Chapter 2
Observational Cosmology

This chapter will present in a more detail the observational foundations of the standard cosmology. We first quickly pass in review the astronomical objects found on various scales. Section 2.1 is concerned with the compact objects that produce the visible photons in the universe either through combustion of nuclear fuel (in stars) or through gravitational collapse (in core collapse supernovae and quasars). The two subsequent sections show how stars are grouped together in galaxies and in clusters of galaxies. In Sect. 2.5 we will review present efforts to elucidate the nature of the dark matter that dominates the mass of the universe. Finally, in Sect. 2.6, we summarize the present measurements of the parameters of the standard model of cosmology i.e., the expansion rate and the various universal densities. A discussion of observations of the CMB photons will be delayed until Chap. 7.

2.1 Stars and Quasi-stars

Among the several thousands of astronomical objects in the sky visible to the naked eye, nearly all are nearby stars ($1\,pc\, <\, R\, <\, 50\,pc$). The exceptions are the Sun ($R = 4.8457 \times 10^{-6}$pc), the Moon, 5 planets, 4 galaxies, and occasional comets. The galaxies are M31 (Andromeda), the Large and Small Magellanic Clouds, and our own Milky Way. While observational cosmology is primarily concerned with supergalactic scales, our understanding of the universe would be impossible without some understanding of stars. In this short section we can only hope to give a superficial account of the astrophysics necessary for cosmology.

Stars begin their lives as diffuse clouds composed primarily of hydrogen and helium. The first generation had the primordial mixture (75% hydrogen and 25% helium by mass) that was produced by nuclear reactions in the early universe. As the clouds contract gravitationally, the negative gravitational binding energy is compensated by increasing the thermal energy of the cloud and by radiating photons (Exercise 2.9). The contraction stops (temporarily) when the core is sufficiently hot to initiate nuclear reactions.[1] At this point, the energy radiated is compensated by

---

[1] Nuclear reactions occur only at high temperature, $T \sim 1\,keV$, because at low temperatures they are exponentially inhibited by the Coulomb barrier between the positively charged nuclei.

the released nuclear energy and the star can exist in a steady state as long as its nuclear fuel holds out.

The first series of nuclear reactions transforms hydrogen into helium. Stars in this phase of nuclear burning are called "main sequence stars." The luminosity of a main sequence star is roughly proportional to the third power of its mass (Exercise 2.9). Since the amount of nuclear fuel is roughly proportional to the mass, the total duration of the hydrogen-burning phase is proportional to M^{-2}. Table 2.1 gives some mass-dependant characteristics of main sequence stars.

After exhausting their hydrogen, helium can be burned to heavy elements in the "giant" phase. The helium is burned to carbon, to oxygen, and, in the most massive stars, to ^{56}Fe, the third most highly bound nucleus.[2]

Some of the heavy elements produced in stars will eventually be dispersed into interstellar space during supernova explosions, the final event in the lives of some stars with masses greater than $\sim 2M_\odot$. The result of this heavy element dispersion

Table 2.1 Characteristics of some representative main sequence stars [36]. Stars are classified according to their spectral type (O, B, A, F, G, K, M: "Oh Be A Fine Gnu, Kiss Me") ranging from hot to cold. Each class is subdivided into subclasses (0-9). The table shows five examples ranging from O5 (heavy, bright, hot) to M5 (light, dim, cold). The first two lines show the correlation between mass and luminosity (heavier-brighter). The lifetime τ on the main sequence corresponds to the time necessary to burn the hydrogen in the stellar core (about 10% of the total hydrogen) at the roughly constant luminosity with bright-heavy stars having a shortest lifetimes. The surface effective temperature T_s and the radius R are related by Stephan's law with bright-heavy stars being hot and big. The next two lines show how the luminosity and surface temperature are reflected in the absolute visual magnitude M_V and color index $B - V$. The last line gives the number density of stars per unit magnitude near our position in the Milky Way. It shows that most stars have relatively small masses

Type	O5	A0	G2	K5	M5
$M/M_\odot$	60	2.9	1.0	0.67	0.21
$L_V/L_{V\odot}$	8×10^5	50	1.0	0.15	0.01
$\tau/10^{10}$ yr	10^{-4}	0.05	1	4	200
T_s (Kelvin)	44500	9520	5860	4350	3240
$R/R_\odot$	12.	2.4	1.0	0.72	0.27
M_V	-5.7	0.6	4.7	7.4	12.3
$B - V$	-0.33	-0.02	0.65	1.15	1.64
n (pc^{-3}mag^{-1})	10^{-8}	10^{-4}	3×10^{-3}	3×10^{-3}	10^{-2}

