits, we may choose

$$P = +\sum_{i=1}^{K} J_i^2 w_i, \qquad (35)$$

which maximizes the mean-square angular momentum. It is not hard to construct models with an isotropic velocity-dispersion tensor[31)] or even to construct EOM models with an appropriate choice of the profit function. The profit function is flexible enough to permit us to examine a wide range of models consistent with the data.

As an example of the use of linear programming, in Figure 2 I show the velocity-dispersion profiles $\sigma_p(R)$ for three models with surface-brightness profiles that satisfy de Vaucouleurs' empirical law,

$$I(R) = I_0 \exp[-(R/R_e)^{0.25}]. \tag{36}$$

In these models the mass-to-light ratio is assumed to be independent of radius, and the velocity dispersion is measured in units of $(GM/R_e)^{1/2}$, where M is the total mass. In general, models with near-radial orbits have line-of-sight dispersions that are lower in the outer parts and higher near the center than models with near-circular orbits. In principle, therefore, fitting the dispersion profile to observations could provide information on the orbit distribution; however, the effects of the orbit distribution are difficult to disentangle from a possible radial dependence of the mass-to-light ratio.

Geometrically, the set of orbit weights consistent with the data can be viewed as a $K - N$ dimensional polygon in the K dimensional space whose coordinate axes are the orbit weights. The orbit weights that maximize the profit function lie at a vertex of the polygon. This is both a virtue and a defect: on one hand, linear programming is useful because it generates extreme models—for example, the most nearly radial orbits or the maximum mass-to-light ratio—and provides an efficient way to explore the limits of the range of consistent models; on the other hand, it seems more likely that real galaxies lie somewhere in the middle of the range of possible models, and therefore linear programming cannot easily produce a single "most likely" model.

2.3 *The Maximum-Entropy Method*

One plausible way to produce a "most likely" model consistent with the observational constraints is to maximize the entropy,

$$S = -\int d\mathbf{x} d\mathbf{v} f \log f. \tag{37}$$

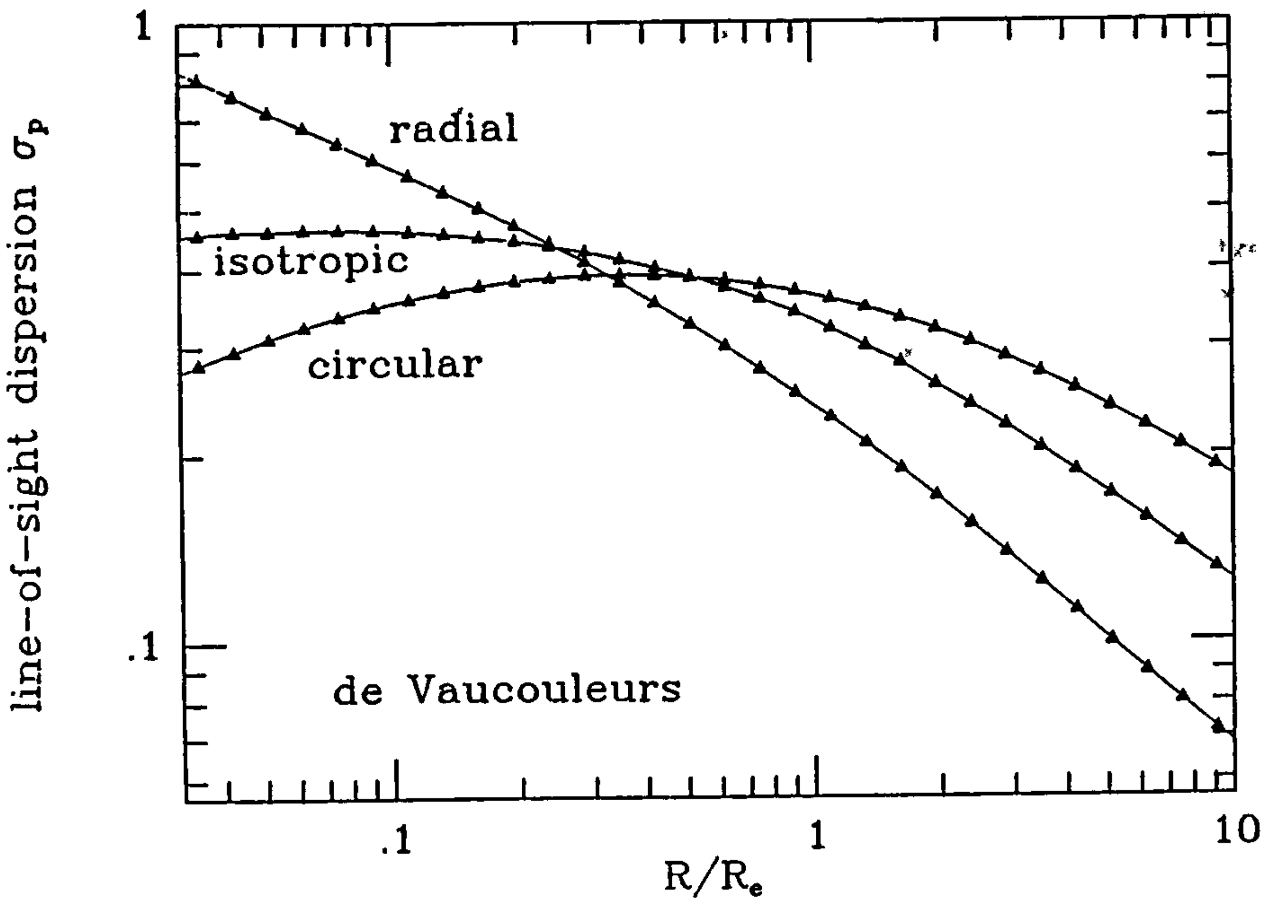

Figure 2. The root-mean-square velocity dispersion along the line of sight as a function of radius, for a galaxy with constant mass-to-light ratio and de Vaucouleurs' surface-brightness profile (eq. 36). The dispersion is measured in units of $(GM/R_e)^{1/2}$. The curves are for models with near-radial orbits, circular orbits, and isotropic velocity-dispersion tensor (from reference 31).

