DARK MATTER IN GALAXIES AND GALAXY SYSTEMS

Scott Tremaine and Hyung Mok Lee
Canadian Institute for Theoretical Astrophysics
University of Toronto
Toronto, Canada

ABSTRACT

We review the dynamical evidence for dark matter in galaxies, groups of galaxies and clusters of galaxies. A summary, expressed in terms of the mass-to-light ratios of various systems, is presented in Figure 3 near the end of the review.

I. INTRODUCTION

In astronomy, most information regarding the presence of different kinds of mass comes from photons at various wavelengths. Very hot gases emit X-rays, while stars produce most of their energy at optical wavelengths. Some atomic or molecular gases in interstellar space show emission lines at radio wavelengths (e.g., HI, CO, etc.). In addition, there are non-luminous objects whose existence can be inferred from other considerations. For example, interstellar dust grains are known to exist because of interstellar reddening, and the number of non-luminous stellar remnants (e.g., black holes, neutron stars or white dwarfs) can often be estimated from stellar population and stellar evolution theory.

Sometimes, certain masses manifest themselves only through gravitational interaction. We will use the term *dark matter* (abbreviated as DM) to denote matter whose existence is inferred only through its gravitational effects. Therefore the best—indeed only—way of studying dark matter is to accurately determine the mass of astronomical objects from their dynamics and to compare this mass with the mass inferred from the light emitted by the objects. A discrepancy indicates the presence of dark matter.

The determination of dynamical mass, however, is not a trivial task. The main difficulty is that even perfect observations (i.e., observations without error) cannot always provide enough information to completely constrain theoretical models. This is because we always observe projected positions on the plane of the sky and line-of-sight velocities at a given instant rather than complete three-dimensional orbits over an extended period of time. The subject of these lectures is the methods of determining dynamical masses of galaxies and systems of galaxies. The plan of the lectures is as follows: the basic theoretical framework is discussed in the next lecture. The subsequent four lectures describe the determination of masses of various systems. The final lecture provides a summary. Additional details and other topics are provided in many review articles, conference proceedings, and textbooks,

including Binney and Tremaine (1987), Faber and Gallagher (1979), Kormendy and Knapp (1987), Primack (1987), and Trimble (1987).

We would like first to discuss some introductory subjects that do not fit easily into the main arguments in the other lectures.

1 Virial theorem

One of the simplest ways to determine the mass of a stellar system is through the virial theorem.

Consider a self-gravitating system composed of N point masses. We denote by m_i, $\vec{r}_i$, and $\vec{v}_i$ the mass, position and velocity of each particle. The moment of inertia of such a system is

$$I = \sum_{i=1}^{N} m_i \vec{r}_i{}^2. \tag{1.1}$$

Now take first and second time derivatives of the moment of inertia to get

$$\dot{I} = 2\sum_{i=1}^{N} m_i \vec{r}_i \cdot \vec{v}_i \tag{1.2}$$

and

$$\ddot{I} = 2\sum_{i=1}^{N} m_i (\vec{v}_i{}^2 + \vec{r}_i \cdot \ddot{\vec{r}}_i). \tag{1.3}$$

Since the system is self-gravitating, the acceleration of particle i can be computed by summing over the contribution from all the particles

$$\ddot{\vec{r}}_i = \sum_{j \neq i} G m_j \frac{(\vec{r}_j - \vec{r}_i)}{|\vec{r}_j - \vec{r}_i|^3}, \tag{1.4}$$

so that equation (1.3) can be rewritten

$$\tfrac{1}{2}\ddot{I} = 2K + \sum_{i=1}^{N}\sum_{j \neq i} G m_i m_j \frac{\vec{r}_i \cdot (\vec{r}_j - \vec{r}_i)}{|\vec{r}_j - \vec{r}_i|^3} = 2K - \frac{G}{2}\sum_{i=1}^{N}\sum_{j \neq i} \frac{m_i m_j}{|\vec{r}_i - \vec{r}_j|} = 2K + W, \tag{1.5}$$

where

$$K = \tfrac{1}{2}\sum_{i=1}^{N} m_i \vec{v}_i{}^2, \tag{1.6}$$

is the total kinetic energy and

$$W = -\frac{G}{2}\sum_{i=1}^{N}\sum_{j \neq i} \frac{m_i m_j}{|\vec{r}_i - \vec{r}_j|}, \tag{1.7}$$

is the total gravitational potential energy of the system.

If the system is in a stationary state, we may assume that $\langle \ddot{I} \rangle_t = 0$, where $\langle \rangle_t$ is the time average taken over several times the dynamical timescale of the system. The dynamical timescale is the time required for typical particles to complete one orbit, that is, to move across the whole system. In practice, we observe at only one epoch. Thus, although the rigorous form of the virial theorem is $2\langle K \rangle_t + \langle W \rangle_t = 0$, in practice the theorem is used without the time average as follows:

$$2K + W \approx 0. \tag{1.8}$$

Equation (1.8) is valid only if the system is in equilibrium and in a stationary state. Such a condition is achieved if the age of the system is much longer than its dynamical timescale.

The virial theorem formed the basis for most mass determinations of galaxies and galaxy systems for decades, and although it has now largely been superseded by more sophisticated methods, it is still used to provide rough mass estimates for systems such as groups of galaxies with few members where detailed modelling is inappropriate.

2 History of dark matter

The discovery of dark matter can be clearly traced to a seminal paper by Zwicky (1933). By the early 1930's, Hubble's law relating radial velocities v of distant galaxies to their distances d through

$$v = H_0 d, \tag{1.9}$$

was established, although the Hubble constant H_0 was very poorly determined. The above relation holds for galaxies with $v \gtrsim 2000\,\mathrm{km\,s^{-1}}$. We will parametrize the Hubble constant as

$$H_0 = 100h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}, \tag{1.10}$$

where $1\,\mathrm{Mpc} = 1$ megaparsec $= 3.086 \times 10^{24}$ cm and h is a dimensionless parameter. In the early 1930's, h was known to be 5.58, while it is currently believed to lie between 0.5 and 1.

At the time of Zwicky's paper, rotation curves were available for some spiral galaxies so that the mass-to-light ratios of galaxies were determined to be

$$\frac{M}{L} \approx 1h \left(\frac{\mathrm{M}_\odot}{\mathrm{L}_\odot} \right)_V. \tag{1.11}$$

The subscript V in the above equation indicates that the luminosity is measured in the visual band. The unit for mass-to-light ratios adopted in the above equation, i.e., $(\mathrm{M}_\odot/\mathrm{L}_\odot)_V$, will be used throughout these lectures, and we will usually omit the units in expressing M/L from here on.

For $h = 5.58$, the mass-to-light ratio in equation (1.11) was consistent with observations of the Solar neighbourhood, which gave $M/L \approx 3$, not far from the Solar value. This agreement, together with the fact that galaxy spectra resembled the Solar spectrum, led to the simple and compelling picture that galaxies are made up of stars like the ones around us and that the Sun is a typical star in a galaxy.

Zwicky applied the virial theorem to the Coma cluster of galaxies. Radial velocities were known for seven galaxies in Coma at that time. The mean and root mean square velocities Zwicky used were

$$\overline{v_\parallel} \approx 7300\,\mathrm{km\,s^{-1}}; \quad v_{\parallel,\mathrm{rms}} = \left[\frac{1}{N-1}\sum_{i=1}^{N}(v_{\parallel,i} - \overline{v_\parallel})^2\right]^{1/2} \approx 700\,\mathrm{km\,s^{-1}}. \quad (1.12)$$

For spherical systems, the gravitational potential energy can be expressed in the form

$$W = -\alpha\frac{GM^2}{R}, \quad (1.13)$$

where M and R are the total mass and radius of the system, respectively, and α is a constant depending on the density distribution. For a uniform sphere $\alpha = 3/5$, while $\alpha = 3/(5-n)$ for a polytropic sphere with polytropic index n. Note that this constant is not very sensitive to the specific form of the density distribution, which was poorly known at that time; thus the value for the uniform sphere ($\alpha = 3/5$) was used by Zwicky. If the cluster is spherical and the galaxies have the same mass, the total kinetic energy can be computed from the formula $K = \frac{3}{2}Mv^2_{\parallel,\mathrm{rms}}$. After plugging these numbers into equation (1.8), he obtained the total mass of the Coma cluster

$$M \approx \frac{5v^2_{\parallel,\mathrm{rms}}R}{G} \approx 1\times10^{15}h^{-1}\,\mathrm{M_\odot}, \quad (1.14)$$

where he took the size of the cluster to be $1.5^\circ \approx 2h^{-1}\,\mathrm{Mpc}$. The total luminosity of the cluster was known to be

$$L_V \approx 2\times10^{13}h^{-2}\,\mathrm{L_\odot}, \quad (1.15)$$

which gives the mass-to-light ratio $M/L \approx 50h$ in Solar units. This is about a factor of 50 larger than the M/L's of individual galaxies [eq. (1.11)], whatever the value of h may be. Therefore he concluded that a large amount of non-luminous matter is required if the Coma cluster is in dynamical equilibrium.

By modern standards, Zwicky's analysis has several problems, including crude estimates for the cluster radius, luminosity, and density distribution, poor statistics due to the small number of radial velocities, and possible contamination from background or foreground galaxies. Nevertheless his principal result has survived. The mass-to-light ratios for both galaxies and clusters of galaxies have gone up, but the discrepancy of a factor of 50 found by Zwicky still remains. More detailed analyses show that the best present estimate of M/L for the Coma cluster is about $400h$, as will be discussed in greater detail in lecture 6.

Zwicky's remarkable result implies that at least 95% of the mass in Coma is in some invisible form. It suggests that on scales larger than $\approx 1\,\mathrm{Mpc}$, visible stars represent only a minor contaminant in a vast sea of dark matter of completely unknown nature.

3 A quick review of cosmology

One of the most important assumptions in standard cosmological models is that the Universe is isotropic and homogeneous on large scales (say, between $30h^{-1}$ and

$3000h^{-1}$ Mpc). This means that there exists a set of "fundamental observers" for whom the Universe looks isotropic, and a cosmic time such that Universe is homogeneous for all fundamental observers at any fixed time. Any observer living in a galaxy is a fundamental observer to a good approximation. Of course, this is not the only available cosmological model, but one of the simplest and the most widely used one, and the assumption of large-scale homogeneity and isotropy is consistent with observations of galaxy counts and the microwave background. [This part of the lecture is mainly based on Gunn (1978).]