---

[2] The most highly bound nuclei are ^{58}Fe and ^{62}Ni but they cannot be produced in stars by two-body reactions.

is that later generations of stars will have small admixtures of elements heavier than helium ("metals" according to the astronomical jargon). For example, the Sun started its life 4.5×10^9 years ago with $\sim 28\%$ helium (by mass) and $\sim 2\%$ metals. We see that even recently formed stars have a nuclear composition that is not too far from primordial. The great variety of stars is thus due mostly to differences in stellar masses which vary from $\sim 0.1 M_\odot$ to $\sim 50 M_\odot$ (Table 2.1).

The observable spectrum of photons radiated from stellar surfaces is only indirectly related to stellar structure. Photons in stellar interiors random-walk through the star until they happen to reach a radius where the density is sufficiently low that they escape. The shell from which average photons escape is called the "photosphere." If photons of all wavelengths interacted with the same cross-section, the escape radius would be wavelength independent and stellar spectra would be nearly perfect blackbody spectra reflecting the temperature of the photosphere. The presence near the photosphere of atoms and molecules that are not completely ionized results in a wavelength-dependent photon cross-section. Photons with high cross-sections escape at larger, and therefore colder, radii leading to lower fluxes at the corresponding wavelengths. Two examples of spectra are shown in Fig. 2.1.

In the (rather poor) approximation that the photosphere of a star is a blackbody of a unique temperature, the luminosity of the star is given by Stefan's law, $L = \sigma T_{\mathrm{s}}^4 \pi D^2$ where D is the diameter of the photosphere. A measurement of the flux $f = L/4\pi R^2$ then gives an estimation of the angular size of the star:

$$\Delta\theta = \frac{D}{R} = 2\left(\frac{f}{\sigma T_{\mathrm{s}}^4}\right)^{1/2} , \tag{2.1}$$

where R is the distance to the star.

Of the quantities listed in Table 2.1 only the surface temperature, T_{s}, and "color index," $B - V$, are directly measurable from the shape of the observed photon spectrum. The other quantities can be deduced only if the distance to a star is known. We therefore now turn to the fundamental problem of determining stellar distances.

Apart from the distance to the Sun, which can be accurately determined by radar, stellar distances are extremely difficult to measure. Distances to stars of known diameter can be estimated through (2.1). Diameters can be directly determined for stars in binary systems that happen to be oriented so that the two stars periodically eclipse each other (Exercise 2.8). The diameter of a star whose photosphere is expanding, e.g. pulsing stars or supernovae, can be determined from the photosphere's velocity as determined by the position and shape of its spectral lines (Sect. 2.6.1).