(I have suppressed an irrelevant factor of Boltzmann's constant.) A stellar system is not necessarily in a maximum-entropy state because it is not in thermodynamic equilibrium (relaxation to thermodynamic equilibrium is mainly due to gravitational interactions between individual stars and takes about 10^{18} yr, far longer than the typical age of 10^{10} yr). In fact, it can be shown that there is no maximum-entropy state of an isolated self-gravitating system of fixed mass and energy, because the entropy can increase without limit (see §4). Nevertheless, given the observational constraints on the emissivity profile and dispersion profile, and an assumed form for the potential, we can ask what is the maximum-entropy state consistent with these constraints, and there are several reasons why the maximum-entropy model—in this limited sense—provides the most natural model to represent the real galaxy. First, entropy increases in any relaxation process (see §4 for a more detailed discussion), and hence in the absence of other considerations, models with large entropy are more plausible than models with small entropy. Second, models with large entropy tend to have smooth DFs, which seems to be a desirable feature. Finally, the determination of the DF from observations is essentially a deconvolution process, and maximum-entropy methods have proved to be extremely efficient in many deconvolution problems in radio astronomy, image processing, etc.

To apply the maximum-entropy method to galaxies, we can use the same set of K bins in energy-angular momentum space and N radial bins that we used in linear programming. The entropy (37) may be written

$$S = -\sum_{i=1}^{K} w_i \log(w_i/\Delta V_i), \tag{38}$$

where ΔV_i is the phase-space volume associated with bin i (eq. 30), and I have dropped an unimportant multiplicative factor of the mass-to-light ratio Υ. A more general form is

$$S = -\sum_{i=1}^{K} C_i(w_i), \tag{39}$$

where $C_i(w)$ is any convex function, that is, a function with $d^2C/dw^2 > 0$. In this more general form, it is not guaranteed that the entropy increases in a relaxation process, but it is still usually true that smooth DFs correspond to large values of the entropy.

Thus, we simply instruct the program to maximize S from equation (39) subject to the constraints (33), plus any constraints arising from velocity-dispersion measurements (see reference 32 for details). We do not explicitly impose the constraints $w_i \geq 0$; these constraints tend to be satisfied automatically since the maximum-entropy solution is "smooth" and hence avoids negative weights. It is straightforward to check that the orbit weights are non-negative after the solution been obtained.

As an example, I show in Figure 3 the maximum-entropy galaxy with an emissivity profile given by Plummer's law,

$$j(r) = \frac{j_0}{(1+r^2)^{5/2}}. \tag{40}$$

One shortcoming of the maximum-entropy method is evident from Figure 3: the radial and tangential velocity dispersions are roughly equal at all radii,

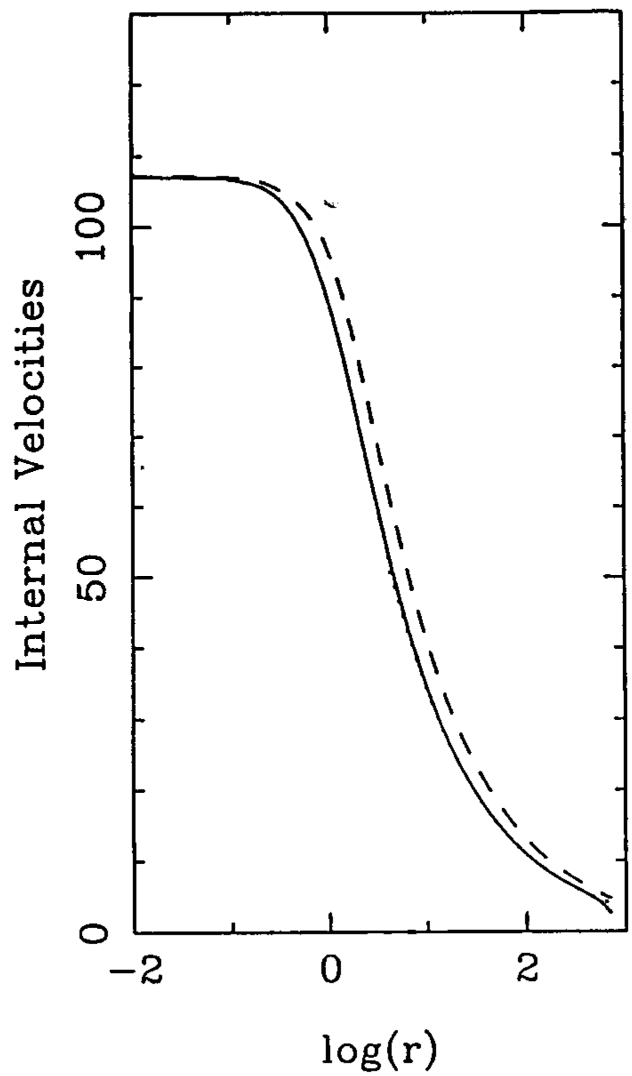

Figure 3. The maximum-entropy galaxy with a brightness profile given by Plummer's law (40). The model is normalized so that the central line-of-sight dispersion is 100 units. The solid and dashed lines show $\sqrt{\sigma_{rr}}$ and $\sqrt{\sigma_{\theta\theta}}$ respectively (from reference 32).

while in fact we expect that the orbits should be predominantly radial at large radii. This problem raises the interesting issue of how the maximum-entropy method can be modified to find the "most likely" model consistent with both the observations and our beliefs about the structure of galaxies. An interesting first step in this direction has been made by Dejonghe[9)].

2.4 *Future Prospects*

In the last few years we have developed a flexible and efficient set of tools for constructing galaxy models. In the study of stellar structure, the existence of similar tools for the numerical construction of stellar models led to explosively rapid progress. In the study of galaxies, the prospects are less hopeful, for two main reasons. First, there is no analog of the Hertzsprung-Russell diagram, which showed strong and distinctive correlations between many properties of stars (e.g., color and luminosity); in the case of galaxies it is much less clear which observational features demand explanation. Second, models of stars are unique once the mass and chemical composition are given (the Vogt-Russell theorem), while many different models of a given spherical galaxy are likely to be consistent with even the best available data.

Despite this uncertain future, there are a number of promising avenues of research in this area:

(i) Richstone and I have found that very careful programming is needed to produce reliable dispersion profiles in models based on a discrete set of orbits. The reason is basically that each individual orbit has a square-root singularity in its contribution to the emissivity at its turning points (apocenter and pericenter), and the dispersion profiles depend on the gradient of the emissivity (eq. 18). Further work is needed to produce accurate, flexible codes that can be used as "black boxes" by observers who wish to construct galaxy models.

