5 Dark Matter in the universe

5.1 Introduction

Consideration of the geometry of the universe as described by General Relativity, combined with the particle and thermal history of the universe outlined via BBN and CMB analyses leads to an overall model that may be dubbed the Hot Big Bang (HBB) model of the universe.

As stated at the beginning of the course, observational tests of the cosmological model change in response to the developing model. The growing realisation that the universe must contain a new, unknown form of matter represents one of the most important developments in cosmology (I would argue that the growing body of evidence pointing to the existence of a cosmological constant is the other development).

Accounting for the properties of dark matter has entered the "standard" set of theoretical descriptions required by any successful model of the evolving universe. This lecture covers the contribution of dark matter to the HBB model, observations of dark matter in the universe and physical theories of dark matter and their consequences.

5.2 The cosmological context – The mass–energy budget

1. Prior to measurements of CMB anisotropies, constraints upon the total baryon density of the universe were derived solely from BBN observations/analyses. The baryon density – expressed as $\Omega_{\rm b}h^2$ – was estimated to be ~ 0.02 (with $h = 0.7$ this implies $\Omega_{\rm b} \sim 0.04$).

2. Total matter estimates were derived from a number of observations probing different physical scales, i.e. dynamical measurements of single galaxies and clusters of galaxies, lensing measurements of the Milky Way, single galaxies and clusters of galaxies, in addition to measurements of X–ray emitting gas in clusters of galaxies. Simple mass–to–light ratio arguments pointed to the existence of dark matter in the sense that it was either extremely faint (or non–emitting) or emitting the bulk of its radiation in an unobserved waveband.

   Though estimates of the total matter density varied as a function of observed scale (see below), on the largest scales the total matter density converged to $\Omega_{\rm M} = 0.2 - 0.3$. Comparison of these observations to BBN results provided the first evidence for the existence of dark matter in the sense to which we refer today, i.e. an unknown or exotic form of matter.

3. Prior to the appearance of improved constraints upon the total geometry of the universe (Point 4), the only constraint upon the total matter–energy content of the universe came from the inflationary model of the big bang. This model predicted a spatially flat universe. The success of the model in explaining other aspects of the observed universe brought considerable favour to the theory. Thus, observations that, on the largest scales, the universe contained less matter than required for flatness resulted in considerable speculation as to where and in what form

the 'additional matter' was located. Inflation is currently not covered in this course. However, the subject remains one of considerable interest[1].

4. In the present day, the strongest constraints upon the total, dark and baryonic matter contributions comes from the measurement of acoustic peaks in the CMB temperature power spectrum (combined with additional cosmological tests). While not wanting to give away the whole story just yet, the most–favoured set of cosmological parameters is the following:

$$\Omega_{\rm DM} + \Omega_{\rm b} + \Omega_{\Lambda} = 0.25 + 0.05 + 0.7 = 1 \tag{1}$$

   This equation may be considered as the *cosmic mass–energy budget*. Not only does it define the geometry of the universe but it sets the relative contribution of each matter/energy constituent to the universe as a whole. It provides a simple but exceptionally useful physical picture of the universe and helps define broad trends in large scale structure and galaxy formation.

5. So what can we currently say about dark matter. Being dark, i.e. not interacting with light, we know that dark matter does not interact via the electromagnetic force. In addition, it is likely that dark matter does not couple to baryonic matter via either the weak or strong nuclear forces. If this were not the case we would expect departures from observed nuclear and BBN observations.

   Dark matter therefore appears to interact solely via gravity. The idealised self-interaction cross-section is zero and dark matter is said to be **dissipationless**, i.e. energy of motion is never lost via interactions. This concept goes a long way toward explaining the observed trends of mass-to-light ratios of objects as a function of physical scale.

5.3 Mass to light ratios

When we observe an object in the universe – a star, a galaxy or a cluster of galaxies – we measure the apparent luminosity. In addition, we can apply additional observations to determine the mass of the object, for example using dynamical, lensing or X-ray arguments. Expressing these observations as the mass–to–light ratio M/L for that object provides a measure of the relative contribution of visible and non-visible matter.

We could start with the Sun. The B-band luminosity of the Sun is $L_{\odot,B} = 4.7 \times 10^{25}$W. The mass of the Sun is $M_{\odot} = 2 \times 10^{30}$kg. Therefore the B-band mass-to-light ratio of the Sun is

$$\frac{M_{\odot}}{L_{\odot,B}} \approx 4.25 \times 10^{4}\ \mathrm{kg\,W^{-1}}. \tag{2}$$

[1] One could also think of this in terms of the evolution of Ω as a function of time. $\Omega = 0, 1$ and ∞ represent three stable positions of Ω. For us to be anywhere close to $\Omega = 1$ today and not exist at some special time in the universe (an extension of the cosmological principle) requires that $\Omega = 1$.

Although this may seem a large amount of mass per unit energy released one must remember that a star releases its energy over a timescale of up to several billion years. The mass-to-light ratio of the Sun provides a basic unit, i.e. one solar mass of stars liberating one solar luminosity of B-band radiation possesses $\langle M/L_B \rangle = 1$.

Over the spectral sequence of stars (OBAFGKM) the mass-to-light ratio varies from $\langle M/L_B \rangle \sim 10^{-3}$ for the bright, massive O stars to $\langle M/L_B \rangle \sim 10^3$ for faint, low mass M stars. The average value of the stellar mass-to-light ratio with 1 kpc of the Sun (the Solar neighbourhood) is $\langle M/L_B \rangle = 4$.

Within 1 Mpc of the Galaxy the total luminosity density of stars is

$$j_{\star,B} = \frac{L_B}{V} = 1.2 \times 10^8 \ \mathrm{L_{\odot,B} Mpc^{-3}}. \tag{3}$$

Assuming that the mean stellar mass-to-light ratio in the solar neighbourhood is representative of this larger volume then the mean stellar mass density is

$$\rho_\star = \langle M/L_B \rangle \, j_{\star,B} \approx 5 \times 10^8 \mathrm{M_\odot Mpc^{-3}}. \tag{4}$$

If we go further and assume that the stellar mass density in the local volume is representative of the universe as a whole then the normalised stellar mass density of the universe is

$$\Omega_{\star,0} = \frac{\rho_{\star,0}}{\rho_{c,0}} \approx \frac{5 \times 10^8 \mathrm{M_\odot Mpc^{-3}}}{1.4 \times 10^{11} \mathrm{M_\odot Mpc^{-3}}} \approx 0.004. \tag{5}$$

So normal stars make up only 0.5% of the density required to generate a spatially flat universe. Consideration of BBN and CMB analyses indicates that $\Omega_b = 0.04$ and we immediately confront the fact that most of the baryons in the universe do not exist in the form of "normal" stars. As we shall see later, additional baryons in galaxies exist as low-mass stars such as brown dwarfs, stellar remnants such as faint white dwarfs and diffuse clouds of hot gas – the warm/hot interstellar medium – WHIM. On the scale of galaxy clusters, the dominant structures in the universe, we shall see that most of the baryons exist in the form of a X-ray emitting plasma.