Other than these and other relatively rare exceptions, distances can be directly determined only for nearby stars via their "parallax," i.e. their apparent annual movement with respect to more distant stars that results from the Earth's movement around the Sun. The principle of this technique is illustrated in Fig. 2.2. The excellent angular resolution of the telescope on the Hipparcos satellite [40] that
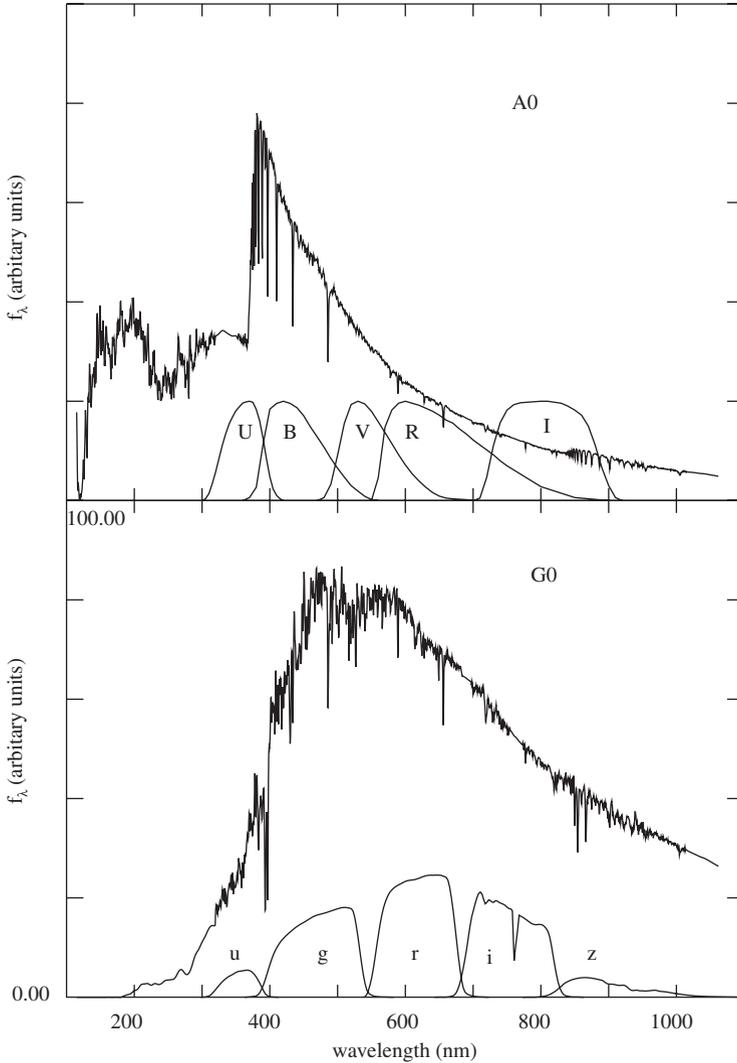
Fig. 2.1 The spectrum (energy per unit wavelength) of an "A0" star (*top*) and a "G0" star (*bottom*) [37]. For A0 stars (like Vega), the photosphere is sufficiently hot that much of the hydrogen is in the $n = 2$ atomic state so the Balmer series of hydrogen lines ($n = 2 \rightarrow n'$) is clearly present. The strong absorption at wavelengths shorter than the "Balmer break" at ~ 370 nm is due to the large cross-section for photo-ionization of $n = 2$ hydrogen and to absorption by closely spaced lines. For G stars (like the Sun), the photosphere is much cooler and there is little flux in the ultraviolet. Superimposed on the A0 spectra are the transmissions of the standard UBVRI filters of the Johnson–Cousins system [38] and on the G0 spectrum the ugriz transmissions (CCD response included) of the SDSS filter system [39]
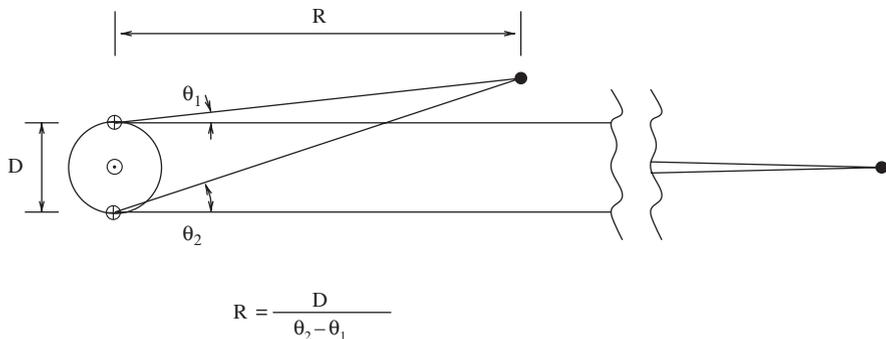
$$R = \frac{D}{\theta_2 - \theta_1}$$

Fig. 2.2 The measurement of the distance R of a nearby star by parallax. The angular separation θ between the nearby star and a distant star undergoes an annual modulation resulting from the Earth's movement around the Sun. For the realistic case of small modulations, the amplitude is $\theta_2 - \theta_1 = D/R$ where $D = 2AU = 3 \times 10^{11}$m is the diameter of the Earth's orbit. The modulation is 1 arcsec for a star at a distance of 1 pc (which explains the origin of the unit)

observed from 1989 to 1993 allowed the measurement of the distances of $\sim 10^5$ stars within ~ 200 pc to a precision of order 10%.

For stars with distances R determined by parallax, the luminosity can be calculated from the measured photon flux $f = L/4\pi R^2$, after corrections are made for absorption by interstellar dust.[3] The calculated luminosity as a function of observed surface temperature is shown in the "color-magnitude" diagram for Hipparcos stars in Fig. 2.3. The luminosity L_V in the wavelength band "V" ($\lambda \sim 550$ nm) is given on an inverse logarithmic scale of "absolute magnitude," M_V:

$$M_V = -2.5 \log L_V + \text{constant.} \tag{2.2}$$

(The magnitude system is explained in Appendix D.) The horizontal scale gives the "color index" $B - V$ which is the difference in magnitude in the "B" band ($\lambda \sim 450$ nm) and the V band. Stars with relatively low surface temperatures have large $B - V$ and stars with relatively large surface temperatures have small $B - V$.