(ii) The construction of axisymmetric or triaxial galaxy models often requires only a straightforward generalization of the methods used for spherical galaxies (indeed, linear programming was first used to construct a triaxial galaxy[33)]). However, the problems mentioned in item (i) become even more severe, and we are far from being able to produce good velocity-dispersion maps for such systems by any of the methods described above. (An alternative approach due to Bishop[8)] does yield smooth dispersion maps at some cost in generality.)

(iii) What other methods are available for constructing galaxy models besides linear programming and maximum entropy? Bishop's approach[8)] has already been mentioned; an iterative algorithm devised by Lucy[18)] and applied by Newton and Binney[27)] and Statler[34)] is efficient but produces a model whose properties depend in a complicated way on the starting point of the iteration.

(iv) In a certain sense the main problem we face at present is that constructing DFs of spherical galaxies is too easy—so many different models match the data that we don't know what to do next! We are badly in need of other constraints, arising perhaps from stability arguments or from a better understanding of the formation process, to limit the range of models that we must consider.

3. STABILITY OF SPHERICAL GALAXIES

Somewhat surprisingly, the study of the stability of spherical stellar systems is still in a primitive state. Much more effort has been expended on studying the stability properties of disklike stellar systems than of spheres. Thus, there are many good prospects for important research in this field.

So far I have generally let the DF represent the density of luminosity in phase space. In this section the DF will represent the density of mass, since we are now interested in dynamics rather than observations. Hence Poisson's equation takes the form

$$\nabla^2\Phi = 4\pi G \int f(\mathbf{x},\mathbf{v})d\mathbf{v}. \tag{41}$$

To analyze stability, one must linearize the collisionless Boltzmann equation around an equilibrium spherical initial state and then solve the linearized collisionless Boltzmann equation and the Poisson equation, which together determine the perturbed DF $f_1(\mathbf{x},\mathbf{v},t)$, the perturbed density $\rho_1(\mathbf{x},t)$, and the perturbed potential $\Phi_1(\mathbf{x},t)$. It can be shown (most clearly in reference 5) that all such solutions can be written as a sum of functions of the form

$$\Phi_1(\mathbf{x},t) = \Phi_{\ell m}(r)\mathrm{Y}_{\ell m}(\theta,\phi)e^{-i\omega t}, \tag{42}$$

where $\mathrm{Y}_{\ell m}(\theta,\phi)$ is a spherical harmonic. The squared eigenfrequency ω^2 is not necessarily real; that is, spherical systems can be either overstable [$\mathrm{Im}(\omega) > 0$, $\mathrm{Re}(\omega) \neq 0$] or unstable [$\mathrm{Im}(\omega) > 0$, $\mathrm{Re}(\omega) = 0$].

The eigenfunction $\Phi_{\ell m}(r)$ is determined by an integrodifferential equation that is straightforward to derive (see, e.g., reference 30), but as yet no general code has been written to find the solutions. I expect and hope that someone will write the required code within the next year or two. Until this is done we must be content with a number of results that have been derived for special cases.

The most powerful results have been derived in the case where the DF depends only on energy and not angular momentum, $f = f(E)$. Many of these results were derived by the Soviet physicist V. A. Antonov in two seminal papers in the early 1960's[1,2]. In particular:

(i) *A spherical stellar system with* DF $f(E)$ *and* $df/dE < 0$ *is stable if the barotropic gaseous system with the same unperturbed density distribution is stable.* Here a barotrope is a gas in which the pressure is a function only of the density, $p = p(\rho)$. This result permits many known stability results for stars to be carried over to galaxies.

(ii) *A spherical stellar system with* DF $f(E)$ *and* $df/dE < 0$ *is stable to all non-radial perturbations, that is, to perturbations with* $\ell > 0$.

(iii) *A spherical stellar system with* DF $f(E)$ *and* $df/dE < 0$ *is stable to radial perturbations if its unperturbed density* $\rho_0(r)$ *and potential* $\Phi_0(r)$ *satisfy the inequality*

$$\frac{d^3\rho_0}{d\Phi_0^3} \leq 0. \tag{43}$$

Together, these three theorems are powerful enough to prove stability for almost all spherical systems with DF $f = f(E)$ and $df/dE < 0$ that are plausible models for real galaxies.

An even more powerful theorem was later proved by Doremus, Feix and Baumann[11,17,35]:

Any spherical stellar system with DF $f(E)$ *and* $df/dE < 0$ *is stable.*

Unfortunately the proof of this remarkably general theorem is very difficult to follow. There is also some concern because the same group has claimed that anisotropic systems with DF $f(E, J^2)$ satisfying $\partial f/\partial E < 0$, $\partial f/\partial J^2 < 0$ are also stable, and this proof is contradicted by numerical experiments[26]. Thus it would be a worthwhile exercise to check the proof of this theorem and to determine whether there is a simpler and more physically appealing approach than the ones in the literature.

The stability of spherical shell models ($\sigma_{rr} = 0$) is still largely unexplored. Here the analysis is much simpler, because eigenfunctions are determined by an ordinary second-order differential equation rather than by an integrodifferential equation; nevertheless, few general results are known. In the case where the density falls as r^{-2}, the shell model is stable[36]. Some N-body simulations of other shell models have been carried out and show no sign of instability.

The stability of models composed of stars on radial orbits ($\sigma_{\theta\theta} = \sigma_{\phi\phi} = 0$ was first explored by Antonov[3]. Antonov showed that all such models were unstable. This result suggests that models composed of stars on near-radial orbits may also be unstable. This interesting possibility, although recognized fifteen years ago, has only been subject to intense investigation in the last three or four years, largely because (i) Antonov's result was published in Russian in an obscure conference proceedings; (ii) analytic work[14], which is now believed to be incorrect, implied that most interesting systems composed of stars on near-radial orbits were stable. The Soviet work on what has come to be called the *radial-orbit instability* only became widely known in the West with the publication of an English translation of a two-volume work