5.4 Dynamical measurements of galaxies and clusters of galaxies

5.4.1 Galaxies

The Milky Way galaxy is a spiral galaxy. The Sun moves around the Galactic centre on an approximately circular orbit of radius $R = 8.5$ kpc and velocity $v = 220$ kms^{-1}. If the orbit is stable then the outward centripetal acceleration is balanced by the inward acceleration due to gravity, i.e.

$$\frac{v^2}{R} = \frac{GM(R)}{R^2} \tag{6}$$

$$\text{or} \quad v = \sqrt{\frac{GM(R)}{R}}, \tag{7}$$

where $M(R)$ is the mass of the galaxy contained within a radius R.
The surface brightness profile of a galaxy $I(R)$ describes the luminosity emitted per unit surface area. Spiral galaxies typically display exponentially declining surface brightness profiles (Figure 1)such that

$$I(R) = I(R=0)\exp\left(-\frac{R}{R_s}\right). \tag{8}$$

The scale length of the Milky Way is $R_s \approx 4$ kpc. For M31 it is $R_s \approx 6$ kpc. Note that the luminosity enclosed within a radius R is

$$L(< R) = \int_0^R 2\pi\, R\, I(R)\, \mathrm{d}R, \tag{9}$$

and approximately 80% of the total luminosity is contained within 3 R_s. The majority of the luminosity of spiral galaxies is contributed by stars and, assuming that such stars dominate the mass of the spiral galaxy, one expects that at $R \gtrsim 3R_s$ the circular velocity should behave as

$$v(R) \propto \frac{1}{\sqrt{R}}. \tag{10}$$

The velocity behaviour is referred to as Keplerian as Kepler discovered that this relationship describes the velocities of planets in the solar system.

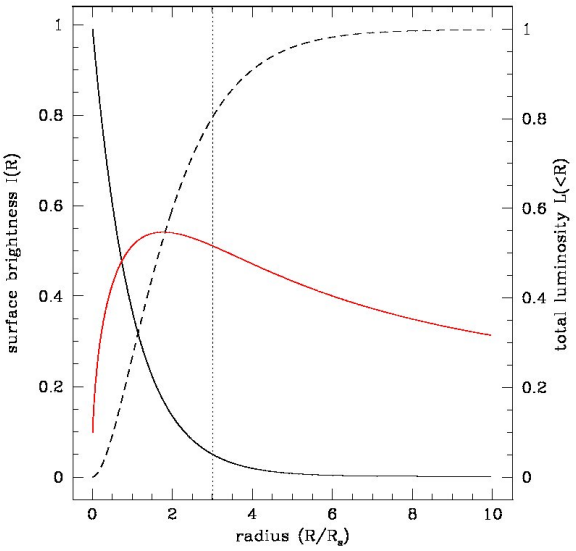

Figure 1: Surface brightness (solid line) and total luminosity within a given radius (dashed line) versus radius for an exponentially declining surface brightness distribution. The red line indicates the shape of the circular velocity curve predicted by this model.

To perform this test upon other spiral galaxies is relatively straightforward: one measures the circular velocity of "test particles" – stars or hydrogen gas clouds – at increasing radius from the centre. Close to the centre of the spiral galaxy the stellar surface brightness is still sufficiently large that one can measure the rotational velocity from the shift of spectral lines in the spectrum of the integrated starlight. However, at larger radii the surface density of stars is not sufficient to generate enough flux to register on a spectrum. One pushes to larger radii either by focussing on the bright optical emission lines generated by individual HII regions or by measuring the shift in the neutral hydrogen 21cm emission line in radio spectra.

A further complication is that the spiral galaxy may not be completely edge on to the observer, with the effect that the radial velocities measured by spectroscopy do not correspond to the true circular velocity. In this case one simply corrects for the estimated inclination angle. Finally, one must also account for the apparent velocity shift between the observer and the spiral galaxy due to either the galaxy peculiar velocity or the Hubble flow. The total radial velocity at a distance R from the centre of the spiral galaxy, as measured from a spectrum, will be

$$v_r(R) = v_{gal} + v(R)\,\sin i, \tag{11}$$

such that

$$v(R) = \frac{v_r(R) - v_{gal}}{\sin i}, \tag{12}$$

where i is the inclination angle and v_{gal} is the peculiar velocity (alternatively $v_{gal} = cz_{gal}$ for a galaxy locked in the Hubble flow).

In 1970 Vera Rubin and Kent Ford published a ground breaking paper on the dynamics of M31 (Andromeda). Rubin and Ford used spectroscopic observations of ionised HII regions (i.e. tracking the spectra shift of the prominent Hα line) within M31 as test particles to constrain the gravitational potential as a function of radius. The Rubin and Ford measurements extended to of order $2-3$ R_s and showed no strong decline in the circular velocity. However, at $2-3$ R_s one can see from Figure 1 that a strong decline is not always expected[2]. However, in 1975 Roberts and Whitehurst extended velocity measurements of M31 to $4-5$ R_s using 21cm observations and confirmed a flat rotation curve. The combined measurements were strongly in disagreement with the prediction of a Keplerian system with the mass concentrated in the central bulge. In fact the mass did not follow the light at all; as the radial light profile decreased at large radii, the mass profile was still increasing – out to the limit of the observations (R = 30 kpc).