Suppose $l_{ij}(t)$ denotes the distance between two fundamental observers i and j, which can be written

$$l_{ij}(t) = l_{ij}(t_0)R(t), \tag{1.16}$$

where t_0 is the present time and $R(t)$ is the scale factor, which depends only on t because the Universe is isotropic and homogeneous. Note that $R(t_0) = 1$. The subscript 0 in the above equation represents the value at the present epoch. The relative velocity between i and j is simply the time derivative of l_{ij}

$$v_{ij}(t) = l_{ij}(t_0)\dot{R}(t) = \frac{\dot{R}(t)}{R(t)} l_{ij}(t) = H(t)l_{ij}(t), \tag{1.17}$$

which is equivalent to Hubble's law of expansion with the Hubble constant being

$$H(t) = \frac{\dot{R}(t)}{R(t)}; \quad H_0 = H(t_0). \tag{1.18}$$

Next, consider a non-relativistic free particle passing a fundamental observer at point A with a velocity v_p at time t and passing an observer at point B at a later time $t + dt$. If the separation between the points A and B is dl, the relationship between dt and dl is

$$dl = v_p dt. \tag{1.19}$$

The velocity seen at the point B is

$$v_p' = v_p - H(t)dl, \tag{1.20}$$

so that

$$\frac{dv_p}{dt} = -H(t)v_p, \tag{1.21}$$

which can be integrated to give the relation $v_p \propto 1/R(t)$. For the case of relativistic particles like photons, a similar argument gives the relationship between the frequency at which the photon is emitted and the frequency observed by us,

$$\frac{\nu_0}{\nu_e} = \frac{R(t_e)}{R(t_0)} \equiv \frac{1}{1+z}, \tag{1.22}$$

where the subscript e represents the time of emission and z is the redshift of the photon.

Now let us consider dynamics. We will use Birkhoff's theorem of general relativity, which states that in a spherical system the acceleration at any radius

depends only on the mass distribution within that radius. Thus if we consider two galaxies separated by a distance $l(t)$, when l is sufficiently small their relative acceleration is described by the Newtonian formula

$$\frac{d^2l}{dt^2} = -\frac{GM}{l^2}, \tag{1.23}$$

where M is the mass inside a sphere of radius l. If the Universe is matter-dominated, then M is constant in time [since the Hubble flow is smooth, galaxies neither enter nor leave the comoving sphere of radius $l(t)$], and equation (1.23) can be integrated over t to yield

$$\tfrac{1}{2}\dot{l}^2 - \frac{GM}{l} = E, \tag{1.24}$$

where

$$M = \frac{4}{3}\pi\rho(t)l^3 \quad ; \quad l = l_0 R(t), \tag{1.25}$$

and E is the integration constant, which is equivalent to binding energy in Newtonian dynamics. Dividing through by $\frac{1}{2}l_0^2$, we get

$$\dot{R}^2 - \frac{8}{3}\pi G\rho R^2 = \epsilon, \tag{1.26}$$

where $\epsilon = 2E/l_0^2$. At the present epoch, equation (1.26) can be written

$$H_0^2 - \frac{8}{3}\pi G\rho_0 = \epsilon, \tag{1.27}$$

or

$$1 - \Omega_0 = \frac{\epsilon}{H_0^2}, \tag{1.28}$$

where the density parameter Ω_0 is defined to be the value at the present epoch of

$$\Omega = \frac{8\pi G\rho}{3H^2}. \tag{1.29}$$

We may also express the density parameter as

$$\Omega_0 = \rho_0/\rho_c, \tag{1.30}$$

where the critical density

$$\rho_c \equiv \frac{3H_0^2}{8\pi G} = 1.88 \times 10^{-29} h^2\,\mathrm{g\,cm^{-3}} = 2.76 \times 10^{11} h^2\,\mathrm{M_\odot\,Mpc^{-3}}, \tag{1.31}$$

which corresponds the present mean density required to make the Universe bound ($\epsilon < 0$).

The solutions to equation (1.26) can be divided into three cases: $\Omega_0 < 1$, $\Omega_0 > 1$ and $\Omega_0 = 1$:

(1) $\Omega_0 < 1$: In this case the parametrized solution to equation (1.26) is

$$R = A(\cosh\eta - 1); \qquad t = B(\sinh\eta - \eta), \tag{1.32}$$

where

$$\Omega_0 = \operatorname{sech}^2\left(\tfrac{1}{2}\eta_0\right); \qquad \frac{A^3}{B^2} = \frac{4\pi G\rho_0}{3}. \tag{1.33}$$

This solution represents a model for an open Universe. The curvature of three-dimensional space for this solution is negative and the Universe expands forever.

(2) $\Omega_0 > 1$: The parametrized solution for this case is

$$R = A(1 - \cos\eta); \qquad t = B(\eta - \sin\eta), \tag{1.34}$$

where

$$\Omega_0 = \sec^2\left(\tfrac{1}{2}\eta_0\right); \qquad \frac{A^3}{B^2} = \frac{4\pi G\rho_0}{3}. \tag{1.35}$$

This solution represents a closed Universe which eventually recollapses after maximum expansion. The space part of this solution has positive curvature.

(3) $\Omega_0 = 1$: Finally, in this case, the solution for equation (1.26) becomes

$$R = \left(\frac{3H_0 t}{2}\right)^{2/3}, \tag{1.36}$$

which represents a flat Universe. This solution also expands forever, but the expansion rate becomes asymptotically zero. The space part of this solution has zero curvature, and can be represented by Euclidean geometry.

The determination of Ω_0 from observation is of great importance because this parameter determines the future evolutionary path of the Universe. Notice that the value of Ω changes through the evolution, but the sign of $\Omega - 1$ remains unchanged.

4 Mass-to-light ratio in the Solar neighbourhood

The most important indicator of the presence of dark matter is the mass-to-light ratio, M/L. It is very instructive to know M/L in the region close to the Sun, where high precision observations are possible. Since we have a reasonably good idea of the constituents in the Solar neighbourhood—certainly better than in any other system—comparison of the dynamically determined M/L with that deduced from an inventory of the constituents serves to indicate the amount of dark matter more accurately than in more distant systems.

Since different stellar populations have different scale heights, the surface density integrated along the direction perpendicular to the Galactic plane is a more fundamental quantity than the volume density. Table 1 summarizes the surface density of mass and luminosity per square parsec integrated within $|z| < 700$ pc, where z is the height above the Galactic midplane.

From this table we deduce M/L in the Solar neighbourhood to be $\Sigma/I \approx 3.3$ for all known constituents.

Table 1. Surface density and brightness of Solar neighbourhood

Species	$\Sigma(M_\odot/pc^2)$	$I(L_\odot/pc^2)$
Visible stars	27	15
Dead stars	18	0
Gas	5	0
Total	50	15

The total mass density can be determined from carefully selected star samples by analyzing the velocity dispersion and density profile in the direction normal to the Galactic plane. These studies give the mass-to-light ratio within 700 pc,

$$\frac{M}{L}(|z| < 700\text{pc}) \approx 5, \tag{1.37}$$

which is slightly larger than that given by Table 1. Therefore, there is marginal evidence for dark matter in the Solar neighbourhood. This subject is discussed in greater detail in the lectures by John Bahcall.

The M/L's derived for the Solar neighbourhood are only benchmarks and should not necessarily be expected to apply to other systems. For example, roughly 95% of the light in the Solar neighbourhood comes from stars brighter than the sun but about 75% of the mass is contained in the stars fainter than the sun. Therefore, any slight variation of the initial mass function can change M/L significantly.

5 Classification scheme of dark matter

Dark matter has been discussed in various astronomical contexts, and it is worth bearing in mind that both the reliability of the evidence and the nature of the dark matter may be quite different in different contexts. There are at least four different categories of possible dark matter (DM): (1) DM in the Solar neighbourhood, (2) DM in galaxies, (3) DM in clusters and groups of galaxies, and (4) DM in cosmology.

As we discussed above, there is marginal evidence for DM in the Solar neighbourhood. This is based on the discrepancy between the mass determined by local vertical dynamics and the mass detected by direct observations. The study of the Solar neighbourhood is important even though not much DM may be present, because it is here that we have best hope of determining the nature of the DM. One further property of the DM in the Solar neighbourhood is that it is concentrated in a disk, and hence must almost surely be composed of baryons, since dissipation in a rotating gas is by far the most common way to form disks.

Stronger evidence for DM can be found in galaxies. Recent 21-cm radio observations have revealed the ubiquity of flat rotation curves in spiral galaxies out to radii much larger than the radii containing most of the visible stars. (It is much more difficult to measure the rotation curves for elliptical galaxies.) If the mass is proportional to the light, the rotation curves should exhibit a Keplerian falloff at large radii (that is, $V_{\rm rot} \propto r^{-1/2}$). The flat rotation curves, instead, suggest that the mass within the radius r

$$M(r) \approx \frac{V_{\rm rot}^2 r}{G} \propto r, \tag{1.38}$$

which yields M/L's of up to 30–40 for individual galaxies and possibly much more depending on the extent of the flat rotation curve beyond the last measured point. If the rotation curves remain flat to several hundred kiloparsecs (as proposed by Ostriker *et al.* 1974), the M/L's of individual spiral galaxies may be comparable to those in clusters.

Clusters of galaxies have provided the best evidence for DM ever since Zwicky's original work. The Coma cluster of galaxies provides perhaps the single strongest piece of evidence for DM because of the dramatic difference in its M/L from that of the Solar neighbourhood (see lecture 6).