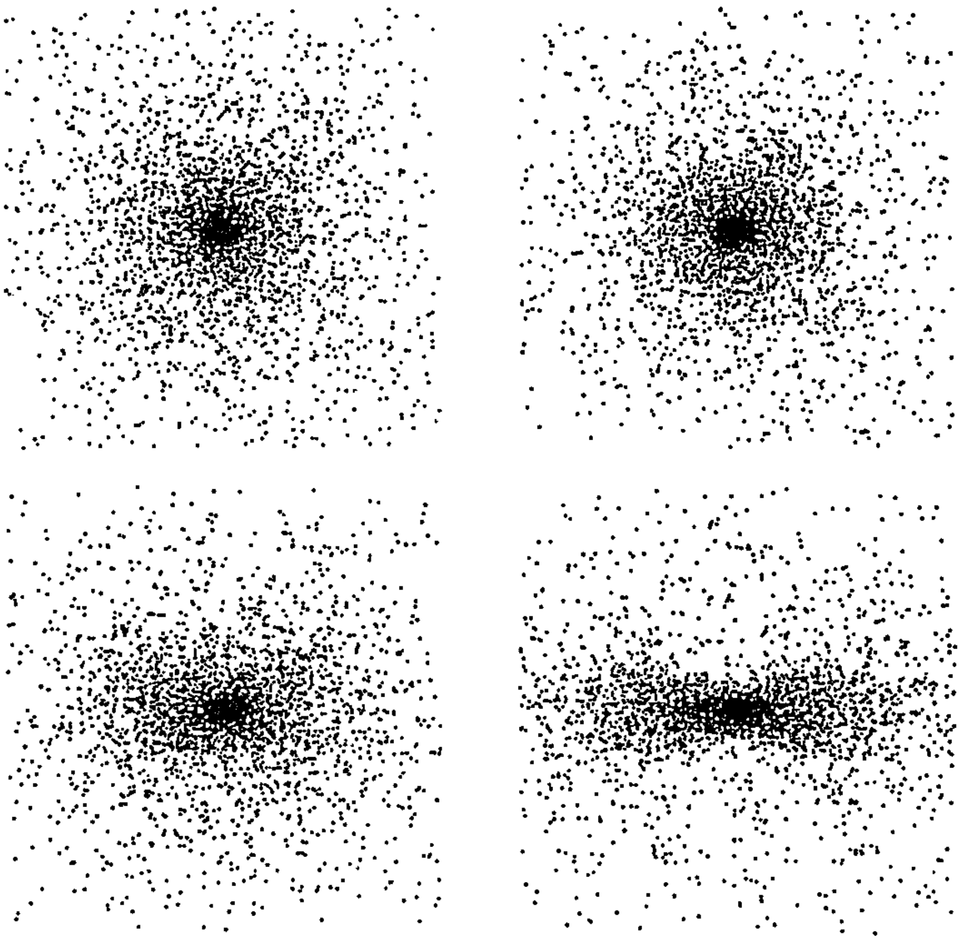

Figure 4. Development of the radial-orbit instability in an initially spherical stellar system with near-radial orbits (from reference 25).

on the equilibrium and stability of collisionless stellar systems by Fridman and Polyachenko[13]. At about the same time, Barnes[4] independently carried out N-body experiments that demonstrated the existence of an instability in spherical systems composed of stars on near-radial orbits. These developments stimulated more extensive numerical experiments that are described in papers by Merritt and Aguilar[26] and Merritt[25]. Figure 4 is taken from the latter paper and shows the evolution of a typical system that is subject to the radial-orbit instability. The instability deforms the spherical system into a triaxial one.

The physical origin of the radial-orbit instability has been discussed by a number of authors. Let me first give two simple but incorrect explanations: (i) Since the tangential dispersions $\sigma_{\theta\theta} = \sigma_{\phi\phi} = 0$ in a galaxy composed of stars on radial orbits, it is sometimes argued that the galaxy is Jeans unstable to short-wavelength perturbations in the tangential directions. This argument is misleading because the growth time for the Jeans instability is of order $(G\rho)^{-1/2}$ where ρ is the *local* density, while the orbital period is of order $(G\overline{\rho})^{-1/2}$, where $\overline{\rho}$ is the *mean* density interior to the point in question. Since in general $\overline{\rho} > \rho$, the orbital period is shorter than the growth time, so that the approximation on which the Jeans instability is based, that the surroundings are static and homogeneous, is false. (ii) It is also sometimes argued that radial orbits have no "restoring force" that prevents their orientations from drifting, and hence that it is natural for self-gravity to lead to clumping of neighboring orbits. This argument neglects the fact that orbits are not rigid bodies; the response of a radial orbit to a torque is to gain angular momentum and become less radial, not necessarily to shift its orientation.

The correct explanation is due to Palmer and Papaloizou[30]; an earlier paper by Lynden-Bell[20] gives essentially the same explanation in a different context. An orbit in a spherical potential has two periods, the *radial period* T_r and the *azimuthal period* T_a. The radial period is the time required to travel from apocenter to pericenter and back, while the azimuthal period is defined so that $2\pi/T_a$ is the mean angular speed. These periods are given by

$$
\begin{aligned}
T_r &= 2\int_p^a \frac{dr}{[2(E-\Phi)-J^2/r^2]^{1/2}},\\
\frac{T_r}{T_a} &= \frac{J}{\pi}\int_p^a \frac{dr}{r^2[2(E-\Phi)-J^2/r^2]^{1/2}}.
\end{aligned}
\tag{44}
$$

Here J is the angular momentum, p is the pericenter, and a is the apocenter. An orbit is closed if T_r and T_a are commensurable; for example, orbits in a Keplerian potential have $T_r = T_a$ and are ellipses with one focus at $r = 0$, while orbits in a quadratic potential $\Phi(r) = Kr^2$ have $T_r = \frac{1}{2}T_a$ and are ellipses centered on $r = 0$. Another example of a closed orbit is a radial ($J = 0$) orbit in a potential that is non-singular near $r = 0$; these orbits have $T_r = \frac{1}{2}T_a$ are are simply straight lines through $r = 0$.