Returning to Equation 7 one can write the mass enclosed within a radius R as

$$M(R) = \frac{v^2 R}{G} = 9.6 \times 10^{10} \mathrm{M}_\odot \left(\frac{v}{220 \mathrm{kms}^{-1}}\right)^2 \left(\frac{R}{8.5 \mathrm{kpc}}\right), \tag{13}$$

[2]Note that Figure 1 considers an exponential disk with no other galactic components. M31 has a massive central bulge of stars that significantly adds to the central mass concentration and suppressing the velocity curve beyond the bulge.

where the values scale the mass enclosed to that within the Sun's location within the Galaxy. If we assume a maximum radius for the halo of matter surrounding the Galaxy and a asymptotic velocity value we can determine the halo mass, i.e. taking $v = 220$ kms^{-1} and $R_{halo} = 100$ kpc we obtain $M_{halo} = 1.1 \times 10^{12}\mathrm{M}_{\odot}$. The B-band luminosity of the Galaxy is $L_{MW,B} = 2.3 \times 10^{10}\mathrm{L}_{\odot,\mathrm{B}}$ and therefore

$$\langle M/L_B \rangle_{MW} \approx 50 \langle M/L_B \rangle_{\odot} \left(\frac{R_{halo}}{100 \text{ kpc}} \right). \tag{14}$$

So how big is the halo of the Milky Way? Assuming the outermost globular clusters and satellites such as the Large and Small Magellanic Clouds are bound to the Milky Way then the halo extends to 75 kpc and $\langle M/L_B \rangle_{MW} \approx 40 \langle M/L_B \rangle_{\odot}$. If the Milky Way extends some four times further, i.e. to 300 kpc and halfway to M31, then we obtain $\langle M/L_B \rangle_{MW} \approx 150 \langle M/L_B \rangle_{\odot}$. The implies that the Milky Way contains some 10 to 40 times more mass than can be accounted for by its visible stars. Should the Milky Way be typical in this respect then we can write the normalised mass density in galaxies as

$$\Omega_{gal} = (10 \rightarrow 40)\Omega_{\star} \approx 0.04 \rightarrow 0.16. \tag{15}$$

Although this simple prediction contains many assumptions we are therefore forced to contemplate the fact that some fraction of the dark halos surrounding galaxies is composed of non-baryonic dark matter.

5.4.2 Comments on galactic dark matter

It is interesting to note that the results of Rubin and Ford were not cited by Ostriker and Peebles in their 1973 N–body simulation of the stability of spiral galaxies embedded in massive halos. Ostriker, Peebles and Yahil (1974) presented 'measurements' of spiral galaxy masses as a function of radius drawn from a number of different sources (rotation curves, tidal stability of satellite dwarf galaxies, binary galaxies and timing arguments for the Local Group) in order to claim that galaxy mass–to–light ratios continue to rise out to radii of 1 Mpc. Only in 1985 did Burstein and Rubin present the results for the 'full' data set of 60 spiral galaxies that indicated that the rotation curve phenomenon observed in M31 and the implied mass distribution was common to majority of spiral galaxies observed. As always, a number of caveats are present:

1. Measured velocities could be non-circular and may have been disrupted by gravitational interaction with companion galaxies.

2. Recent star formation may cause one to overestimate $L(R)$.

3. Measuring HII regions is OK for spirals but such gas clouds are not common in elliptical galaxies. PNe are instead employed as tracers.

5.4.3 Clusters of Galaxies

The core of a typical rich galaxy cluster may display a projected extent of approximately 1 Mpc and contain hundreds to thousands of galaxies. Early observations of rich galaxy clusters (e.g. Coma and Virgo) indicated that the member galaxies of such clusters displayed well defined line–of–sight velocity distributions with a typical dispersion $\sigma_v \sim 10^3$ kms^{-1}. The crossing time for a cluster of extent R may be defined as

$$t_{cr} = \frac{R}{\sigma_v} \simeq 1 \left(\frac{R}{1\,\mathrm{Mpc}} \right) \left(\frac{\sigma_v}{10^3\,\mathrm{kms}^{-1}} \right)^{-1} \ \mathrm{Gyr}. \tag{16}$$

Sufficient time therefore exists for the central $1h^{-1}$ Mpc of rich galaxy clusters to relax dynamically within a Hubble time, $t_H \simeq 10h^{-1}$ Gyr. Note that the same cannot be said with confidence for the outer regions (~ 10 Mpc) of such clusters. We can describe the orbital motions of the cluster galaxies by the virial equation

$$\ddot{I} = 2W + 4K, \tag{17}$$

where $\ddot{I}$ is the rate of change of inertia and W and K are the potential and kinetic energies of the cluster galaxies. The cluster is said to be in virial equilibrium if $\ddot{I} = 0$, i.e. it is neither expanding or contracting with time. In this case we can write

$$K = -\frac{W}{2}, \tag{18}$$

where

$$K = \frac{1}{2} M \langle v^2 \rangle \tag{19}$$

is the total kinetic energy contained in the cluster galaxies and

$$\langle v^2 \rangle = \frac{1}{M} \sum_i m_i |v_i|^2 \tag{20}$$

is the mean square velocity weighted by galaxy mass. The potential energy of the cluster may be written as

$$W = -\alpha \frac{GM^2}{r_h}, \tag{21}$$

where $M = \sum m_i$ is the total mass of the cluster, α is a numerical factor of order unity that depends upon the density profile of the cluster and r_h is the radius of a sphere centred on the cluster and containing one half of the total cluster mass. For observed clusters, $\alpha \approx 0.4$ provides a good fit the potential energy. Assuming virial equilibrium then, the typical cluster mass within this central region is

$$M = \frac{\langle v^2 \rangle \, r_h}{\alpha G}. \tag{22}$$

Therefore, determining the mass of a galaxy cluster via virial arguments has reduced to determining the mean square velocity $\langle v^2 \rangle$ and the half mass radius r_h.

Consider the Coma cluster of galaxies – studied by Fritz Zwicky in 1937. The galaxies constituting Coma lie at a mean redshift $\langle z \rangle = 0.0232$. Observation of individual cluster galaxies provides their redshift which contains two components: the cosmological redshift and a peculiar velocity generated by the orbital motion of the galaxy within the cluster. The rest-frame radial line of sight velocity dispersion of a galaxy cluster may be written as

$$\sigma_r = \langle (v_r - \langle v_r \rangle)^2 \rangle^{1/2}. \tag{23}$$

This can be written in terms of the cluster galaxy redshifts by noting that $v_r = cz$ and $\langle v_r \rangle = c\langle z \rangle$. Therefore

$$\sigma_r = \left\langle \left(\frac{cz - c\langle z \rangle}{1 + \langle z \rangle} \right)^2 \right\rangle^{1/2}. \tag{24}$$

The measured value of σ_r for Coma is 880 kms^{-1}. Assuming the galaxy orbits are isotropic then

$$\langle v^2 \rangle = 3\,\sigma_r^2 = 2.32 \times 10^{12} \mathrm{m}^2\mathrm{s}^{-2}. \tag{25}$$