The principal cosmological evidence for DM is that theoretical prejudice, as well as some specific models of the early Universe such as inflationary models, require $\Omega_0=1$. To determine what M/L this implies we must estimate the mean luminosity density. The determinations of the present average luminosity density in the Universe by Davis and Huchra (1982) and Kirshner *et al.* (1983) can be averaged to give

$$j_0 \approx 1.7 \times 10^8 h(\mathrm{L}_\odot\,\mathrm{Mpc}^{-3})_V. \tag{1.39}$$

Using equation (1.31), the density parameter can be expressed as

$$\Omega_0 = \frac{\rho_0}{\rho_c} = \frac{M}{L}\frac{j_0}{\rho_c} \approx 6.1 \times 10^{-4} h^{-1} \left(\frac{M}{L}\right)_V \approx \frac{(M/L)_V}{1600h}, \tag{1.40}$$

which means that M/L should be about $1600h$ in order for Ω_0 to be unity. This M/L is larger than that of the Coma cluster by about a factor of four. There is no known system of galaxies whose dynamical mass implies a mass-to-light ratio as high as $1600h$; hence, if the Universe is closed, the mass-to-light ratio must be much larger outside galaxy systems—even those as large as Coma—than within them.

An independent constraint on the total amount of baryonic matter in the Universe comes from the study of primordial nucleosynthesis. The present amount of deuterium and helium is mainly produced during an early phase of the Universe. The observed abundances of deuterium and helium set a limit on Ω_0 in the form of baryonic matter

$$\Omega_B \approx (0.011 - 0.048)h^{-2}, \tag{1.41}$$

which translates to $20h^{-1} < (M_B/L)_V < 80h^{-1}$ through equation (1.40), where M_B is the mass of baryons.

If all these arguments are taken seriously, then (1) there must be *both* baryonic DM (to provide the DM in the Solar neighbourhood) and non-baryonic DM [so that $\Omega_0 = 1$ withoout violating equation (1.41)]; (2) the DM in galaxies may be baryonic but the DM in clusters like Coma must be non-baryonic (unless h is as small as 0.5, in which case the upper limit to M_B/L implied by nucleosynthesis may be barely consistent with the M/L's of rich clusters); (3) the ratio of DM to luminous mass must be larger outside galaxies, groups, and clusters than inside, since the mass-to-light ratios of these systems are not sufficient to close the Universe.

The candidates for non-baryonic DM in the context of particle physics are discussed by H. Harari in his lectures.

II. THEORY OF STELLAR DYNAMICS

In most cases the determination of masses of galaxies and systems of galaxies is based on the dynamical theory of stellar systems. Here by "stellar systems" we

mean systems composed of self-gravitating point masses, which may be either stars or galaxies. For the following lectures, we will restrict ourselves to spherical systems, although many of the objects we are interested in are not spherical. The main reason for using the spherical approximation is simplicity. Also the potential distribution is much more round than the density distribution so that even for flattened mass distributions a spherical potential is often not a bad approximation.

1 Collisionless Boltzmann equation

We will first consider the equation of continuity for a fluid, and draw the analogy to the dynamics of discrete point masses in the next paragraph. The rate of change of mass within a volume V with surface S is

$$\frac{dM}{dt} = \int_V \frac{\partial \rho}{\partial t} d^3\vec{r} = -\int_S \rho \vec{v} \cdot d^2\vec{S}, \tag{2.1}$$

where $\rho(\vec{r}, t)$ is the density distribution and $\vec{v}(\vec{r}, t)$ is the velocity field. Using Gauss's theorem,

$$\int_S \rho \vec{v} \cdot d^2\vec{S} = \int \nabla \cdot (\rho \vec{v}) d^3\vec{r}, \tag{2.2}$$

so that equation (2.1) can be written,

$$\frac{dM}{dt} = \int_V \frac{\partial \rho}{\partial t} d^3\vec{r} = -\int_V \nabla \cdot (\rho \vec{v}) d^3\vec{r}. \tag{2.3}$$

Since the equation must hold for an arbitrary volume in the absence of source and sink terms, we have

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0, \tag{2.4}$$

which is the usual equation of continuity for fluids.

The dynamics of a system of point masses can be conveniently described by employing the phase space density distribution $f(\vec{r}, \vec{v}; t)$, where $f(\vec{r}, \vec{v}; t) d^3\vec{r} d^3\vec{v}$ is either the number, luminosity or mass contained in a phase space volume $d^3\vec{r} d^3\vec{v}$. Then one can derive a similar equation to the equation of continuity for the phase space density distribution using the same particle conservation argument in six-dimensional phase space rather than in normal three-dimensional space:

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{3} \left[\frac{\partial}{\partial x_i}(f\dot{x}_i) + \frac{\partial}{\partial v_i}(f\dot{v}_i) \right] = 0. \tag{2.5}$$

Now notice that $\dot{x}_i = v_i$, and $\partial v_i / \partial x_i = 0$. Furthermore $\dot{v}_i = -\partial\Phi/\partial x_i$, where Φ is the gravitational potential, so that $\partial \dot{v}_i / \partial v_i = 0$. Thus we find

$$\frac{\partial f}{\partial t} + \sum_{i=1}^{3} \left(v_i \frac{\partial f}{\partial x_i} - \frac{\partial \Phi}{\partial x_i} \frac{\partial f}{\partial v_i} \right) = 0, \tag{2.6}$$

which is called the Vlasov or collisionless Boltzmann equation. Now define a convective derivative operator

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + \sum_i \left(v_i \frac{\partial}{\partial x_i} - \frac{\partial \Phi}{\partial x_i} \frac{\partial}{\partial v_i} \right), \tag{2.7}$$

which gives the rate of change of a quantity as seen by an observer moving with a given star. Thus the collisionless Boltzmann equation is simply

$$\frac{Df}{Dt} = 0. \tag{2.8}$$

The collisionless Boltzmann equation states that the local phase space density as viewed by an observer moving with a given star is conserved. This is analogous to a phenomenon seen in a marathon race. In the beginning of the race the spatial density of runners is high but their speeds vary over a wide range. As the race progresses, runners whose speed is nearly the same stay together so that near the finish, the runners in any given location have a low spatial density but travel at nearly the same speed. Therefore, phase space density remains roughly constant throughout the race. (We have assumed that all the runners who started the race finish so there are no sinks, and that each runner travels at constant speed.)

2 The Jeans theorem

If the stellar system is in a steady state, the partial derivative with respect to time in equation (2.6) vanishes. We define integrals of motion $I(\vec{x}, \vec{v})$ to be functions such that

$$\frac{d}{dt} I\left[\vec{x}(t), \vec{v}(t)\right] = 0, \tag{2.9}$$

along any trajectory $[\vec{x}(t), \vec{v}(t)]$. The integrals satisfy the relation,

$$\frac{dI}{dt} = \sum_{i=1}^{3} \left(\dot{x}_i \frac{\partial I}{\partial x_i} + \dot{v}_i \frac{\partial I}{\partial v_i} \right) = \sum_{i=1}^{3} \left(v_i \frac{\partial I}{\partial x_i} - \frac{\partial \Phi}{\partial x_i} \frac{\partial I}{\partial v_i} \right) = 0. \tag{2.10}$$

Therefore, the integrals satisfy the time-independent collisionless Boltzmann equation and thus the phase space density distribution is a function only of the integrals, i.e., $f = f(I_1, I_2, ...)$. This is known as the Jeans theorem. In general an arbitrary stellar system can have six different integrals. In spherical systems only four of these are important for stellar dynamics: the energy per unit mass, E, and three components of the angular momentum per unit mass, $\vec{L}$. If the system is spherically symmetric in all respects (that is, any variable depends only on the distance from the center), the distribution function depends only on E and the square of the angular momentum, L^2. Therefore, the general solution of the collisionless Boltzmann equation for a spherically symmetric stellar system is any function of the form $f(E, L^2)$.

The number density distribution in a spherical system can be computed from

$$\nu(r) = \int_{-\infty}^{\infty} dv_r \int_0^{\infty} 2\pi v_t dv_t f\left(\tfrac{1}{2}v_r^2 + \tfrac{1}{2}v_t^2 + \Phi(r), r^2 v_t^2\right), \tag{2.11}$$

where v_t and v_r are the tangential and radial velocities, respectively. If f represents the luminosity density or mass density, we can replace $\nu(r)$ by $j(r)$ or $\rho(r)$, respectively. The gravitational potential Φ satisfies Poisson's equation

$$\frac{1}{r^2}\frac{d}{dr}\left[r^2\frac{d\Phi(r)}{dr}\right] = 4\pi G\rho(r). \tag{2.12}$$

Ideal observations could give us, at best, the distribution of $I(\vec{R}, v_\parallel)$, where $I(\vec{R}, v_\parallel)\, d^2\vec{R}\, dv_\parallel$ is the luminosity in the area $d^2\vec{R}$ and velocity interval $dv_\parallel$ at projected position $\vec{R}$ and line-of-sight velocity $v_\parallel$. Even this information—far more than we are able to glean from present-day observations—is insufficient to give the distribution function $f(E, L^2)$, since we do not know the potential Φ to use in equation (2.11). Therefore, the determination of f is always an intrinsically underdetermined problem. One must make additional assumptions to get the distribution function and mass distribution [for example, $j(r) \propto \rho(r)$].

3 Examples of distribution functions

In general the distribution functions for real stellar systems are not well-known. However, one often can make reasonable models using known analytic distribution functions. We will give two very simple examples of such distribution functions.

3.1 Plummer model

This is a power-law model for the phase space distribution function

$$f(E, L^2) = \begin{cases} K|E|^{7/2}, & \text{for } E < 0; \\ 0, & \text{for } E > 0. \end{cases} \tag{2.13}$$

In this case, the density distribution becomes

$$\nu(r) = 4\pi K \int_0^{\sqrt{2|\Phi|}} (|\Phi| - \tfrac{1}{2}v^2)^{7/2} v^2 dv = 7\pi^2 \cdot 2^{-11/2} K|\Phi|^5. \tag{2.14}$$

Now assume that $\nu = \rho$, i.e., $\nu(r)$ satisfies Poisson's equation

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d|\Phi|}{dr}\right) = -4\pi G\nu = -4\pi G \cdot 7\pi^2 \cdot 2^{-11/2} K|\Phi|^5. \tag{2.15}$$

A solution is a potential of the form

$$\Phi = -\frac{\Phi_0}{\sqrt{1 + r^2/a^2}}, \tag{2.16}$$

where Φ_0 is the central potential and the characteristic length scale a satisfies the relation,

$$a^2 = \frac{3 \cdot 2^{7/2}}{7\pi^3 G K \Phi_0^4}. \tag{2.17}$$

We have normalized the potential so that it vanishes at infinity. The Plummer model is one of the simplest models of stellar systems. The density distribution extends to infinity, but the mass is finite. It is obvious from equations (2.14) and (2.16) that the density falls off as r^{-5} at large radius. Generally speaking, such a rapid falloff of density at large r is not compatible with the observed brightness distribution of galaxies, which decays somewhat more slowly (r^{-3} to r^{-4}). The Plummer model has the same density distribution as a gaseous polytrope with polytropic index $n = 5$.