Near-radial orbits in a potential that is non-singular near $r = 0$ can be approximated as nearly straight lines whose orientation slowly drifts or *precesses* (Figure 5). It is not hard to see from equations (44) that the orientation angle ψ of a near-radial orbit precesses at a rate

$$\dot{\psi} = \frac{2\pi}{T_a}\left(1 - \frac{T_a}{2T_r}\right). \tag{45}$$

(Note that ψ is measured in the plane normal to **J** and increases in the direction of orbital motion.) The term in brackets is zero for radial orbits and

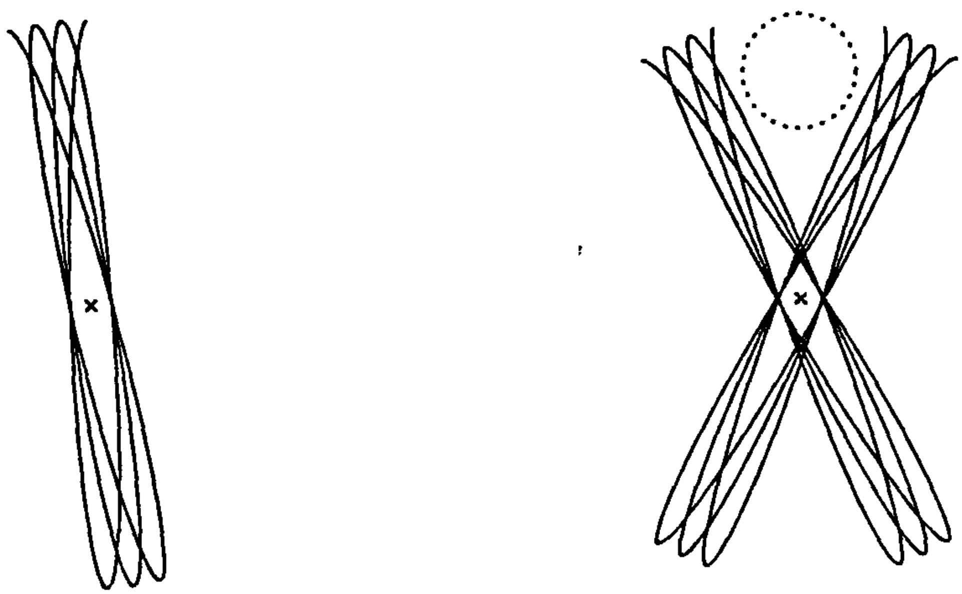

Figure 5. Near-radial orbits in a potential that is smooth near $r = 0$. (a) A typical orbit can be approximated as a highly elongated ellipse that precesses slowly. The center of the galaxy is marked by an $\times$. (b) A mass concentration (dotted lines) exerts a torque on near-radial orbits on either side.

increases as the angular momentum of the orbit increases; in fact it can be shown that

$$\left(1 - \frac{T_a}{2T_r}\right) = Jh(E) + O(J^2), \tag{46}$$

where

$$h(E) = \frac{2}{\pi} \lim_{r_1 \to 0} \left(\int_{r_1}^{r_2} \frac{dr}{r^2\sqrt{2(E-\Phi)}} - \int_{r_1}^{\infty} \frac{dr}{r^2\sqrt{2E}} \right), \tag{47}$$

$\Phi(r_2) = E$, and I have assumed that $\Phi(r = 0) = 0$.

Now consider orbits in a given plane with polar coordinates (r, ϕ), where ϕ increases in the direction of orbital motion. Suppose that a mass concentration develops slowly along the line $\phi = 0$. The energy of an orbit does not change, because the perturbation is nearly stationary; however, the angular momentum changes because the perturbation is non-axisymmetric. The concentration exerts a positive torque on near-radial orbits with $-\frac{1}{2}\pi < \psi < 0$ and a negative torque on orbits with $0 < \psi < \frac{1}{2}\pi$ (see Figure 5). If $h(E) > 0$ then the positive torque increases both J and $(1-\frac{1}{2}T_a/T_r)$ when $-\frac{1}{2}\pi < \psi < 0$ so these orbits precess more rapidly towards the mass concentration. Similarly, the precession rate $\dot{\psi}$ of orbits with $0 < \psi < \frac{1}{2}\pi$ is slowed. In this case near-radial orbits can be trapped to oscillate with their long axes aligned with the mass concentration, thus enhancing its effects and leading to instability. On the other hand, for $h(E) < 0$ near-radial orbits can only be trapped at right angles to the mass concentration, where their gravity suppresses the disturbance. (Lynden-Bell[20] calls orbits with $h(E) < 0$ "donkey" orbits, since they resist a pull by twisting in the opposite direction.) In most potentials that behave smoothly near $r = 0$, $h(E) > 0$ so that instability can arise if near-radial orbits make a sufficiently large contribution to the overall potential.[5]

The radial-orbit instability is one of the most interesting aspects of the dynamics of spherical galaxies to arise in recent years. Moreover, it seems likely that the instability sets significant constraints on the structure of real galaxies. For example, Merritt[24] has shown that one of the best available models of the giant elliptical galaxy M87, constructed by Newton and Binney[27], is unstable; and Dejonghe and Merritt[10] have argued that spherical galaxies obeying Plummer's law (eq. 40) that are constructed using the EOM method are unstable unless the anisotropy radius $r_a > 1.1$.

4. DISSIPATIONLESS FORMATION OF GALAXIES

In this part of the lecture I consider only a very limited model problem connected with galaxy formation. In particular, I will assume that galaxy formation is *dissipationless*, that is, that stars form from the intergalactic gas *before* the galaxy forms, so that galaxy formation is a problem of collisionless dynamics rather than gas dynamics.

Let us imagine an irregular distribution of stars with small random velocities. If the stars are released at $t = 0$, they will fall together due to their mutual self-gravity, and the distribution will collapse, undergoing a complicated motion which eventually leads to the formation of a stationary, smooth, final distribution of stars that turns out to strongly resemble many real galaxies. During the collapse, many of the stars experience a rapidly changing gravitational potential, which causes their energies and angular momenta to change in an irregular manner. This process redistributes the energies and angular momenta of the stars, in much the same way that two-body collisions

[5] The criterion $h(E) > 0$ is a special case of Lynden-Bell's criterion, which applies to the more general situation where the orbits are not necessarily radial and the potential perturbation is stationary in a rotating frame. In this case the so-called *fast action* rather than the energy is conserved.

redistribute energy and angular momentum in a gas. Lynden-Bell[19] suggested that this redistribution, which he called ***violent relaxation***, could lead to a unique final state that was largely independent of the initial conditions. Subsequent numerical calculations have verified this suggestion, at least so long as the initial state is sufficiently irregular and its random velocities are sufficiently small[41,42,22].

What I want to discuss here is whether there are simple and general arguments that determine the nature of this unique final state.