To determine the half mass radius we assume that the unknown dark matter component displays the same distribution as the visible galaxies – which is the same as assuming that the mass–to–light ratio is constant at all radii. The half mass radius can then be computed as the half light radius which for Coma is

$$r_h \approx 1.5 \text{ Mpc} \approx 4.6 \times 10^{22} \mathrm{m}. \tag{26}$$

We can then write the mass of the Coma cluster as

$$M_{Coma} = \frac{\langle v^2 \rangle\, r_h}{\alpha G} \approx 4 \times 20^{45} \mathrm{kg} = 2 \times 10^{15} \mathrm{M}_\odot. \tag{27}$$

The total B-band luminosity of the Coma cluster is $L_{Coma,B} = 8 \times 10^{12} \mathrm{L}_{\odot,\mathrm{B}}$, implying

$$\left\langle \frac{M}{L_B} \right\rangle_{Coma} \approx \frac{250\ \mathrm{M}_\odot}{\mathrm{L}_{\odot,\mathrm{B}}}. \tag{28}$$

However, this is not the complete story as only approximately one percent of the mass of Coma is the form of galactic stars, i.e. $M_{Coma,\star} \approx 3 \times 10^{13} \mathrm{M}_\odot$. A further ten percent exists in the form of hot baryonic gas – the intracluster gas – that can only be observed at X-ray wavelengths, i.e. $M_{Coma,X} \approx 2 \times 10^{14} \mathrm{M}_\odot$.

A mass-to-light ratio of 250 implies that the total mass of the cluster is $250/4 \approx 60$ times the baryonic mass contained in stars. If, for every one baryon in the form of stars, there are ten residing in the X-ray emitting gas, then the ratio of baryonic to total mass in the cluster is 1:6 – roughly the same as the ratio of Ω_b to Ω_M in the universe as a whole.

Fritz Zwicky used these virial arguments in 1937 to argue that Coma was dominated by dark matter. Although determining the virial mass of a galaxy cluster is conceptually very simple, this simple model may be complicated by the influence of number of additional factors, i.e.

1. For a distribution of isotropic orbits, the relationship between radial and orbital velocity is $\langle \sigma_r^2 \rangle = \langle v^2 \rangle / 3$. However, for a more complicated orbit structure, this simple relation will not hold.

2. Mass–to–light ratio is a function of cluster radius.

3. Sub–regions of the cluster may not be relaxed or "virialised". The cluster may be experiencing a local interaction with a 'sub–clump' (see examples from Virgo and recent studies in CL0024+17 below) or, in the outer regions of the cluster the virial assumption does not hold due to ongoing infall of galaxies into the cluster (this may also be expressed by thinking of the cluster as a growing region decoupled from the Hubble flow by the influence of its self gravity).

4. Not all of the light is counted. The contribution of faint cluster galaxies not observed due to the limited sensitivity of a particular observation has to be estimated by assuming a given form for the cluster luminosity function (LF). In a similar manner, the contribution from intracluster light or extended galaxy surface brightness profiles must be accounted for (though are a less serious effect). The contribution from unassociated/projected galaxies along the line–of–sight will also bias the estimated M/L ratio.

5.5 Lensing analyses: Micro, Macro, Weak and Strong

5.5.1 A lensing primer

General Relativity describes the apparent deflection of light by massive objects as seen by a suitably located observer. Assuming the spatial extent of the mass distribution to be small compared to the total light travel distance, the predicted deflection angle is

$$\alpha = \frac{4GM(< b)}{c^2} \frac{1}{b}, \tag{29}$$

where b is the physical impact parameter of the apparent location of the image relative to the mass. Basic geometry and the small angle approximation permit the follow relations to be written (see Figure 2)

$$\theta - \beta = \hat{\alpha}, \;\; \hat{\alpha} = \frac{d_{LS}}{d_{OS}} \alpha, \;\; b = \theta \, d_{OL}. \tag{30}$$

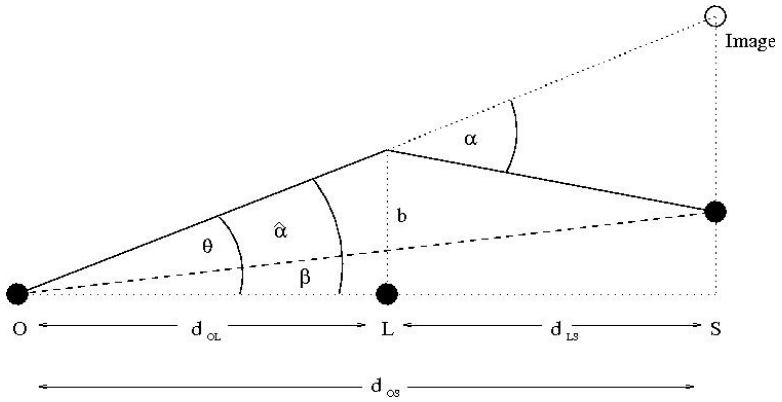

Figure 2: The basic geometry of a gravitational lens.

Note that d_{OL} is the angular diameter distance between the observer and the lens. Substituting for b in Equation 30 generates a quadratic equation for the apparent image position,

$$\begin{aligned} \theta - \beta &= \frac{d_{LS}}{d_{OS}}\,\alpha \qquad (31) \\ &= \frac{d_{LS}}{d_{OS}} \cdot \frac{4GM}{c^2} \cdot \frac{1}{\theta\, d_{OL}} \\ \theta^2 - \theta\,\beta - \frac{4GM}{c^2}\,\frac{d_{LS}}{d_{OS}\,d_{OL}} &= 0. \end{aligned}$$

The solution to this equation indicates that, for a simple point mass lens, there exist two lensed images of the source. For the special case where the observer, lens and source are aligned one obtains the angular Einstein radius applicable for that system, i.e.

$$\theta_E\,(\beta = 0) = \sqrt{\frac{4GM}{c^2}\,\frac{d_{LS}}{d_{OS}\,d_{OL}}}. \qquad (32)$$

The point mass system is circularly symmetric by definition, therefore, θ_E corresponds to a circular lensed image centred on the lens – an Einstein ring. **A few notes:**

1. The separation of the two images is always fairly well represented by θ_E in most strong lensing systems (multiple images observed). Therefore, the observed image separation provides a good general estimate of the total matter along the projected line of sight contained within the two images. Though the universe contains more complex mass distributions (e.g. $M(b)$, ellipticity, etc.) this basic method of estimating total masses remains robust.

2. The magnification of each image follows from solving $\mathrm{d}\theta/\mathrm{d}\beta$ as a function of source or image position. The later discussion will introduce the term *shear*. This arises from the fact that the

image magnification can be expressed as a combination of two terms, *convergence*, or uniform radial expansion/contraction of the image, and *shear*, or image stretching approximately along gravitational equipotentials.