3.2 Isothermal sphere

Consider a distribution function that follows the Maxwell-Boltzmann law

$$f(E) = Fe^{-E/\sigma^2}. \tag{2.18}$$

Such a distribution is achieved in gases through relaxation by collisions. However, in galaxies, where the relaxation timescale due to two-body gravitational encounters is much longer than the Hubble time, there is no fundamental reason to reach such a state. Nonetheless, the cores of elliptical galaxies and bulges of spirals are often found to be fitted with this model very well.

The density distribution for this model is

$$\nu(r) = 4\pi F \int_0^\infty v^2 dv e^{-\Phi/\sigma^2} e^{-v^2/2\sigma^2} = F(2\pi\sigma^2)^{3/2} e^{-\Phi/\sigma^2}. \tag{2.19}$$

If we set $\nu = \rho$, and choose the potential at the center to be zero, the density distribution becomes

$$\rho(r) = \rho_0 e^{-\Phi/\sigma^2}. \tag{2.20}$$

By substituting equation (2.20) into Poisson's equation, we get

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\Phi}{dr}\right) = 4\pi G\rho_0 e^{-\Phi/\sigma^2}. \tag{2.21}$$

If we introduce the dimensionless variables, $\psi = -\Phi/\sigma^2$, $r_0 = \sqrt{9\sigma^2/4\pi G\rho_0}$, $s = r/r_0$, we get the dimensionless equation for ψ

$$\frac{1}{s^2}\frac{d}{ds}s^2\frac{d\psi}{ds} = -9e^{\psi}, \tag{2.22}$$

which also should satisfy the boundary conditions,

$$\psi(0) = 0, \quad \psi'(0) = 0. \tag{2.23}$$

Figure 1 shows the density and surface density distribution for the solution of equation (2.22), which is called the isothermal sphere. The central surface density satisfies $\Sigma_0 \approx 2\rho_0 r_0$ (more precisely the constant is 2.018), and the volume density falls as r^{-2} at large r. The radius parameter r_0 is called the core radius and it

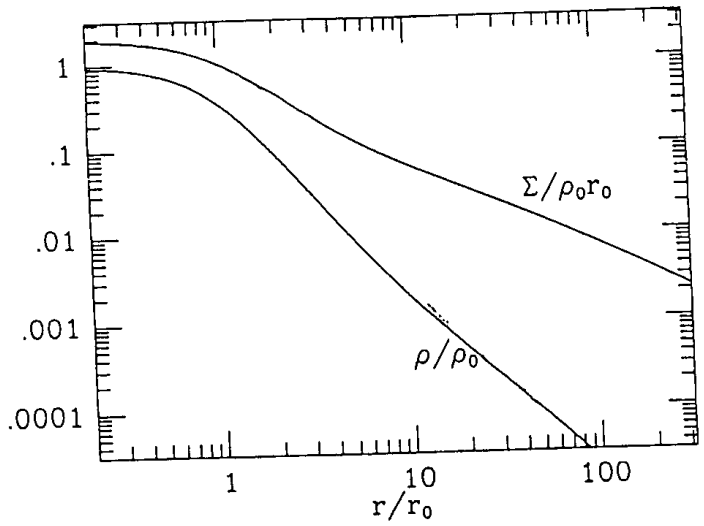

Figure 1: Density $\rho(r)$ and surface density $\Sigma(r)$ for an isothermal sphere, in units of the central density ρ_0 and core radius r_0.

roughly corresponds to the radius where the surface density becomes half of the central value.

However, the isothermal sphere is unrealistic in that the total mass is infinite since $\rho \propto r^{-2}$ at large radii. More realistic models can be obtained by decreasing the phase space density at high energy. Models of this type are known as King or Michie models, but we will not discuss them here since the truncation affects only the outer parts, while we are concerned mostly with the inner parts in fitting observations of elliptical galaxies.

4 Moments of the collisionless Boltzmann equation

By taking moments of the collisionless Boltzmann equation, we can get some useful equations relating observable quantities. The first step is just to integrate equation (2.6) over velocity space to get

$$\frac{\partial \nu}{\partial t} + \frac{\partial}{\partial x_i}(\nu \bar{v}_i) = 0, \tag{2.24}$$

where $\nu \bar{v}_i = \int f v_i d^3\vec{v}$, and we have adopted the Einstein summation convention for simplicity of notation. The above equation is simply the equation of continuity (2.4). We can also multiply equation (2.6) by v_j and integrate over velocity space to obtain

$$\frac{\partial}{\partial t}(\nu \bar{v}_j) + \frac{\partial}{\partial x_i}(\nu \overline{v_i v_j}) + \frac{\partial \Phi}{\partial x_j}\nu = 0, \tag{2.25}$$

One can generalize the above formula a little if one assumes that the cores of galaxies satisfy only the first three assumptions: spherical symmetry, isotropic velocity distribution and constant M/L. In this case, one can write the mass-to-light ratio as

$$\frac{M}{L} = \eta \cdot \frac{9\sigma_p^2}{2\pi G I_0 r_h}, \tag{3.4}$$

where σ_p^2 is the line-of-sight mean square velocity dispersion at the center, η is a constant that depends on the actual distribution function of the stellar system, and r_h is the radius at which the surface brightness drops to half of its central value. The constant η is found to be very insensitive to the details of the actual models. For example, $\eta = 0.971$ for a Plummer model. One finds empirically that $\eta = 1$ to within a few percent for most stellar systems satisfying the three assumptions that have density distributions resembling those of real galaxies (Richstone and Tremaine 1986). Equation (3.4) with $\eta = 1$ is known as King's core-fitting formula and is commonly used for the determination of the mass-to-light ratios in the cores of galaxies.

How accurate is the core-fitting formula likely to be? The most serious concerns are the assumptions of velocity isotropy and spherical symmetry. A spherical galaxy composed entirely of stars on circular orbits would have $\sigma_p = 0$ and hence $\eta \to \infty$. However, such models are unrealistic; in the more plausible case of radial anisotropy ($\overline{v_r^2} > \overline{v_\phi^2}$), the constant η can be as small as 0.65 (Merritt 1987b). For non-spherical models η can be as low as 0.4 if the galaxy is viewed along its long axis (Merritt 1987b). The sensitivity to shape and velocity anisotropy is reduced if the dispersion is averaged over the central core rather than being measured precisely at the center.

Despite these concerns, in the absence of information on the orientation and velocity anisotropy of a given galaxy, the core-fitting formula (3.4) with $\eta = 1$ offers the best available estimate of the central M/L.

This technique has been applied to the cores of elliptical galaxies and bulges of spiral galaxies by Kormendy (1987a). No systematic differences are found between bulges and elliptical cores. The median value for the elliptical cores is found to be $M/L \approx 12h$, which is similar to the mass-to-light ratio expected from the Solar neighbourhood if gas and young stars are removed. This suggests that the DM in the Solar neighbourhood is also present in ellipticals and probably is a normal component of the stellar population.

The same technique has been applied to dwarf spheroidal galaxies by Aaronson and Olszewski (1987). There are about a half dozen low mass dwarf spheroidals that are satellites of our own Galaxy. Such galaxies are good places to look for DM since the density of luminous matter is very low. However, there are some difficulties in obtaining reliable velocity dispersion data for dwarf spheroidals since the velocity dispersion is low and the surface brightness is small. One has to measure accurate velocities of a number of individual stars to get a reliable velocity dispersion.

The velocity dispersion may be misleading if the measured stars are members of binary systems. To avoid contamination due to binaries, Aaronson and Olszewski measured the velocities of individual stars more than once to see if there is any velocity variation over the observation interval (typically one year). Any stars showing velocity variation are excluded in obtaining the velocity dispersion. The resulting M/L's for the dwarf spheroidals are listed in Table 2, as determined by equation (3.4) with $\eta = 1$. Notice that the stars used in estimating σ_p are spread

Table 2. Mass-to-light ratios of dwarf spheroidal galaxies

Name	N	$\sigma_p(\mathrm{km\,s^{-1}})$	$I_0(\mathrm{L_\odot\,pc^{-2}})$	$r_h(\mathrm{kpc})$	M/L
Fornax	5	6.4 ± 2	16.2	0.50	1.7
Sculptor	3	5.8 ± 2.4	9.7	0.17	6.8
Carina	6	5.6 ± 1.6	4.0	0.22	12
Ursa Minor	7	11 ± 3	1.2	0.15	220
Draco	9	9 ± 2	2.5	0.15	71

Notes: N denotes the number of stars used to determine the velocity dispersion. Parameters from Kormendy (1987b).

out over the galaxy rather than being concentrated at the center; this implies that our estimates of M/L will be systematically low, by up to a factor of two or so.

The M/L's of some dwarf spheroidals are normal (similar to the Solar neighbourhood or central parts of elliptical galaxies), but two, Draco and Ursa Minor, show very large M/L. If these large M/L's are correct, they show that some dwarf spheroidals are composed mostly of dark matter.

It has often been asked whether the exceptionally large M/L for Draco and Ursa Minor could be due to binary stars. Thus, one needs to know how much velocity dispersion can be attributed to the orbital motion of binaries, given the selection criterion against velocity variation employed by Aaronson and Olszewski. To address this question quantitatively, we have made a simple simulation. We have calculated the velocity dispersion due to orbital motions of binaries assuming that (1) all stars are in binary systems; (2) the primary component has mass $M_1 = 0.8\,\mathrm{M_\odot}$ and the cumulative mass distribution of the secondaries is $\propto M_2^{0.4}$ with $M_2 < M_1$, similar to the distribution for binaries in the Solar neighbourhood; (3) the binary periods P are uniformly distributed in $\log(P)$ over a interval ± 0.25, centered on some value $\log(P_0)$; and (4) the eccentricity distribution is uniform in e^2, as expected in statistical equilibrium. Stars showing velocity variation $\geq 4\,\mathrm{km\,s^{-1}}$ between two epochs separated by 1 year are excluded in computing the line-of-sight velocity dispersion since the observers exclude them as well. Figure 2 shows the line-of-sight velocity dispersion σ_p as a function of period P_0. Also shown is the ratio f_v of the number of stars showing velocity variation to the number showing no variation. This graph shows that there is a good correlation between the velocity dispersion due to binary systems and the fraction of stars showing velocity variation. Therefore, once the fraction of stars showing velocity variation among the sample is known, one can estimate the contribution to the velocity dispersion from binary orbital motions.