As a preliminary step, it is useful to examine the evolution of the entropy during violent relaxation. Differentiating equation (37) and using the collisionless Boltzmann equation (3), we have

$$\begin{aligned}\frac{dS}{dt} &= -\int d\mathbf{x}d\mathbf{v}(1+\log f)\frac{\partial f}{\partial t}\\ &= \int d\mathbf{x}d\mathbf{v}(1+\log f)\left[v_i\frac{\partial f}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial f}{\partial v_i}\right]\\ &= \int d\mathbf{x}d\mathbf{v}\left[v_i\frac{\partial (f\log f)}{\partial x_i} - \frac{\partial \Phi}{\partial x_i}\frac{\partial (f\log f)}{\partial v_i}\right].\end{aligned} \tag{48}$$

The first and second terms in the square brackets on the last line can be integrated over x_i and v_i and yield zero if $f \to 0$ as $|\mathbf{x}|, |\mathbf{v}| \to \infty$. Hence

$$\frac{dS}{dt} = 0, \tag{49}$$

and the entropy is conserved. This is in contrast to systems like the ideal gas, in which particles interact via short-range forces, where the entropy is known to increase by the Boltzmann H-theorem.

For collisionless systems, it turns out to be more useful to define the entropy in terms of the coarse-grained DF $\overline{f}$ defined in §1.1. Thus I write

$$S_c = -\int d\mathbf{x}d\mathbf{v}\overline{f}\log\overline{f}. \tag{50}$$

With this definition of the entropy, there is an analog of the Boltzmann H-theorem, which can be stated as follows: *Assume that at the initial time $t = 0$, the DF has no fine-scale structure (or, what is equivalent, choose the macrocells small enough so that the DF is nearly constant within each macrocell at $t = 0$). Then at any other time t we have $S_c(t) \geq S_c(t = 0)$.* I shall call this the *H-theorem for galaxies*.

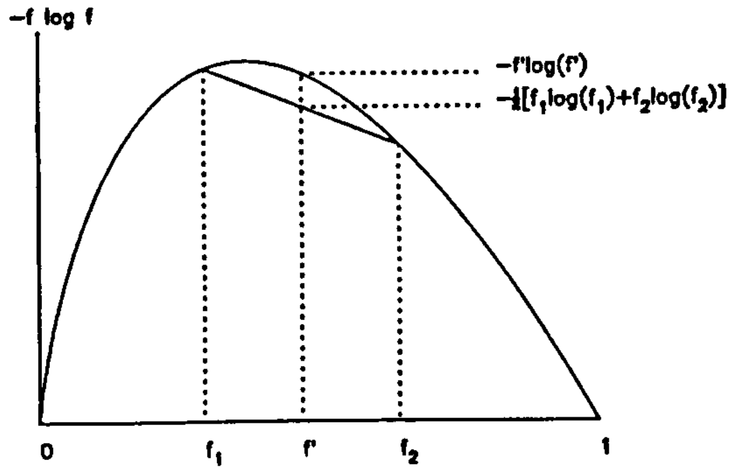

Figure 6. An analogy to the H-theorem for galaxies: If equal volumes of paint with colors f_1 and f_2 are mixed to give paint with color $f' = \frac{1}{2}(f_1 + f_2)$, the selling price $-f \log f$ goes up.

The proof is given in a number of places[39], and rather than repeat it here I will describe an analogy. Suppose that I own a paint store but that I have only red and white paint in stock. A customer comes in and wants to buy pink paint. I can make pink paint by mixing red and white paint; to be quantitative I assign the value $f = 0$ to white paint, $f = 1$ to red paint, and $f = \lambda$ to paint made by mixing red and white paint with relative volumes λ and $1 - \lambda$ (of course, strictly f is a "coarse-grained" color, since on a sufficiently fine scale the red paint globules and the white paint globules are still distinct). The customer is willing to purchase paint of color f at a rate per unit volume $-f \log f$, while the cost per unit volume of red and white paint is the same. How can I mix up the red and white paint to maximize my profit?

Consider mixing unit volumes of paint of colors f_1 and f_2. The final color will be $f' = \frac{1}{2}(f_1 + f_2)$. The selling price per unit volume before mixing is $-\frac{1}{2}(f_1 \log f_1 + f_2 \log f_2)$, and the price after mixing is $-f' \log f'$. It is easy to see from Figure 6 that the selling price and hence the profit has increased as a result of the mixing.

It is clear that *any* possible mixing strategy increases my profit, no matter what the relative volumes and initial colors of the paints may be. Thus we can formulate a kind of Boltzmann H-theorem for paint, which says that the profit can never decrease as a result of mixing. This may seem like a curious story, but I hope that the analogy with the H-theorem for galaxies is clear. The selling price is the entropy and the color is the coarse-grained DF ; both paint and the phase-space fluid are incompressible so that the effects of mixing are very similar.

There are important differences between the H-theorem for galaxies and the usual Boltzmann H-theorem[9,16,21]:

(i) The H-theorem for galaxies does not say that the entropy can never decrease; it only says that the entropy at all times other than $t = 0$ is larger than the entropy at $t = 0$. Thus the entropy could increase for a while and then decrease, so long as it never drops below its initial value.

(ii) There is no arrow of time in the H-theorem for galaxies. Thus, the entropy is larger than the entropy at $t = 0$ not only for all times $t > 0$ but also for all times $t < 0$. The arrow of time is only defined implicitly, in that $t = 0$ is the time when the DF is smooth on small scales, and this is more natural for an initial state than a final state.

(iii) The function $-\int d\mathbf{x}d\mathbf{v}\overline{f}\log\overline{f}$ is not the only function for which an H-theorem can be proved. In fact, the theorem has equal validity if $\overline{f}\log\overline{f}$ is replaced by *any* convex function $C(\overline{f})$ (a convex function is one for which $d^2C(x)/dx^2 \geq 0$). Thus the entropy, equation (50), loses much of the central role that it plays in conventional statistical mechanics, and we can state a generalized version of the H-theorem for galaxies: *Assume that at the initial time $t = 0$, the* DF *has no fine-scale structure. Then at any other time t we have $H_C(t) \geq H_C(t = 0)$, where*

$$H_C = -\int d\mathbf{x}d\mathbf{v}C(\overline{f}), \tag{51}$$

and C is any convex function.

Having established that there is an H-theorem for galaxies, it is natural to ask whether the final state after violent relaxation is a maximum-entropy state [i.e. a maximum of equation (50) for fixed mass and energy]. The answer is no, since the following simple argument due to James Binney shows that there is *no* maximum-entropy state. Take a small fraction ϵM, $\epsilon \ll 1$, of the mass M of the galaxy and place it in an extended halo of typical radius r. The typical velocity of material in this halo is $v \approx \sqrt{GM/r}$. Thus the typical DF in the halo is $\overline{f} \approx \epsilon M/(r^3v^3)$. The contribution to the entropy is $-\epsilon M\log\overline{f} \approx \epsilon M[\log M - 3\log(rv)] \approx \epsilon M(-\frac{1}{2}\log M - \frac{3}{2}\log G - \frac{3}{2}\log r)$. Now let $r \to \infty$. The change in energy is $\approx G\epsilon M/r$, which tends to zero, but the contribution of the halo to the entropy diverges; hence there is no maximum-entropy state.