3. Modeling the extended mass distribution responsible for observed lens systems is more complex; lensed images are required on a number of angular scales in order to constrain either the radial or two dimensional mass distribution. The combination of strong plus weak lensing methods for particular systems gives the best results.

5.5.2 Understanding d_{LS}

The angular diameter distance to the source as measured by an observer at the lens redshift, $d_{LS} \neq d_S - d_L$. Instead we must consider the definition of angular diameter distance. In Lecture 2 we defined the angular diameter distance to be

$$d_A = a_e S(r) = \frac{a_0 S(r)}{1+z}, \tag{33}$$

Following Hogg, we write the angular diameter distance between a lens at z_L and a source at z_S as

$$\begin{aligned} d_A(z_L, z_S) &= \frac{1}{1+z_S}([a_0 S(r)]_{z=z_S} - [a_0 S(r)]_{z=z_L}) && (34)\\ &= \frac{1}{1+z_S}\left(\frac{2c}{H_0}\right)\left[\left(1 - \frac{1}{\sqrt{1+z_S}}\right) - \left(1 - \frac{1}{\sqrt{1+z_L}}\right)\right] && (35)\\ &= \frac{1}{1+z_S}\left(\frac{2c}{H_0}\right)\left[\frac{1}{\sqrt{1+z_L}} - \frac{1}{\sqrt{1+z_S}}\right], && (36) \end{aligned}$$

for an EdS universe.

5.5.3 Micro versus Macro

Consider the following table of 'typical' lensing values.

Lens	Mass ($M_\odot$)	d_{OL}	θ_E (arcseconds)
MACHO	1	10 kpc	1.7×10^{-3}
Galaxy	10^{12}	900 Mpc	1.4
Cluster	10^{14}	900 Mpc	14

Note that $d_{OL} = 900$ Mpc approximates to a redshift $z = 0.5$.

5.5.4 Microlensing

Microlensing describes lens events where the expected image separations are not resolvable. Microlensing has been applied most successfully to determine the properties of lensing bodies in the Galaxy acting upon sources located in either the Magellanic Clouds or the Galactic Bulge (i.e. providing a high surface density of potential sources). The likelihood that a microlensing event occurs is most readily expressed as an optical depth, i.e. the integrated probability that at any given time a background star lies within θ_E of a foreground lens. This requires an equation of the form

$$\tau = \int_0^{d_S} P(d_L)\,\mathrm{d}(d_L) \tag{37}$$

where d_L is the distance to the lens and d_S is the distance to the source. Considering a single distribution of sources located in a plane at d_S we can also express τ as the fraction of the source plane covered by intervening Einstein rings (if we expect τ to be small we need not worry about overlaps). For a lens number density of lenses n_L we may write the optical depth as

$$\tau = \frac{1}{\delta\omega}\int n_l\,\pi\theta_E^2\,\mathrm{d}V, \tag{38}$$

where $\mathrm{d}V = \delta\omega\, d_L^2\mathrm{d}(d_L)$. This equation expresses the solid angle subtended by a population of lenses divided by the solid angle of the observations. Substituting for θ_E^2 we can write

$$\tau = \int_0^{d_S} \frac{4\pi G\rho}{c^2}\frac{d_{LS}d_L}{d_S}\,\mathrm{d}(d_L) = \frac{4\pi G}{c^2}\,d_S^2\int_0^1 \rho(x)\,x(1-x)\mathrm{d}x, \tag{39}$$

where $x = d_L d_S^{-1}$ and $\rho = n_L M$ is the mass density of lenses. Note that when writing this equation we have assumed that space is locally flat and that $d_{LS} = d_S - d_L$. Comparing this expression to Equation 37 we see the lensing probability per unit distance $P(x) \propto \rho(x)\,x\,(1-x)$. Assuming ρ to be constant along the line of sight, this integral reduces to

$$\tau = \frac{2\pi}{3}\frac{G\rho}{c^2}\,d_S^2. \tag{40}$$

Paczynski (1986) was the first to perform this calculation and, based upon a reasonable model of the Galactic halo, he computed values of $\tau \sim 10^{-6}$. This calculation formed the basis for a number of microlensing experiments conducted over the next 15 years, that is, given this value of τ, the monitoring of $\sim 10^6$ stars within the LMC, SMC or the Galactic bulge on timescales of order a few hours to a few days should generate up to several tens of microlens events over the course of several years of monitoring. In addition to this, by imaging in two colours, microlensing events may be separated from much more common variable star events by the fact that they are *achromatic* in nature and (as far as seems reasonable) not recurring.

The timescale on which a microlensing event occurs is essentially the time take for the Einstein radius of a the lens to cross a given source. For a point source we may write

$$t_0 = \frac{d_L \theta_E}{v_{LS}} = 0.214 \text{ yr } \left(\frac{M}{M_\odot}\right)^{1/2} \left(\frac{d_L}{10 \text{ kpc}}\right)^{1/2} \left(\frac{d_{LS}}{d_S}\right)^{1/2} \left(\frac{200 \text{ kms}^{-1}}{v_{LS}}\right), \quad (41)$$

where v_{LS} is the relative velocity of the lens and the source. The ratio $d_{LS} d_S^{-1}$ is close to unity for lenses in the Galactic halo and sources in the LMC. Therefore, lens masses in the range 10^{-6} to 10^2 $M_\odot$ produce microlensing timescales ranging from hours to years. The microlensing timescale constrains a combination of M, v_{LS} and d_L. The value of v_{LS} may be determined in a statistical manner by assuming an orbital model for the halo lenses. In addition, the value of d_L may also be determined in a statistical manner by assuming a density distribution of lenses within the halo. In addition to the lens event timescale, each event may be characterised by the peak magnification. At fixed impact parameter, there is a one-to-one mapping between the peak magnification and the angular size of the Einstein radius. Therefore, for a sample of lens events, the distribution of peak magnifications constrains the combination of M and d_L in a statistical manner. Therefore, assuming a lens orbital model, the distribution of peak magnifications plus typical brightening timescale for a sample of microlensing events constrains the combination of the mass plus distance of the lenses.

Microlensing searches were initiated to determine the possible nature of dark matter within the Galaxy. The result of the MACHO search was to confirm that the observed microlensing rate was about 5 times greater than that expected from known Galactic components. The typical mass of the unseen lenses was estimated to be just under one solar mass and they can contribute no more than about 20–30% of the halo mass. This would appear to support some proper motion studies (e.g. Ibata et al.) that point to a large fraction of white dwarf stars in the Galactic halo. Previous MACHO candidates were primordial black holes, Jupiter mass planets, and brown dwarfs (still a viable candidate).