This result can be applied to the two observed dwarf spheroidals with high M/L's. The Draco system has 11 stars observed with radial velocities. Two of them show velocity variation, so $f_v \approx 0.2$. From Figure 2, the velocity dispersion contribution due to binaries is then at most about $2\,\mathrm{km\,s^{-1}}$. The contribution of binaries to the observed dispersion of $9\,\mathrm{km\,s^{-1}}$ is therefore less than about 5% in quadrature. The same analysis for Ursa Minor, which has a dispersion of $11\,\mathrm{km\,s^{-1}}$ from 7 stars (3 out of 10 measured stars show velocity variation, so $f_v \approx 0.43$), shows that the dispersion due to binaries is about $4\,\mathrm{km\,s^{-1}}$ so that the observed velocity dispersion is in error by at most 13%. Similar results are obtained for other period distributions. Therefore it may safely be concluded that the effects of

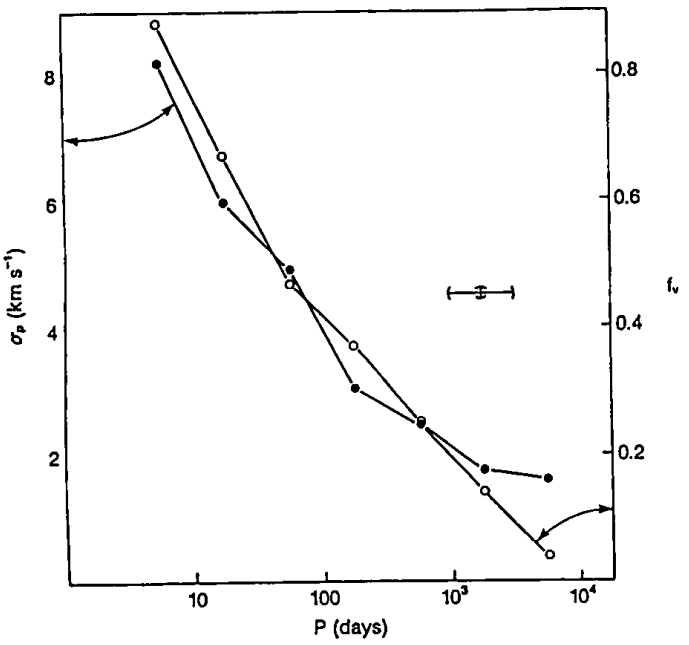

Figure 2: Plot of line-of-sight velocity dispersion σ_p due to binaries as a function of binary period, as well as the ratio f_v of the number of stars showing velocity variation exceeding $4\,\mathrm{km\,s^{-1}}$ over a 1 year interval to the number showing variation $< 4\,\mathrm{km\,s^{-1}}$. We have assumed that all stars are in binaries with $\log(P)$ uniformly distributed over the interval ± 0.25 around the plotted value. The error bars indicate the width of the assumed distribution in $\log(P)$ and the statistical uncertainty in the results. The curves are not completely smooth because of resonance effects between the mean binary period and the observation interval.

binaries are negligible in Draco and Ursa Minor.

Are there other effects that might contribute to erroneous M/L's? The radial velocity data could be contaminated by motions of the stellar atmospheres. However, K-giants, which are the main sources of velocity determinations, have very small atmospheric motions. The core-fitting formula (3.4) is already biased toward low M/L's because it is based on the central velocity dispersion while the measured stars often lie outside the core where the dispersion is lower. Statistical errors are also biased toward smaller M/L since the χ^2 distribution is asymmetric.

Perhaps the most interesting alternative to the existence of large quantities of dark matter is the possibility that Draco and Ursa Minor are unbound systems. Their extremely low surface brightness, and their location in the plane of the orbit of the Magellanic Clouds, suggest that they may be only apparent density enhancements arising, for example, from the crossing of streams of tidal debris from the Clouds.

IV. THE EXTENT OF THE GALACTIC HALO

Most spiral galaxies, including our own, have flat rotation curves as far as one can measure. This naturally leads to the question, how far beyond the last measured point do rotation curves stay flat, that is, how far do dark galactic halos extend? Here we will discuss various methods to determine the extent of the dark halo of our Galaxy.

We shall employ a very simple spherical model in which the Galaxy has a flat rotation curve up to the radius r_*; thus

$$V_{\rm rot}^2 = \begin{cases} V_0^2 & \text{for } r < r_*, \\ \propto \dfrac{1}{r} & \text{for } r > r_*; \end{cases} \tag{4.1}$$

which gives the mass distribution

$$M(r) = \begin{cases} \dfrac{V_0^2 r}{G} & \text{for } r < r_*; \\ M_* = \dfrac{V_0^2 r_*}{G} & \text{for } r > r_*. \end{cases} \tag{4.2}$$

Now our aim is to determine the value of r_*. Rotation curve measurements show constant rotation velocity up to $r \approx 2R_0$, where R_0 is the Galactocentric distance of the sun; thus $r_* \gtrsim 2R_0 \approx 17\,\mathrm{kpc}$.

1 Local escape speed

The velocity distribution of stars in the Solar neighbourhood with respect to the rest frame of our Galaxy shows a cut-off near $v_{\rm max} \approx 500\,\mathrm{km\,s^{-1}}$ (Carney and Latham 1987). This is interpreted as a lower limit to the escape speed of stars in the Solar neighbourhood. This means that the following equation holds,

$$\tfrac{1}{2} v_{\rm max}^2 + \Phi(R_0) < 0. \tag{4.3}$$

The gravitational potential satisfies the equation,

$$\frac{d\Phi}{dr} = \frac{GM(r)}{r^2}, \tag{4.4}$$

which in the model of equations (4.1) and (4.2) can be integrated to yield

$$\Phi = \begin{cases} V_0^2 \left[\ln \dfrac{r}{r_*} - 1 \right] & \text{for } r < r_*, \\ -\dfrac{V_0^2 r_*}{r} & \text{for } r > r_*, \end{cases} \tag{4.5}$$

where we have normalized the potential to vanish at infinity. By plugging $V_0 = 220\,\mathrm{km\,s^{-1}}$ and $v_{\rm max} = 500\,\mathrm{km\,s^{-1}}$ into equation (4.3), we get $r_* \gtrsim 4.9R_0$. If we

further use $R_0 \approx 8.5\,\mathrm{kpc}$, the extent of the halo of our galaxy is roughly $r_* \gtrsim 41\,\mathrm{kpc}$. This corresponds to a mass $M_* \gtrsim 4.6 \times 10^{11}\,\mathrm{M}_\odot$. Since the total luminosity of our galaxy at the V band is $L_V \approx 1.4 \times 10^{10}\,\mathrm{L}_\odot$, the mass-to-light ratio becomes

$$\frac{M}{L} \gtrsim 33, \tag{4.6}$$

at least a factor of six larger than the M/L in the Solar neighbourhood.

2 Magellanic Stream

The Large Magellanic Cloud (LMC: Galactocentric distance $d = 52\,\mathrm{kpc}$ and $M \approx 2 \times 10^{10}\,\mathrm{M}_\odot$) and Small Magellanic Cloud (SMC: $d = 63\,\mathrm{kpc}$ and $M \approx 2 \times 10^9\,\mathrm{M}_\odot$) are the two nearest satellite galaxies to us. They contain $\approx 5 \times 10^9\,\mathrm{M}_\odot$ of neutral hydrogen gas. In addition, there is a stream or trail of HI extending away from the clouds. This HI trail follows a great circle and contains almost $10^9\,\mathrm{M}_\odot$ of neutral hydrogen. It is presumably composed of gas from the Clouds that has been stripped off by the tidal field of the Galaxy.

Most of this gas is probably on free Kepler orbits, not too dissimilar from that of the Clouds. At the tip of the stream, gas is falling towards us at high speed ($v \approx -220\,\mathrm{km\,s^{-1}}$ in the Galactic rest frame). If the Galactic potential were that of a point mass, this large velocity would suggest that the material at the tip has fallen deep into the potential, to a Galactocentric distance $\lesssim 15\,\mathrm{kpc}$. However, parallax effects due to the offset of the Sun from the Galactic center should then spoil the great circle shape of the stream. This argument suggests that there is a massive halo in our Galaxy, since then the required infall velocity can be achieved at a larger radius.

Detailed dynamical models have been constructed by Murai and Fujimoto (1980) and Lin and Lynden-Bell (1982). If a flat rotation curve is assumed for our Galaxy, the best fitting circular velocity is $244 \pm 20\,\mathrm{km\,s^{-1}}$, and the rotation curve must remain flat to at least 70 kpc. This suggests that there is a very large amount of DM in our Galaxy, much more than the lower limit obtained above using the local escape velocity. The mass-to-light ratio of the Galaxy would exceed 56 in Solar units if such an extended halo were present.