This result is consistent with the generalized H-theorem for galaxies, which tells us that all H-functions of the form (51) are non-decreasing during violent relaxation, and hence that there is no obvious reason why the entropy—which is just one of many H-functions—is maximized in the final state. This lack of a unique final state shows that statistical mechanics is less powerful in its application to galaxies than to gases.

Although the generalized H-theorem does not predict a unique final state for violent relaxation, it does provide a strong inequality constraint on the final state, since every H-function must be larger in the final state than the initial state. We say that a DF $\overline{f}'(\mathbf{x},\mathbf{v})$ is *more mixed* than a DF $\overline{f}(\mathbf{x},\mathbf{v})$ with the same total mass if $H_C[\overline{f}'] \geq H_C[\overline{f}]$ for all convex functions C; thus, violent relaxation leads to a more mixed DF.

The most convenient characterization of the constraint provided by the generalized H-theorem is provide by the *mixing theorem*, which may be stated as follows. Let $V(\phi)$ be the volume of phase space in which the coarse-grained phase-space density exceeds ϕ,

$$V(\phi) = \int d\mathbf{x}d\mathbf{v}\Theta[\overline{f}(\mathbf{x},\mathbf{v}) - \phi], \tag{52}$$

where $\Theta(x)$ is the step function, $\Theta(x) = 1$ for $x \geq 0$, $\Theta(x) = 0$ for $x < 0$. Similarly let $M(\phi)$ be the mass contained in the volume $V(\phi)$,

$$M(\phi) = \int d\mathbf{x}d\mathbf{v}\overline{f}(\mathbf{x},\mathbf{v})\Theta[\overline{f}(\mathbf{x},\mathbf{v}) - \phi]. \tag{53}$$

Eliminating ϕ between equations (52) and (53), we obtain a function $M(V)$. The mixing theorem states: *The* DF $\overline{f}'$ *is more mixed than the* DF $\overline{f}$ *if and only if* $M'(V) \leq M(V)$ *for all* V. (See reference 39 for a proof.)

Notice that the slope of the $M(V)$ curve at $V = 0$ is simply the maximum value of the coarse-grained DF, $\overline{f}_{\max}$. Thus a corollary of the mixing theorem is that $\overline{f}'_{\max} \leq \overline{f}_{\max}$, a constraint that we have used already in §1.1.

The mixing theorem is the strongest general constraint of which I am aware on the final state after violent relaxation. This constraint is not strong enough to specify the final state uniquely. Therefore I think that it is unlikely that general arguments from statistical mechanics provide a useful approach to determining the outcome of violent relaxation. Instead, I think it is more fruitful to examine the physics of violent relaxation, to search for an crude description of the relaxation process to guide the construction of a heuristic DF.

In particular, violent relaxation redistributes the energies and angular momenta of stars through rapid potential fluctuations in the central region of the collapsing system. These fluctuations only affect stars that pass through the central region and hence cannot give stars angular momenta that are so large that their pericenters lie outside the central region. This argument suggests that the DF should have the general form

$$f \propto (\text{function of energy}) \times \begin{pmatrix} \text{cutoff in peri-} \\ \text{center or } J^2 \end{pmatrix}. \tag{54}$$

Furthermore, the number of stars per unit energy should be smooth near $E = 0$, since the potential fluctuations redistribute stars over a wide range of energy and there is no physical reason why the location $E = 0$ should be special. Since the phase-space volume corresponding to unit energy interval is proportional to the radial period T_r (eq. 30), we expect that[15,37]

$$f \propto \frac{1}{T_r(E, J^2)} \times \begin{pmatrix} \text{cutoff in peri-} \\ \text{center or } J^2 \end{pmatrix} \times \begin{pmatrix} \text{smooth function} \\ \text{of } E \text{ which is} \\ \text{non-zero at } E = 0 \end{pmatrix}. \tag{55}$$

This prescription is still pretty vague, but I believe that it contains much of the physics of violent relaxation, and it would be worthwhile to compare some families of functions of this kind with numerical simulations of violent relaxation. One simple functional form consistent with the criterion (55) is the Bertin-Stiavelli[6] DF ,

$$f(E, J^2) = K|E|^{3/2} \exp(-\beta E - \alpha J^2). \tag{56}$$

Once the mass and energy are fixed, the Bertin-Stiavelli DF has one free parameter, which would reflect the size of the region in which the relaxation occurs.

This approach to violent relaxation suggests many interesting research possibilities. How sensitive is the DF to the exact choice of functional form, so long as the constraint (55) is satisfied? Are the resulting DFs subject to the radial-orbit instability? Can these arguments be made more accurate or rigorous to give us more clues about the best analytic choice for a family of DFs? To what extent are these concepts valid when mergers or infall contribute to the formation process?

I am grateful to Zhou Zhi-hong for preparing a careful and accurate transcription of my lecture, from which this account was prepared.

References

1. Antonov, V. A., "Remarks on the problem of stability in stellar dynamics". *Astr. Zh.* **37**, 918–926 (1960). Translated in *Sov. Astr.-AJ*, **4**, 859–867 (1961).

2. Antonov, V. A., "Solution of the problem of stability of a stellar system with Emden's density law and a spherical distribution of velocities". *Vest. Leningrad Univ.* **19**, 96–111 (1962). Translated in de Zeeuw, T., ed., *Structure and Dynamics of Elliptical Galaxies*, IAU Symposium No. 127, Dordrecht: Reidel, 531–548 (1987).

3. Antonov, V. A., "On the instability of stationary spherical models with purely radial motion", in Omarov, T. B., ed., *Dynamics of Galaxies and Star Clusters* (1972). Translated in de Zeeuw, T., ed., *Structure and Dynamics of Elliptical Galaxies*, IAU Symposium No. 127, Dordrecht: Reidel, 549–552 (1987)

4. Barnes, J., "Dynamical instabilities in spherical stellar systems", in Goodman, J., and Hut, P., eds., *Dynamics of Star Clusters*, IAU Symposium No. 113, Dordrecht: Reidel, 297–299 (1985).