In summary, MACHO surveys have provided good evidence for the nature of about 20% of the dark matter estimated to be present in the Galactic halo. It is perhaps no surprise that the remaining 80% is roughly in the right proportion based upon the dark to baryonic matter ratio given by the cosmological mass budget. An important caveat is that the LMC events that constrain the Galaxy halo mass are accompanied by an excess of bulge lensing events based upon the standard model of the galaxy (bulge/thin disk/thick disk/halo. Therefore, before the MACHO results can be taken as direct evidence for baryonic dark matter, a reassessment of the dynamical components of our Galaxy may have to occur. Therefore, although we may have accounted for the nature of baryonic dark matter in the Galactic halo, the nature of the larger fraction of the missing matter remains elusive.

5.5.5 Lensing by galaxies

The first gravitational lens discovered was Q0957+561. It was originally discovered as a pair of close-separation QSO images displaying remarkably similar spectra. No lensing galaxy was immediately apparent and it is to the credit of the original authors (Walsh, Carswell and Weymann 1979) that the system was flagged as a potential lens. The apparent image separation (5.7 arcseconds) required a large potential – apparently too large for a galaxy to be responsible. Follow–up CCD imaging and spectroscopy revealed a massive elliptical galaxy at a redshift $z = 0.36$ embedded within a cluster of galaxies. The combination of the two effects – the galaxy plus the cluster mass – produced the wide separation image pair. This original system indicates some of the persistent problems inherent in estimating galaxy masses from gravitational lens systems:

1. Modeling the lens system. Not only must the lensing galaxy mass distribution be modeled (i.e. SIS, NSIS, M/L, NFW+M/L) but the effect of external matter distributions must be accounted for, e.g. cluster potential or 'external shear'.

2. The lens model is only as good as the geometry. Not all lens systems studies have source and lens redshifts. Some lenses/images remain too faint even for large telescopes and the mass models reflect these uncertainties.

3. The location of multiply–imaged background sources only provides the galaxy mass at one scale – typically 1–2 arcseconds (approx. 10–20 kpc). Though mass limits are robust, they do not provide detailed information on the dark matter distribution, i.e. the slope of the potential.

Despite these caveats, galactic gravitational lens systems display mass–to–light ratios $M/L \sim 40$, roughly in agreement with the figure obtained from dynamical studies of local galaxies.

5.5.6 Lensing by clusters

In 1986, Lynds and Petrosian noted the presence of large, blue arc–like structures in distant ($z \sim 0.2 - 0.3$) galaxy clusters. It is interesting to note in their original conference abstract detailing observations of the unknown phenomenon that the arcs displayed radii of curvature $\sim 10 - 20$ arcseconds (i.e. very similar to the naive lensing expectation). At first though the gravitational lensing hypothesis was in competition with theories of interacting galaxies or giant star forming regions. However, spectroscopic observations of the arcs (initially inconclusive due to relatively low surface brightness) indicated that the arcs were in fact background galaxies and the lensing hypothesis advanced by some authors was soon confirmed. Initial lens models applied to these systems generated mass–to–light ratios $M/L \sim 300$ – once again, in acceptable agreement with dynamical models (note that we will push the term 'acceptable agreement' much further in the following case study).

The same caveats mentioned above for isolated galaxy lenses also affect cluster gravitational lens models, with a couple of modifications:

1. While a given mass model must be applied to described the overall effect of the cluster, the local effects of individual cluster galaxies may also be very important for the lensing analysis. A model involving more free parameters requires additional constraints if the resulting mass model is to be believed.

2. A cluster lens strongly magnifies a larger region of the background source plane than an individual galaxy. Therefore, it is not uncommon to observed lensed images of many galaxies at slightly varying cluster radii (the variation is largely driven by varying angular Einstein radius as a function source redshift). Though these additional images provide additional constraints upon the mass model, they do increase the observational investment required to obtain all redshifts and thus the full lens geometry.

5.5.7 Weak lensing

Weak lensing describes systems were multiple images of background sources are not formed by a particular lens. However, the gravitational lens effect may still be observed and quantified via the effect of shear (tangential stretching) on the source. In this context, 'weak' may mean that an image is distorted by $\sim 1\%$ along the tangent direction of the gravitational equipotential. Given random variations in galaxy positions and properties (as well as observational errors) the effect must be identified in a statistical manner (i.e. the collective distortion of many galaxies within a small area). Weak lensing effects have been reported by averaging over observations of many individually stacked galaxy deflectors and in individual galaxy clusters. A number of considerations are common to both:

1. The image quality of the survey material is of paramount importance. Weak lensing causes image distortions at the $\sim 1\%$ level. Therefore absolute size of the detector PSF and it's variation with both time and position across the detector must be understood with considerable accuracy.

2. The shear or lensing signal varies as a function of position. Average over background sources spread over too large an area and you will smooth the (already weak) signal – fail to average over sufficient sources and the signal will not be detected. The solution is therefore to image to fainter limits where the surface density of background galaxies will be larger. However, this method reaches an ultimate limit when the sources begin to overlap and the confusion limit is reached (this limit is overcome by increasing the spatial resolution of the imaging device).

3. Some selection criterion must be applied to ensure that the objects deemed to be sources are really behind the objects deemed to be lenses. For individual field galaxies and for clusters

this has been attempted by applying a simple brightness cut, i.e. bright equals foreground, faint equals background. Though robust, the slight overlap in redshift that will result from this approach will 'wash-out' or reduce the measured lensing signal. An improved approach is to incorporate colour (i.e. two photometric bands) information or, in the limit of many bands, photometric redshifts.

4. Weak lensing measures the effect of all matter projected along the line–of–sight. Therefore, the creation of weak lensing maps for structures at increasing redshift is hampered by the accumulated 'noise' of large scale structure variations along the sight line.

5.6 X–ray observations of galaxy clusters

X–ray emission in galaxy clusters arises from collisionally excited high temperature gas within the cluster gravitational potential. The emission is dominated by thermal Bremsstrahlung with some contribution from heavy atomic line emission. Assuming that galaxy clusters are relaxed, virialised systems, the distribution of the X-ray emitting gas reflects the gravitational potential. Dynamical studies of galaxy clusters employ individual galaxies as test particles to probe the cluster potential. X–ray studies employ individual nuclei in the high temperature plasma in the same manner, i.e. equating the thermal energy of the gas to its kinetic energy one obtains,

$$\frac{3}{2}kT = \frac{1}{2}m_{gas}\langle v^2\rangle \tag{42}$$

$$kT = m_{gas}\frac{\langle v^2\rangle}{3} \tag{43}$$

$$kT = \mu m_p \sigma_r^2, \tag{44}$$

where the mean mass per particle is $m_{gas} = \mu m_p$ and $\mu = 0.6$ is the mean molecular weight for a hydrogen fraction of 76%. Inputting the value of $\sigma_r = 880$ kms^{-1} for the Coma cluster one obtains

$$kT_{Coma} = 4.8 \text{ keV} \quad \text{or} \quad T_{Coma} = 5.6 \times 10^7 \text{ K}. \tag{45}$$

These values are reasonably well matched to X-ray observations of the Coma cluster that show relatively cool regions ($kT \sim 5$ keV) mixed with hotter regions ($kT \sim 12$ keV) – possibly resulting from shocked gas heated as the result of merging structure.