3 Local Group timing

The nearest giant spiral galaxy M31 is about 730 kpc away. Together with the associated satellite galaxies, we and M31 compose a relatively isolated system of galaxies, known as the Local Group. The relative radial velocity of the center of mass of us and M31 is $-119\,\mathrm{km\,s^{-1}}$, which has the opposite sign to the usual Hubble flow. It is possible that the two galaxies are approaching each other by chance. However, a more natural explanation is that the two galaxies were once moving apart due to the Hubble expansion, but that the expansion was slowed and reversed by their mutual gravitational attraction. If we assume M31 and Galaxy to be point masses and ignore the presumably small masses of other Local Group members, the relative orbit of the two galaxies can be written parametrically as

$$r = a(1 - e\cos\eta); \quad t = \sqrt{\frac{a^3}{GM}}(\eta - e\sin\eta) + C, \tag{4.7}$$

where C is a constant, a is the semi-major axis, e is the eccentricity, η is the eccentric anomaly, and M is the total mass of the Local Group [compare eq. (1.34)]. Since the two galaxies were presumed to be at the same place at the beginning of the expansion, $r=0$ at $t=0$ so that we have to choose $e=1$ and $C=0$, corresponding to a radial orbit. The relative radial velocity may be written

$$\frac{dr}{dt} = \sqrt{\frac{GM}{a}} \frac{\sin\eta}{1-\cos\eta} = \frac{r}{t} \frac{\sin\eta(\eta-\sin\eta)}{(1-\cos\eta)^2}. \tag{4.8}$$

By plugging in $dr/dt = -119\,\mathrm{km\,s^{-1}}$ and $t = 10$–$20\,\mathrm{Gyr}$ ($1\,\mathrm{Gyr} = 10^9$ years), we get $\eta = 4.11$ (for $t = 10\,\mathrm{Gyr}$) to 4.46 ($t = 20\,\mathrm{Gyr}$) radians. Then the mass of the Local Group is found to be $M = 5.5 \times 10^{12}\,\mathrm{M_\odot}$ (for $t = 10\,\mathrm{Gyr}$) or $3.2 \times 10^{12}\,\mathrm{M_\odot}$ (for $t = 20\,\mathrm{Gyr}$). This gives the mass-to-light ratio of the Local Group

$$\frac{M}{L} = 76 - 130, \tag{4.9}$$

which is again very large. We expect that this roughly represents the M/L of our Galaxy if there is no big variation of M/L between M31 and the Galaxy.

4 Kinematics of satellite galaxies

In principle, the kinematics of satellite galaxies and globular clusters can provide a useful tool to determine the mass of the Galaxy. However, this determination is intrinsically difficult since there are only a few objects and only the line-of-sight velocities are known. In particular, the satellite galaxies are at distances large compared to the distance to the Galactic center from the Sun; so we see mainly the radial component of their velocity. Thus we must make some statistical assumption about the ratio of tangential to radial velocity.

If we assume r_* is large compared to the distances of the sample objects, the gravitational potential (4.5) has the form

$$\Phi(r) = V_0^2 \ln r + \text{constant}. \tag{4.10}$$

The moment of inertia per unit mass is

$$I = r^2. \tag{4.11}$$

and its first and second derivatives are

$$\dot{I} = 2\vec{r}\cdot\dot{\vec{r}} \quad ; \quad \ddot{I} = 2\vec{r}\cdot\ddot{\vec{r}} + 2v^2. \tag{4.12}$$

By taking an average over several galaxies, we get

$$\langle v^2 \rangle = \left\langle r \frac{d\Phi}{dr} \right\rangle = V_0^2. \tag{4.13}$$

If the velocity distribution is isotropic, $\langle v_r^2 \rangle = \frac{1}{3}\langle v^2 \rangle = \frac{1}{3}V_0^2$. For $V_0 = 220\,\mathrm{km\,s^{-1}}$, this gives $\langle v_r^2 \rangle^{1/2} = 127\,\mathrm{km\,s^{-1}}$, while the actual data exhibit $\langle v_r^2 \rangle^{1/2} = 60\,\mathrm{km\,s^{-1}}$

Table 3. Globular clusters and satellite galaxies used for Galactic mass determination

Name	$d_\odot$(kpc)	$v_\odot$(km s^{-1})†	v_G(km s^{-1})§
LMC+SMC	52	245 ± 5	62
Draco	75	-289 ± 1	-95
Ursa Minor	63	-249 ± 1	-88
Sculptor	79	107 ± 2	75
Fornax	138	55 ± 5	-34
Carina	91	230 ± 1	14
AM1	116	116 ± 15	-42
NGC 2419	90	-20 ± 5	-26
Pal 3	91	89 ± 9	-59
Pal 4	105	75 ± 5	54

Notes: The list contains all known satellites with $d_\odot > 50$ kpc and published velocity error < 20 km s^{-1}. Data sources given by Little and Tremaine (1987).
†: radial velocity with respect to the Sun
§: radial velocity with respect to the Galactic center

(Table 3). This means that the existing data are consistent with a flat rotation curve extending to very large distances only if the velocity distribution is primarily tangential, a conclusion first reached by Lynden-Bell *et al.* (1983). Dissipationless collapse leads to galaxies with a velocity distribution that varies from radial to isotropic, and it is difficult to construct galaxies with predominantly tangential velocities. Thus we are led to question the existence of an extended massive halo for the Galaxy.

There have recently been substantial improvements in the accuracy of velocity data for satellite galaxies. Furthermore, it is unlikely that the quality or quantity of radial velocity data will be improved drastically in the near future. Therefore it is perhaps appropriate to elaborate on the arguments above and to make a careful statistical model of the currently available kinematic data for satellites (Little and Tremaine 1987).

Let us introduce a fictitious mass for an observed satellite i

$$\mu_i \equiv \frac{v_{r,i}^2 r_i}{G}, \tag{4.14}$$

where r_i and $v_{r,i}$ are the distance from the Galactic center and the radial velocity with respect to the rest frame of the Galaxy, respectively. For a Galactic mass M, let us denote $P(\mu_i|M)d\mu$ as the probability of finding μ_i in the range between μ_i and $\mu_i + d\mu$, which can be computed from

$$P(\mu|M) = \frac{\int d^3\vec{r}d^3\vec{v}f(\vec{r},\vec{v})\delta(\mu - v_r^2 r/G)}{\int d^3\vec{r}d^3\vec{v}f(\vec{r},\vec{v})}, \tag{4.15}$$

where $f(\vec{r},\vec{v})$ is the usual phase space density distribution. Two different models for the Galactic potential, a point mass and an infinite halo, and two different models

for the velocity distribution, radial and isotropic, are considered. For the point mass potential,

$$P(\mu|M) = \begin{cases} \dfrac{2}{5\pi}\dfrac{\max[0,(2M-\mu)]^{5/2}}{\mu^{1/2}M^3} & \text{for isotopic orbits;} \\ \dfrac{1}{\pi}\dfrac{\max[0,(2M-\mu)]^{1/2}}{\mu^{1/2}M} & \text{for radial orbits.} \end{cases} \tag{4.16}$$

Now what we want to know is the corresponding probability distribution for M given μ, that is, $P(M|\mu)$. The relation between $P(\mu|M)$ and $P(M|\mu)$ is given by Bayes's theorem

$$P(M|\mu) = \frac{P(\mu|M)P(M)}{\int P(\mu|M')P(M')dM'}, \tag{4.17}$$

where $P(M)dM$ is the *a priori* probability for a galaxy to have mass between M and $M + dM$. If several objects are available the probability becomes

$$P(M|\mu_1,\cdots,\mu_N) = P(M)\frac{\Pi_{i=1}^N P(\mu_i|M)}{\int \Pi_{i=1}^N P(\mu_i|M')P(M')dM'}. \tag{4.18}$$

We now have to specify $P(M)$ to proceed with our analysis. One of the most reasonable choices is that the *a priori* distribution is uniform in $\log M$ so that

$$P(M)dM \propto d\log M \quad \text{or} \quad P(M) \propto \frac{1}{M}. \tag{4.19}$$

This choice is obviously somewhat arbitrary. However, the results become less and less sensitive to the choice of the functional form of $P(M)$ as the number of data points increases, and for the numbers we are dealing with ($N \approx 10$) the choice of $P(M)$ has no strong influence on the results.

The data set has been gathered from various sources. All globular clusters and satellite galaxies at distances $d > 50\,\mathrm{kpc}$ with velocity errors less than $20\,\mathrm{km\,s^{-1}}$ are used. The data are listed in Table 3. The LMC and SMC are treated as one data point at their center of mass since the motion of each Cloud is greatly influenced by the other Cloud.

Using equations (4.16) and (4.18) we find that for isotropic orbits there is a 90% probability that the Galactic mass M lies in the range $[1.4, 5.2] \times 10^{11}\,\mathrm{M_\odot}$ with the median value being $2.4 \times 10^{11}\,\mathrm{M_\odot}$. For radial orbits the mass is even smaller. If we define r_* so that the mass obtained above corresponds to $V_0^2 r_*/G$ with $V_0 = 220\,\mathrm{km\,s^{-1}}$, we get $r_* \leq 46\,\mathrm{kpc}$ at the 95% confidence level for isotropic orbits. This result is self-consistent in that the upper limit to r_* is smaller than the lower limit of 50 kpc for the data so that the point mass approximation is valid over most of the orbit of a typical satellite.

A similar analysis has been made assuming an infinite halo potential. In this case, the unknown is the circular speed V_0 rather than the total mass of the Galaxy. Here we find the range of circular velocity at the 90% confidence level to be $[77, 165]\,\mathrm{km\,s^{-1}}$, with the median value being $107\,\mathrm{km\,s^{-1}}$. This is clearly too small compared to the local speed of $220\,\mathrm{km\,s^{-1}}$, and once again implies that the Galaxy's massive halo has only a limited extent.

Table 4. Values of r_* from various methods

Method	r_*(kpc)
Direct measurement of rotation curve	$\gtrsim 20$
Local escape velocity	$\gtrsim 40$
Magellanic Stream	$\gtrsim 80$
Local Group timing	≈ 100
Kinematics of satellites	≈ 50

This analysis strongly suggests that the Galactic halo does not extend much beyond the outermost visible components of the Galaxy at $r \approx 30$ kpc. However, it should be noted that the value for r_* obtained here is considerably lower than the values obtained from the Magellanic Stream and Local Group timing. If some unknown formation mechanism has placed the satellites on orbits with predominantly tangential velocities, then the satellite kinematics could be consistent with these other arguments.

5 Summary

We summarize the various results for the determination of the Galactic mass in Table 4.

All of these estimates agree that the Galactic halo is much more extended than the bulk of the visible stars and gas. Hence most of the mass of the Galaxy is dark. However, there are some contradictions between various determinations of the Galactic mass, particularly between the satellite kinematics on the one hand and the Magellanic Stream and Local Group timing on the other. The source of this discrepancy remains mysterious. Perhaps (1) the gas in the Magellanic Stream is not on free Kepler orbits; (2) the Galaxy and M31 are embedded in a low-density mass concentration so that most of the mass in the Local Group is not asssociated with either galaxy; or (3) the satellite galaxies and distant globular clusters are on nearly circular orbits.