5. Barnes, J., Goodman, J., and Hut, P., "Dynamical instabilities in spherical stellar systems". *Astrophys. J.* **300**, 112–131 (1986).

6. Bertin, G., and Stiavelli, M., "Stellar dynamical models of elliptical systems". *Astron. Astrophys.* **137**, 26–28 (1984).

7. Binney, J., and Tremaine, S., *Galactic Dynamics*, Princeton: Princeton University Press (1987).

8. Bishop, J. L., "Models of axisymmetric galaxies". *Astrophys. J.* **305**, 14–27 (1986).

9. Dejonghe, H., "On entropy and stellar systems". *Astrophys. J.* **320**, 477–481 (1987).

10. Dejonghe, H., and Merritt, D., "Radial and nonradial stability of spherical stellar systems". *Astrophys. J.* xx, xx (1988).

11. Doremus, J.-P., Feix, M. R., and Baumann, G., "Stability of encounterless spherical stellar systems". *Phys. Rev. Lett.* **26**, 725–728 (1971).

12. Eddington, A. S., "The distribution of stars in globular clusters". *Mon. Not. Roy. Astron. Soc.* **76**, 572–585 (1916).

13. Fridman, A. M., and Polyachenko, V. L., *Physics of gravitating systems*, 2 vols., Springer-Verlag: New York (1984).

14. Gillon, D., Doremus, J. P., and Baumann, G., "Stability of self-gravitating spherical systems with phase space density a function of energy and angular momentum for aspherical modes". *Astr. Ap.* **48**, 467–474 (1976).

15. Jaffe, W., "The envelopes of spherical galaxies", in de Zeeuw, T., ed., *Structure and Dynamics of Elliptical Galaxies*, IAU Symposium No. 127, Dordrecht: Reidel, 511–512 (1987).

16. Kandrup, H. E., "An H-theorem for violent relaxation". *Mon. Not. Roy. Astron. Soc.* **225**, 995–998 (1987).

17. Kandrup, H., and Sygnet, J. F., "A simple proof of dynamical stability for a class of spherical clusters". *Astrophys. J.* **298**, 27–33 (1985).

18. Lucy, L. B., "An iterative technique for the rectification of observed distributions". *Astron. J.* **79**, 745–754 (1974).

19. Lynden-Bell, D., "Statistical mechanics of violent relaxation in stellar systems". *Mon. Not. Roy. Astron. Soc.* **136**, 101–121 (1967).

20. Lynden-Bell, D., "On a mechanism that structures galaxies". *Mon. Not. Roy. Astron. Soc.* **187**, 101–107 (1979).

21. Mathur, S. D., "Irreversibility due to mixing in collisionless systems". *Mon. Not. Roy. Astron. Soc.* **xx**, xx (1988).

22. McGlynn, T. A., "Dissipationless collapse of galaxies and initial conditions". *Astrophys. J.* **281**, 13–30 (1984).

23. Merritt, D., "Spherical stellar systems with spheroidal velocity distributions". *Astron. J.* **90**, 1027–1037 (1985).

24. Merritt, D., "Dynamical instability of a model of M87". *Astrophys. J.* **319**, 55–60 (1987).

25. Merritt, D., "Stability of elliptical galaxies. Numerical experiments", in de Zeeuw, T., ed., *Structure and Dynamics of Elliptical Galaxies*, IAU Symposium No. 127, Dordrecht: Reidel, 315–329 (1987).

26. Merritt, D., and Aguilar, L., "A new study of the stability of spherical galaxies". *Mon. Not. Roy. Astron. Soc.* **217**, 787–804 (1985).

27. Newton, A. J., and Binney, J., "Constructing distribution functions for spherical galaxies". *Mon. Not. Roy. Astron. Soc.* **210**, 711–730 (1984).

28. Mihalas, D., and Binney, J., *Galactic Astronomy*, San Francisco: W. H. Freeman (1981).

29. Osipkov, L. P., "Spherical systems of gravitating bodies with an ellipsoidal velocity distribution". *Pis'ma Astr. Zh.* **5**, 77–80 (1985).

30. Palmer, P. L., and Papaloizou, J., "Instability in spherical stellar systems". *Mon. Not. Roy. Astron. Soc.* **224**, 1043–1053 (1987).

31. Richstone, D. O., and Tremaine, S., "A general method for constructing spherical galaxy models". *Astrophys. J.* **286**, 27–37 (1984).

32. Richstone, D. O., and Tremaine, S., "Maximum-entropy models of galaxies". *Astrophys. J.* **xx**, xx (1988).

33. Schwarzschild, M., "A numerical model for a triaxial stellar system in dynamical equilibrium". *Astrophys. J.* **232**, 236–247 (1979).

34. Statler, T. S., "Self-consistent models of perfect triaxial galaxies". *Astrophys. J.* **321**, 113–152 (1987).

35. Sygnet, J. F., Des Forets, G., Lachieze-Rey, M., and Pellat, R., "Stability of gravitational systems and gravothermal catastrophe in astrophysics". *Astrophys. J.* **276**, 737–745 (1984).

36. Synakh, V. S., Fridman, A. M., and Shukhman, I. G., "Resonant damping of oscillations in a model globular cluster". *Astrofizika* **8**, 577–585 (1972).

37. Tremaine, S., "Summary", in de Zeeuw, T., ed., *Structure and Dynamics of Elliptical Galaxies*, IAU Symposium No. 127, Dordrecht: Reidel, 367–373 (1987).

38. Tremaine, S., and Gunn, J. E., "Dynamical role of light neutral leptons in cosmology". *Phys. Rev. Lett.* **42**, 407–410 (1979).

39. Tremaine, S., Hénon, M., and Lynden-Bell, D., "H-functions and mixing in violent relaxation". *Mon. Not. Roy. Astron. Soc.* **219**, 285–297 (1985).

40. Turner, M. S., "A cosmologist's tour through the new particle zoo", in Kormendy, J., and Knapp, G. R., eds., *Dark Matter in the Universe*, IAU Symposium No. 117, Dordrecht: Reidel, 445–488 (1987).

41. van Albada, T. S., "Dissipationless galaxy formation and the $r^{1/4}$ law". *Mon. Not. Roy. Astron. Soc.* **201**, 939–955 (1982).

42. Villumsen, J. V., "Violent relaxation and dissipationless collapse". *Astrophys. J.* **284**, 75–89 (1984).