In practical terms, the temperature of the X–ray emitting gas is computed by fitting an emission spectrum of given temperature to the observed X–ray spectrum. The spectral intensity resulting from bremsstrahlung emission is a function of both photon energy, the gas temperature and the gas density (squared), i.e. $I(E, T, \rho_g)$. The X–ray luminosity is obtained by integrating over the spectral intensity over the extent of the cluster. Typical X–ray luminosities for massive clusters are large,

e.g. $\log L_X = 43 - 45$ erg s^{-1}. Given current detector technology (e.g. Chandra and XMM), X–ray emitting galaxy clusters are identified to redshifts $z > 1$ by strong, extended emission. Given the emission properties of the X–ray emitting gas, the radial projected mass profile of the cluster may be estimated from independent observations of the X–ray temperature and luminosity. Assuming spherical symmetry, the condition of hydrostatic equilibrium relates the local gas pressure, p, to the density , ρ_g, according to

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{GM(<r)\,\rho_g(r)}{r^2}. \tag{46}$$

Assuming the X–ray emitting plasma to be a perfect gas (i.e. $p = \rho_g\, kT/\mu m_p$), one may express the total gravitating mass within a radius r to be

$$\begin{aligned} M(<r) &= -\frac{r^2}{G\rho}\frac{\mathrm{d}P}{\mathrm{d}r} && (47)\\ &= -\frac{k}{G\mu m_p}\frac{r^2}{\rho}\frac{\mathrm{d}}{\mathrm{d}r}(\rho T) && (48)\\ &= -\frac{k}{G\mu m_p}\frac{r^2}{\rho}\left(T\frac{\mathrm{d}\rho}{\mathrm{d}r} + \rho\frac{\mathrm{d}T}{\mathrm{d}r}\right) && (49)\\ &= -\frac{kT\,r}{G\mu m_p}\left(\frac{r}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}r} + \frac{r}{T}\frac{\mathrm{d}T}{\mathrm{d}r}\right) && (50)\\ &= -\frac{kT\,r}{G\mu m_p}\left(\frac{\mathrm{d}\ln\rho_g}{\mathrm{d}\ln r} + \frac{\mathrm{d}\ln T}{\mathrm{d}\ln r}\right). && (51) \end{aligned}$$

Temperature profiles for distant galaxy clusters may only be estimated for very bright systems – there simply aren't enough photons to study the spatial variation of the spectral emission. Consequently, many researchers have chose to assume that the X–ray emitting gas is isothermal in order to simplify the above equation. However, a subset of X–ray emitting clusters display complex temperature profiles indicative of cooling inflows of material or (at least) a multi–phase X–ray emitting medium. A popular analytic form for the gas density profile is the β–profile,

$$\rho_g(r) = \rho_g(0)\left[1 + \left(\frac{r}{r_c}\right)^2\right]^{-3\beta/2}. \tag{52}$$

This profile has the advantage that the integral required to express the projected density is straightforward. A number of caveats apply to cluster mass estimation from X–ray observations:

1. As mentioned above, the gas may not be in hydrostatic equilibrium – bulk motions or inflows may be present.

2. The accretion of sub–structure (e.g. a galaxy group) and confuse the interpretation of the X–ray emission.

5.7 A case study: dynamics, lensing and X–rays in galaxy cluster MS1054-03

Three papers are provided – Tran et al. 1999, Luppino et al. 1996 and Neumann et al. 2000. These three papers illustrate well the total matter estimation techniques described in previous sections. They also illustrate the possible sources of discrepancy arising from each mass estimation technique, i.e.

1. The dependence of the weak lensing mass upon the unknown source redshift distribution.
2. The discrepancy between the X–ray temperature and the temperature estimate from cluster member dynamics. Tran et al. claim that there exists "good agreement" within the error bars. However, Neumann et al. examine the idea that MS1054 may be experiencing additional shock heating from infalling substructure.
3. The dynamical mass estimate is performed with only 24 cluster members – Combination with the existing data set of 8 members was not performed due to apparent offsets in the two data sets. The small number of confirmed cluster members is not sufficient to compute anything more complex than a simple line–of–sight velocity dispersion. MS1054 is an optically extended cluster – definitely non–spherical. The authors themselves comment that the application of a simple dynamical model to a complex systems could result in mass uncertainties of a factor two.

5.8 The observed M/L ratio as a function of scale

We have seen in earlier sections that the baryonic mass-to-light ratio of starlight is 4. Furthermore, the normalised cosmological mass density associated with starlight is $\Omega_{\star,0} \approx 0.004$. Therefore, to generate a spatially flat universe, the typical mass-to-light ratio of structures traced by star-light (i.e. including visible baryons, non-visible baryons and dark matter) must be $\langle M/L \rangle_U \sim 1,000$. Galaxy clusters, the largest visible structure, come closest to this value, displaying $\langle M/L \rangle_U \sim 250$. This argument is modified by the realisation that, on average, the mass density associated with galaxy clusters does approximate to the average mass density of the universe, i.e. $\Omega_{clusters} \approx 0.25$. The "rest" of the universe is to first order composed of dark energy (see Lecture 8).

So why does the mass-to-light ratio visible structures vary as a function of scale? Put another way, star light is more clustered than dark matter. The variation arises because visible structures trace star formation and star formation requires baryonic dissipation and collapse. Non–baryonic dark matter is thought to be dissipationless and has not collapsed beyond a given scale. More dissipation implies a lower mass–to–light ratio. Therefore, galaxies display lower mass–to–light ratios than galaxy clusters which do not dissipate their gravitational energy as effectively as galaxies. This argument can be extended all the way to the solar system ($M/L = 1$) – a highly relaxed system.

5.9 The physics of dark matter

1. Putting it all together and the baryons just don't add up. The baryonic dark matter contribution is limited by BBN and CMB observations. There must exist at least two types of dark matter – non–luminous baryonic matter and exotic/unknown particles.