Thus, it is still unclear whether the halo of the Galaxy—and by analogy, the halos of other spiral galaxies—extends only to 50 kpc or so, about twice the optical extent, or out to much larger distances, up to several hundred kiloparsecs.

V. BINARY GALAXIES

Carefully selected samples of binary galaxies provide another opportunity to determine the mass of individual galaxies. However, constructing a well-defined sample is not an easy task. As an example, we discuss the sample used by Turner (1976).

From the Zwicky catalogue of galaxies, he restricted his sample to galaxies in the northern hemisphere (i.e., declination $\delta > 0$) in order to ensure completeness. To avoid heavy Galactic extinction, he also selected only galaxies at Galactic latitude $|b| > 40°$. The flux limit for his sample was taken to be 15th magnitude, again to help ensure completeness. The selection criteria for binary galaxies were: (1) the separation of the binary (θ_{12}) does not exceed 8 arcminutes; (2) the next nearest galaxy should lie beyond $5\theta_{12}$. The second criterion was used to avoid possible contamination from groups and clusters. The final number of binary galaxies was 156 pairs out of some 30,000 galaxies in Zwicky catalogue.

Despite these stringent selection criteria, later close visual examination by White *et al.* (1983) showed that many of Turner's pairs are members of larger clusters or groups. There are two reasons why Turner's selection criteria prove to be insufficiently stringent. First, groups or clusters can easily have one or two bright members that are included in Turner's sample, while the other group members fall below the magnitude limit. Second, if two galaxies happen to be nearby in the plane of the sky, there may well be no other galaxy within five times their separation even in rich clusters. White *et al.* find that only 76 pairs out of Turner's 156 survived the additional culling of all pairs in visible clusters or groups. Clearly, it is very difficult to construct a large, well-defined sample of isolated binaries.

If we ignore this concern, we can carry out some analysis using the existing sample. We shall work with the moment equation (2.28) for spherical systems. If we assume a flat rotation curve so that the potential $\Phi(r) = V_c^2 \ln r + const$, and the degree of anisotropy $\beta = 1 - \overline{v_\phi^2}/\overline{v_r^2}$ is a constant independent of r, then equation (2.28) reduces to

$$\frac{d}{dr}(\nu\overline{v_r^2}) + \frac{2\overline{v_r^2}\nu}{r}\beta = -\frac{\nu V_c^2}{r}. \tag{5.1}$$

The statistical analysis of large catalogs of galaxies shows that the galaxies have a two-point correlation function $\xi \propto r^{-\gamma}$, with $\gamma \approx 1.8$. This implies that the distribution of binary galaxies follows $\nu \propto r^{-\gamma}$. A solution to equation (5.1) is then $\overline{v_r^2} = const = V_c^2/(\gamma - 2\beta)$. If the system has an isotropic velocity dispersion tensor ($\beta = 0$), we get $V_c = 170\,\mathrm{km\,s^{-1}}$ after plugging in the observed rms velocity difference $(\overline{v_\parallel^2})^{1/2} = 127\,\mathrm{km\,s^{-1}}$ from Turner's sample ($\overline{v_\parallel^2} = \overline{v_r^2}$ since the velocities are isotropic). This is more or less consistent with other determinations. However, if the orbits are radial ($\beta = 1$), no solution with $\overline{v_r^2}$ and V_c^2 positive is possible. In general, we can get any value of $V_c < 170\,\mathrm{km\,s^{-1}}$ by adjusting the unknown parameter β between 0 and $0.5\gamma = 0.9$; thus the results are entirely dependent on the unknown anisotropy of the orbits.

Do we learn anything from binary galaxies? Because of the problems we have mentioned above, the answer is "not much". Perhaps the most interesting result so far is that the velocity difference between the galaxies does not correlate either with the projected distance between binaries or the luminosity. This suggests that mass is not related to the luminosity of the host galaxy.

VI. MASSES OF GROUPS AND CLUSTERS OF GALAXIES

1 Groups of galaxies

The virial theorem discussed early in these lectures can be used to determine the masses of groups or clusters of galaxies in much the same way Zwicky did. The virial theorem [eq. (1.8)] states

$$\sum_{i=1}^{N} m_i v_i^2 - \frac{G}{2}\sum_{\substack{i,j=1 \\ i\neq j}}^{N} \frac{m_i m_j}{|\vec{r}_i - \vec{r}_j|} \approx 0, \tag{6.1}$$

where we have used $\approx$ rather than $=$ because the equation is only precisely true in a time averaged sense. We now take an average of equation (6.1) over angle,

denoted by $\langle\rangle_\Omega$. We have

$$v_i^2 = 3\langle v_{\parallel,i}\rangle_\Omega, \tag{6.2}$$

where $v_\parallel$ denotes the line-of-sight velocity. The projected distances $\vec{R}_i$ are likewise related to the three-dimensional distances $\vec{r}_i$ through

$$\left\langle \frac{1}{|\vec{R}_i - \vec{R}_j|} \right\rangle_\Omega = \frac{1}{|\vec{r}_i - \vec{r}_j|} \left\langle \left| \frac{1}{\sin\theta_{ij}} \right| \right\rangle_\Omega = \frac{\pi}{2} \frac{1}{|\vec{r}_i - \vec{r}_j|}, \tag{6.3}$$

where we have assumed that the direction of a vector $\vec{r}$ is random. The averages in equations (6.2) and (6.3) are not meaningful if one looks at only one or two galaxies. However, these relations can be used to estimate average three-dimensional velocities and separations when summed over many galaxies.

Thus we may write equation (6.1) as

$$3\sum_{i=1}^{N} m_i v_{\parallel,i}^2 - \frac{G}{\pi} \sum_{\substack{i,j=1 \\ i\neq j}}^{N} \frac{m_i m_j}{|\vec{R}_i - \vec{R}_j|} \approx 0. \tag{6.4}$$

We now have several different options for determining the mass of the system: (1) we may assume that $m_i \approx (M/L)L_i$ (i.e., the mass-to-light ratio is constant for all galaxies) or (2) we may assume that $m_i = M/N$, where N is the total number of galaxies and M is the total cluster mass (i.e., the number distribution of galaxies in phase space is proportional to the distribution of mass). If we take the second assumption, the total mass-to-light ratio may be written

$$\frac{M}{L} = \frac{3\pi N}{G} \frac{\sum_i v_{\parallel,i}^2}{\sum_k L_k \sum_{i\neq j} 1/|\vec{R}_i - \vec{R}_j|}. \tag{6.5}$$

Several alternative methods to the virial theorem for estimating masses of groups of galaxies have been proposed by Heisler *et al.* (1985). For example, there is the "median mass" estimator

$$M_{\rm Me} = \frac{f_{\rm Me}}{G} \mathrm{med}_{i,j}[(v_{\parallel,i} - v_{\parallel,j})^2 |\vec{R}_i - \vec{R}_j|], \tag{6.6}$$

where $\mathrm{med}_{i,j}$ denotes the median over all pairs of galaxies (i,j). The proportionality constant $f_{\rm Me}$ is determined to be approximately 6.5 from numerical experiments.

The median mass estimator is found to give similar masses to the virial theorem. No one method appears to be superior to another, and in a given case it appears that all known estimators tend to err in the same direction.

These mass estimators are applied to the catalog of groups of galaxies compiled by Huchra and Geller (1982). The median mass-to-light ratios are found to be approximately $400h$ in Solar units with errors being about ± 0.4 in the logarithm between the median and the quartiles (that is, a factor of 3). Thus there is clear evidence for large quantities of DM in groups.

2 Rich clusters: Coma cluster

We can make more elaborate models for rich clusters than for groups, since there are many more galaxies available. For the sake of definiteness, we will concentrate on the Coma cluster, which provides one of the best examples of the evidence for DM. The virial theorem is not useful in this case because it is plagued by issues of cluster membership in the outer parts. The core-fitting method discussed in lecture 3 is not particularly useful either, in part because there is no well-defined core in the Coma cluster.

The moment equation (2.28) can be written

$$\frac{d}{dr}(\nu\overline{v_r^2}) + \frac{2\nu\overline{v_r^2}}{r}\beta(r) = -\nu\frac{d\Phi}{dr} = -\nu\frac{GM(r)}{r^2}, \tag{6.7}$$

where the anisotropy parameter $\beta(r) = 1 - \overline{v_\phi^2}/\overline{v_r^2}$. If we assume spherical symmetry, we can determine the line-of-sight dispersion profile $v_\parallel^2(r)$ and the number density profile $\nu(r)$ from observations. However, equation (6.7) is still underdetermined, since it involves four functions, $\beta(r)$, $M(r)$, $\overline{v_r^2}(r)$, and $\nu(r)$, and we have only three constraints (the two observable quantities and the equation). Merritt (1987a) has made several assumptions to determine the mass of the cluster as follows:

(1) Number traces mass (i.e., $\nu \propto \rho$): With this assumption, the observed dispersion and number density profiles imply a nearly isotropic velocity distribution (i.e., $\beta \approx 0$). The determination of mass in this case is straightforward, and gives $M = 1.8 \times 10^{15} h^{-1}\,\mathrm{M}_\odot$.

(2) Minimum mass model: Such models can be obtained by concentrating the mass in the center as much as possible without violating observations. The density profile is assumed to be

$$\rho(r) \propto \frac{1}{(1 + r^2/r_0^2)^{n/2}}. \tag{6.8}$$

Then Merritt attempts to make r_0 as small as possible. He finds rather implausible models in this case, that is, all the orbits are nearly circular outside the core. For these models the total mass never becomes lower than $0.7 \times 10^{15} h^{-1}\,\mathrm{M}_\odot$. The DM is mostly concentrated within two optical core radii.

(3) Maximum mass model: One can obtain a rather high mass model by distributing the DM more or less uniformly. The total mass and central density vary over a wide range. However, the mass within $1h^{-1}$ Mpc is within 25% of the value obtained for case (1).