2. Baryonic dark matter contribution is limited by BBN and CMB observations. Current candidates display various problems/observational support.

   (a) Low mass and brown dwarf stars. MACHO microlensing results indicated that up to 20% of the Galaxy halo could be composed of bodies of mass $0.1 < M_{\odot} < 1$. This model therefore remains viable and the 'MACHO population' could indeed by completely baryonic. **What has been done to image/detect such lenses directly?**

   (b) White dwarf stars are also supported by the MACHO result and by proper motion studies within our galaxy (e.g. Ibata et al. – various papers). **However, what implications for the stellar IMF does this model have?**

   (c) Free floating Jupiter–mass 'stars' are at the limit of what is permitted by MACHO. No other observational evidence is available.

   (d) Black holes (primordial or otherwise).

   (e) Additional candidates? Failed galaxies, WHIM – what have we learned from the FUSE satellite?

3. Non–baryonic dark matter. Broad categories – hot (relativistic) versus cold (non–relativistic). Non–interacting versus dissipative models.

5.10 Structure formation and the need for dark matter

1. Structure formation from the CMB. Baryon density fluctuations of amplitude 10^{-5} are implied by CMB observations. Linear gravitational theory (i.e. $\delta\rho/\rho \propto (1+z)$) dictates that these fluctuations should have grown to the level 10^{-2} in present universe. As non–linear collapse occurs at fluctuation levels ~ 1, creating structure from baryon fluctuations alone is difficult. However, if a larger dark matter component was not coupled to the photons and interacted with the baryons via gravity alone, then the DM density field could begin to collapse at earlier times and 'seed' the baryon density fluctuations.

2. Hot and Cold dark matter refers to the typical mass–energy of the dark matter particle. Less massive particles remain relativistic for a greater fraction of the age of the universe. Relativistic particles form gravitating structures on much larger scales than non–relativistic particles. The distribution of relativistic dark matter particles will be smoothed within a sphere of constant time of radius $R_* = 2 - 3ct_*$, i.e. the horizon scale corresponding to either

a radiation or matter dominated universe. The most massive coherent structure possible at that epoch is therefore,

$$M = \frac{4}{3}\pi R_*^3 \left(\frac{u}{c^2}\right), \tag{53}$$

where u is the energy density of the sphere. As the universe cools, the particles become sub-relativistic when the temperature drops to

$$kT_* = m_* c^2, \tag{54}$$

where m_* is the mass of the dark matter particle. Recall that within a radiation dominated universe the temperature is related to the age of the universe via $t_* \propto T_*^{-2}$ and that energy density via $u \propto T_*^4$. Combining these relations, one finds that $M \propto T^{-6} \cdot T^4 \propto T^{-2} \propto m_*^{-2}$, with the exact relation of the form

$$M \simeq 10^{15} \left(\frac{30 \text{ eV}}{m_*}\right)^2 M_\odot. \tag{55}$$

This relation permits a broad definition of Hot and Cold dark matter particles. Hot dark matter particles may display masses $mc^2 \sim 30\text{eV} \rightarrow M \sim 10^{15} M_\odot$. whereas cold dark matter particles may display masses in excess of $mc^2 \sim 1\text{MeV} \rightarrow M \sim 10^6 M_\odot$. The rest mass energy of the dark matter particle(s) is closely related to the typical scale formed by the first structures in the universe; hot dark matter models form a "top–down" universe where structure forms on galaxy cluster scales and then fragments to create variety of structure observed today. Cold dark matter models form a "bottom–up" universe where structure forms on globular cluster scales and merges to form the observed large–scale structure of the universe.

Note: that dark matter physics alone does not determine the nature of cosmic structure. The Jeans criterion will be discussed in the later discussion of quantitative measures of large scale structure.

3. The advent of large redshift surveys coupled with large computer simulations in the late 1980's permitted a quantitative comparison between the predictions of a model describing the gravitational clustering of dark matter within an expanding universe and the observed large–scale structure. Though quantitative tests of large–scale structure will be discussed in later lectures, these simulations provided basic constraints upon the nature of non–baryonic dark matter. The universe is strongly clustered, displaying a rich filamentary structure superposed upon a network of galaxy 'voids'. Producing this variety of structure within HDM models was difficult without supplementary assumptions (e.g. cosmic strings). As a consequence, CDM has become a critical ingredient within such structure formation recipes.

4. In 1973, Ostriker and Peebles published a dynamical argument for the existence of massive galaxy halos based upon the required for rotating disk galaxies to be stable. They performed

small N–body simulations representing rotating disk galaxies embedded within static, extended halos. The orbits of the disk population were found to be unstable in the absence of a massive extended halo. Increasing the halo–to–disk mass ratio was found to increase the secular stability of the disk.

5.11 The cutting edge?

Observations of dark matter in the universe have now moved on beyond performing a census of the mass–to–light ratio as a function of scale. The current emphasis of dark matter studies is to provide critical tests of certain aspects of dark matter distributions:

1. Dynamics – When one tries to extend dynamical analyses beyond simple 'weighing' experiments, the simple picture of galaxies and clusters as relaxed systems breaks down. Upon closer inspection galaxies and clusters display a wealth of detailed substructure due to continuing merging activity and infall. The Milky Way provides many examples of local satellite galaxies and chemically distinct (i.e. distinct star formation history) tidal streams. Closer study of distant rich galaxy clusters has demonstrated the presence of dynamically distinct sub–groups within the supposedly relaxed cluster environment, e.g. Ellis et al. and CL0024+17.

2. Lensing – gravitational lens effects, i.e. the morphology of lensed images, are sensitive not only to the mass enclosed within a given physical scale but also to the local gradient of the mass distribution. Therefore, the presence of central (de–magnified) images in individual galaxy lenses and the presence of radial arcs in cluster lenses provides a powerful discriminant of the central structure of dark matter potentials – a key constrain on the extent to which dark matter is really dissipationless. All is not as easy as it sounds however; identification of central images requires very accurate galaxy surface brightness fitting and the cross–section for the creation of radial arcs in cluster lenses is much smaller than the corresponding cross–section for the production of radial arcs – therefore one has to look at many clusters.

3. X–rays – detailed studies of X–ray emitting galaxy clusters appear more likely to probe the dynamical complexity of the gas rather than the detail of the dark matter distribution. Studies of X–ray clusters with Chandra and XMM have revealed complex cooling and heating mechanisms in the central cluster regions and have indicated the presence of substructure associated with shocked, infalling gas. Untangling cluster astrophysics from the dark matter distribution is difficult and is currently addressed using large computer simulations.