(4) Radial dependence of anisotopy: One also can choose a functional form for the radial dependence of the velocity anisotropy,

$$\beta(r) = \frac{r^2}{r^2 + r_a^2}, \tag{6.9}$$

where r_a is a free parameter. This choice has some nice features, in particular, the orbits are more radial in the outer parts, as one expects from a collisionless

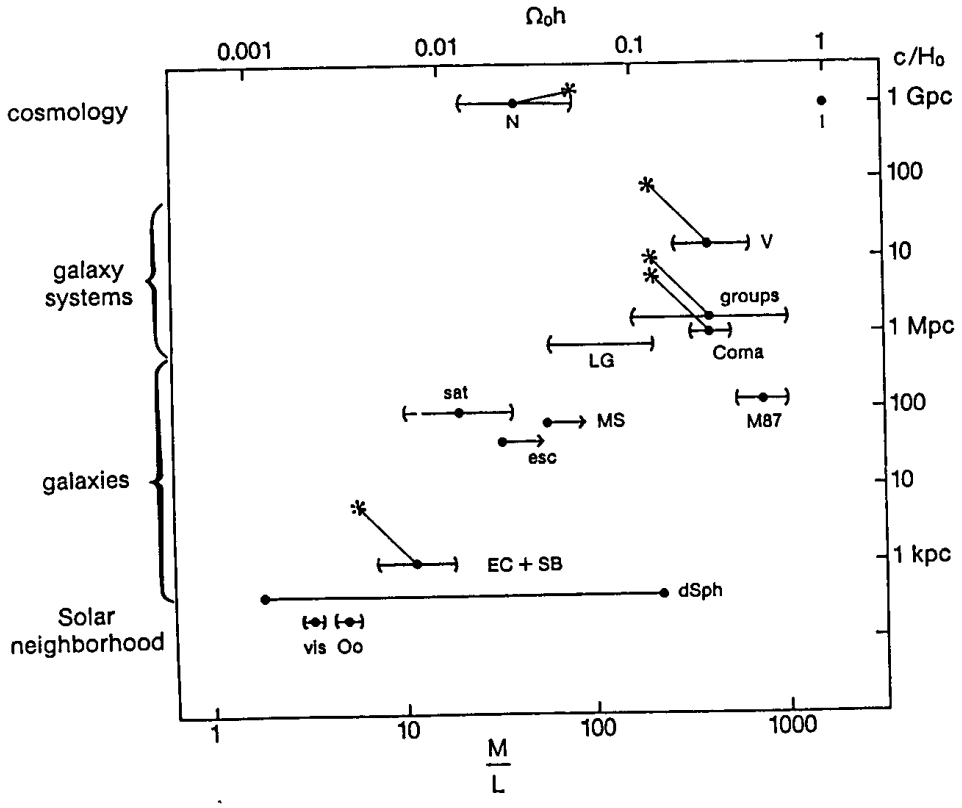

Figure 3: Mass-to-light ratios of galaxies and galaxy systems discussed in the text. The key to the labels is: "N", nucleosynthesis; "I", inflation; "V", Virgo flow; "groups", groups of galaxies; "Coma", Coma cluster of galaxies; "LG", Local Group timing; "M87", X-ray observations of M87 galaxy; "sat", satellites of our Galaxy; "MS", Magellanic Stream; "esc", escape speed from Solar neighbourhood; "EC+SB", elliptical cores and spiral bulges; "dSph", dwarf spheroidal galaxies; "vis", visible components of the Solar neighbourhood; "Oo", Oort limit. All values are plotted for $h = 1$, but a line with an asterisk at the end is used to show how results for M/L and scale change when $h = 0.5$. This is a revised version of a diagram due to Ostriker *et al.* (1974).

collapse process. Furthermore, models of this kind have a known distribution function of the form

$$f = f\left(E + \frac{L^2}{2r_a^2}\right). \tag{6.10}$$

Models with the anisotropy radius r_a exceeding the optical core radius r_0 are found to be consistent with observations. This means that models having predominantly radial orbits outside the core radius, in which the DM is more uniformly distributed than the galaxies, are allowed.

In conclusion, the mass of the Coma cluster is about $2 \times 10^{15} h^{-1}\,\mathrm{M}_\odot$ if the dark matter is distributed like the galaxies; the mass cannot be less than 40% of

this value but may be much more if the dark matter is more extended than the galaxies. For a total luminosity of $5 \times 10^{12} h^{-2}\,L_\odot$, the corresponding mass-to-light ratio is about $400h$ in Solar units. The distribution of DM is not well-constrained, but it seems to be inescapable that most of the mass in the Coma cluster is dark.

There are several other methods to determine the masses of galaxies and galaxy systems, and we will close by mentioning two.

X-ray observations of elliptical galaxies have been used to determine the mass of these galaxies, assuming hydrostatic equilibrium for the X-ray emitting hot gas. The advantage of this technique is that the distribution function of the gas is known to be isotropic, so that the anisotropy parameter $\beta = 0$. However, existing spectral data, which are essential to determine the temperature profile, have no spatial resolution, and the results are sensitive to the assumed temperature profile. The sole exception, where adequate resolution is available, is the giant elliptical galaxy M87. One problem in this case is that M87 is located in the center of the Virgo cluster so that its mass may not be typical of other galaxies. The analyses derive a very large mass-to-light ratio, at least 750 in Solar units (Fabricant and Gorenstein 1983). This technique will become more widespread when high-resolution spectral data are available from the AXAF satellite.

The perturbations to the Hubble flow induced by the nearby Virgo cluster of galaxies have been used to estimate the mass of the cluster. The mass-to-light ratio appears to be comparable to that of Coma and other rich clusters, but substantially less than the value required so that the Universe is closed.

VII. SUMMARY

The determinations of masses of galaxies and systems of galaxies on various scales show large variations in M/L. Generally M/L increases as the scale of the object increases. For example, $M/L \approx 5$ in the Solar neighbourhood (on a scale of a few hundred parsecs) while $M/L \approx 400h$ in the Coma cluster which has a scale of several Mpc. One notable exception to this trend is that some dwarf spheroidal galaxies show large M/L on scales less than a kiloparsec. The mass density implied by these M/L's is never large enough for closure, which corresponds to an average $M/L \approx 1600h$ for the entire Universe.

Baryonic matter can account for the DM on scales up to the size of dark halos of galaxies, but non-baryonic mass is probably required to provide the dark mass in rich clusters unless the Hubble constant is as small as $50\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

To summarize, we have prepared a schematic graph in Figure 3, which shows the various determinations of M/L in different objects and on different scales.

ACKNOWLEDGEMENTS

Much of the material in these lectures is drawn from a book that Tremaine has written with James Binney (Binney and Tremaine 1987), and the treatment given here has been greatly influenced by our joint labours over the past several years. We are grateful to David Merritt for constructive suggestions and criticisms on a number of topics. Of course, the responsibility for errors is our own.

REFERENCES

Aaronson, M., and Olszewski, E., 1987. In: *Dark Matter in the Universe, IAU*

Symposium No. 117, p. 153, eds. Kormendy, J., and Knapp, G.R., Reidel, Dordrecht, The Netherlands.
Binney, J.J., and Tremaine, S., 1987. *Galactic Dynamics*, Princeton University Press, Princeton, NJ.
Carney, B.W., and Latham, D.W., 1987. In: *Dark Matter in the Universe, IAU Symposium No. 117*, p. 39, eds. Kormendy, J., and Knapp, G.R., Reidel, Dordrecht, The Nehterlands.
Davis, M., and Huchra, J., 1982. *Astroph. J.*, **254**, 425.
Faber, S.M., and Gallagher, J.S., 1979. *Ann. Rev. Astron. Astroph.*, **17**, 135.
Fabricant, D., and Gorenstein, P., 1983. *Astroph. J.*, **267**, 535.
Gunn, J.E., 1978. In: *Observational Cosmology, Eighth Advanced Course of the Swiss Society of Astronomy and Astrophysics*, p. 1, eds. Maeder, A., Martinet, L., and Tammann, G., Geneva Observatory, Geneva, Switzerland.
Heisler, J., Tremaine, S., and Bahcall, J.N., 1985. *Astroph. J.*, **298**, 8.
Huchra, J.P., and Geller, M.J., 1982. *Astroph. J.*, **257**, 423.
Kirshner, R.P., Oemler, A., Schechter, P.L., and Shectman, S.A., 1983. *Astron. J.*, **88**, 1285.
Kormendy, J., 1987a. In: *Structure and Dynamics of Elliptical Galaxies, IAU Symposium No. 125*, ed. de Zeeuw, T., Reidel, Dordrecht, The Netherlands.
Kormendy, J., 1987b. In: *Dark Matter in the Universe, IAU Symposium No. 117*, p. 139, eds. Kormendy, J., and Knapp, G.R., Reidel, Dordrecht, The Netherlands.
Kormendy, J., and Knapp, G.R., eds. *Dark Matter in the Universe, IAU Symposium No. 117*, Reidel, Dordrecht, The Netherlands.
Lin, D.N.C., and Lynden-Bell, D., 1982. *Mon. Not. Roy. Astron. Soc.*, **198**, 707.
Little, B., and Tremaine, S., 1987. *Astroph. J.*, **320**, 493.
Lynden-Bell, D., Cannon, R.D., and Godwin, P.J., 1983. *Mon. Not. Roy. Astron. Soc.*, **204**, 87p.
Merritt, D., 1987a. *Astroph. J.*, **313**, 121.
Merritt, D., 1987b. Submitted to *Astron. J.*
Murai, T., and Fujimoto, M., 1980. *Publ. Astr. Soc. Japan*, **32**, 581.
Primack, J. R. 1987. In: *Proceedings of the International School of Physics "Enrico Fermi"*, **92**, ed. Cabibbo, N. Italian Physical Society, Bologna, Italy.
Ostriker, J.P., Peebles, P.J.E., and Yahil, A., 1974. *Astroph. J. Lett.*, **193**, L1.
Richstone, D., and Tremaine, S., 1986. *Astron. J.*, **92**, 72.
Trimble, V., 1987. *Ann. Rev. Astron. Astroph.*, in press.
Turner, E.L., 1976. *Astroph. J.*, **208**, 20.
White, S.D.M., Huchra, J.P., Latham, D., and Davis, M., 1983. *Mon. Not. Roy. Astron. Soc.*, **203**, 701.
Zwicky, F., 1933. *Helvetica Physica Acta*, **6**, 110.