Most of the stars in the color-magnitude diagram are on the diagonal strip that the theory of stellar structure identifies as the main sequence. The heaviest and brightest stars have the highest surface temperatures while the lightest and dimmest stars have the lowest surface temperatures.

Certain post-main sequence stars are on the diagonal extension to the cold side of the main sequence. Especially prominent are the helium-burning "clump" giants near $M_V \sim 0.75$ and $(B - V) \sim 1.0$.

---

[3] The local absorption length of light depends strongly on the line of sight. In the galactic plane it is typically of order 1 kpc for photons with $\lambda \sim 550$ nm [41]. Since absorption by dust is a decreasing function of wavelength, the absorption can be estimated by the "reddening" of a star's spectrum compared to spectra of nearby unreddened stars of the same type.
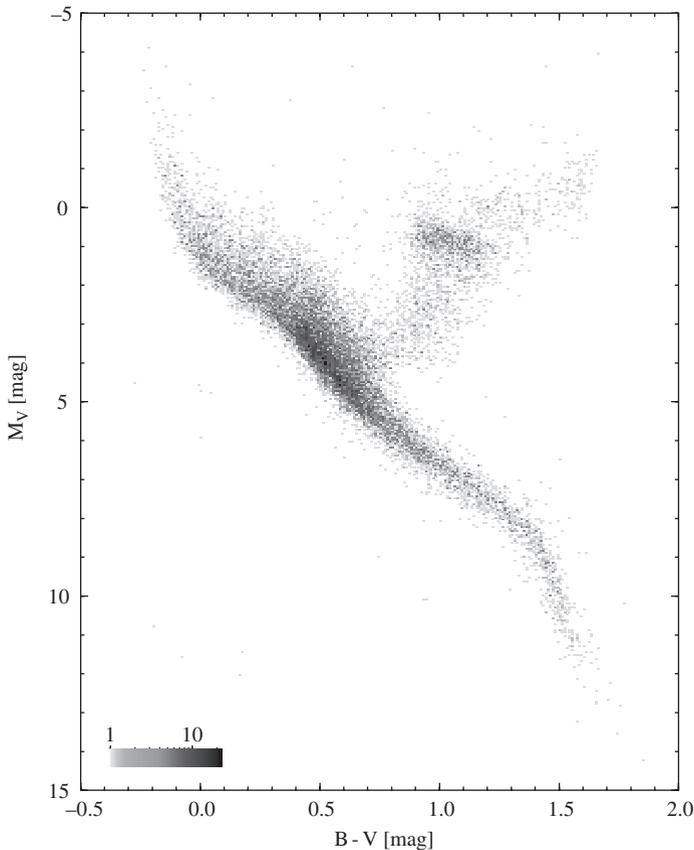
Fig. 2.3 The "color-magnitude" diagram for stars with parallax distances determined with a precision of better than 10% by the Hipparcos satellite [40]. The ordinate is the absolute magnitude in the V band ($\lambda \sim 550$ nm): $M_V = -2.5\log(L_V) +$ constant where L_V is the luminosity in the V band. (The magnitude system is explained in appendix D.) The abscissa is the color index ($B - V$), i.e. the difference in magnitudes between the B ($\lambda \sim 450$ nm) and V bands. The color index is an indicator of a star's surface temperature. Stars that are bright and hot are on the upper left while stars that are dim and cold are on the lower right. The diagonal band corresponds to stars on the main sequence (hydrogen burners). Post-main sequence stars are in the diagonal extension on the cold side of the main sequence. Stars in the accumulation at $M_V \sim 0.75$ and ($B - V$) ~ 1.0 are helium burning "clump" giants. A small number of dead stars (white dwarfs) are present near $M_V \sim 12$, ($B - V$) ~ 0. Courtesy of the European Space Agency

The correlations between *luminosity* and color (Fig. 2.3) generate correlations between *flux* and color for stars grouped at a given distance. Figure 2.4 shows the color-magnitude diagram for stars in the Large Magellanic Cloud (LMC). The flux f_V in the wavelength band V is given on an inverse logarithmic scale of "apparent magnitude," V:

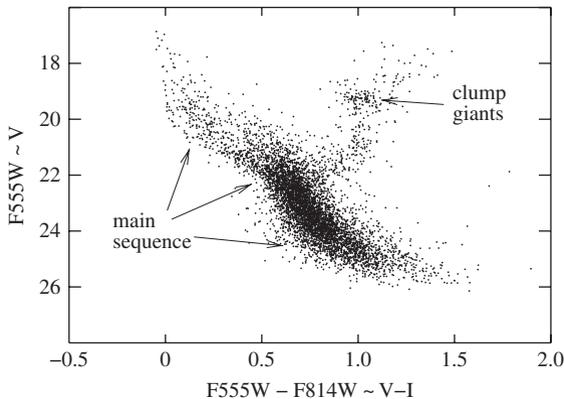$$V = -2.5\log f_V \ + \ \text{constant}'. \tag{2.3}$$

Fig. 2.4 The color-(apparent) magnitude diagram for stars in the Large Magellanic Cloud measured with the Hubble Space Telescope [42]. The ordinate is the apparent magnitude in the F555W band: F555W $= -2.5 \log(f_{555W}) +$ constant where f_{555W} is the flux in the 555W band. (The mean wavelength of F555W, 525 nm, is near that of the standard V band of Fig. 2.3). The abscissa is the color index (F555W−F814W), i.e. the difference in magnitudes in the F555W band and the F814W band ($\lambda \sim 827$ nm). The color index is an indicator of a star's surface temperature. A comparison of the apparent magnitudes of stars in the LMC with the absolute magnitude of the corresponding stars in the Milky Way (Fig. 2.3) allows one to estimate the distance to the LMC

The abscissa is a color index that is an indicator of surface temperature. One is struck by the similarity between this diagram and the Hipparcos color magnitude diagram in Fig. 2.3. The main sequence and clump giants are correctly positioned suggesting that the intrinsic luminosities of the stars in the LMC are nearly the same as those near the solar system. Under this assumption, we can easily estimate the distance to the LMC by using $f_V = L_V/4\pi R^2$. In the magnitude system, this is easy to do because the constants in the magnitude definitions were chosen so that in the absence of absorption, the apparent magnitude is equal to the absolute magnitude of a star at a distance of 10 pc:

$$V = M_V + 5 \log(R/10\,\text{pc}) + A \,, \tag{2.4}$$

where A takes into account absorption. Using the V band magnitudes of the clump giants ($M_V \sim 0.75$ and $V \sim 19.25$ in the LMC) and ignoring correction due to absorption (small in this case), we find

$$R_{\text{LMC}} \sim 10\,\text{pc} \times 10^{0.2(19.25-0.75)} \sim 50\,\text{kpc} \,. \tag{2.5}$$

This turns out to be correct to within 10%.

The traditional "distance ladder" techniques for measuring galactic distances are variations of this calculation of the LMC distance. They all use different types of objects of known luminosity called collectively "standard candles." The technique is accurate only if absorption is negligible or estimable and if the intrinsic difference in

luminosities between the distant and nearby candles can be estimated. In the present example, it is known that the chemical composition of stars in the LMC is slightly different from that of stars in the solar neighborhood. This can make the luminosity of, e.g., clump giants in the solar neighborhood slightly different from those in the LMC and thereby modify distance estimates if not taken into account [43]. We note that the use of different standard candles gives LMC distances that differ by of order 5% [9].

The use of main sequence stars or clump giants as distance indicators is possible only for distances < 1 Mpc beyond which these stars are too dim to be resolved. For distances up to ~ 50 Mpc it is possible to use Cepheid variable stars as distance indicators. Cepheids are a class of post-main sequence stars that have periodic luminosities. Their absolute magnitudes are $M_V \sim -4$ or about 100 times brighter than clump giants and thus identifiable at greater distances. Cepheid luminosities are well-defined functions of their periods as can be seen in Fig. 2.5 for LMC Cepheids. The absolute magnitudes of Cepheids can be determined from the small number that have Hipparcos parallax measurements [44]:

$$M_V = -1.43 - 2.81 \log P \, , \tag{2.6}$$

where P is the period in days. However, in view of the great number of LMC Cepheids, most authors prefer to use LMC Cepheids to calibrate the period–
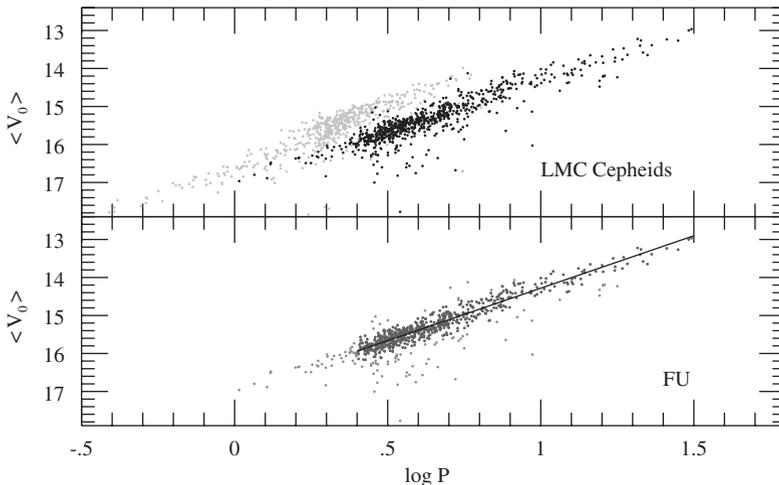

Fig. 2.5 The apparent magnitude, V, $(= V_0)$ as a function of the period P in days for a sample of Cepheids in the LMC from the OGLE collaboration [45]. The *upper panel* shows two types of Cepheids, "fundamental" Cepheids that pulse in the fundamental frequency and "first overtone Cepheids" that have periods one half that of fundamental Cepheids. The *lower panel* shows only the fundamental Cepheids where the best fit is $V = -2.765 \log P + 17.044$. The Cepheid apparent magnitudes can be used to determine the distance to the LMC by using the absolute magnitudes of Hipparcos Cepheids. Alternatively, the apparent magnitudes can be used to find the Cepheid absolute magnitudes by using the LMC distance determined by other methods
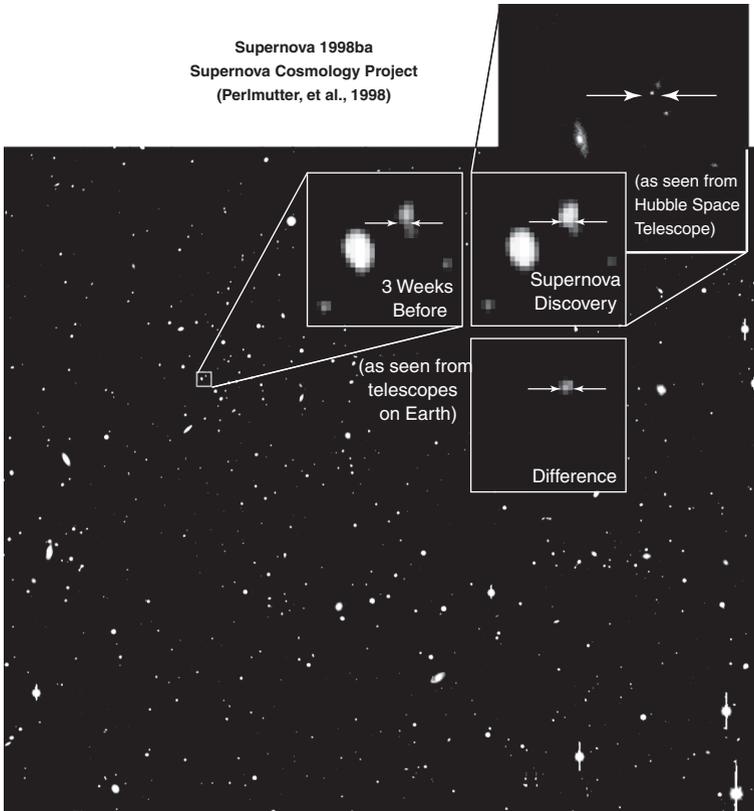
Fig. 2.6 The discovery of a supernova at $z = 0.43$ by the Supernova Cosmology Project [46]. The figure shows a 0.06 deg.^{2} portion of a CCD image of the sky taken with a ground-based telescope. Most objects are distant galaxies. The panels labeled "3 weeks before" and "supernova discovery" show zooms of the large image taken at three-week intervals. The "supernova discovery" panel shows that one of the galaxies has an apparent increase in luminosity and a slight change in shape due to a supernova explosion during the time between the two exposures. In the "difference" panel, the numerically subtracted image shows only the supernova. The final panel shows the same region of the sky taken with the Hubble Space Telescope. Because it does not suffer from the blurring effects of the Earth's atmosphere, the HST image shows the supernova explosion took place at the edge of its host galaxy. Courtesy of Saul Perlmutter

luminosity relation. This strategy, of course, must use a LMC distance determined using other objects, e.g. clump giants or eclipsing binaries (Exercise 2.8).

The brightest "stars" are supernova explosions (Fig. 2.6) that can occur in sufficiently heavy stars at the end of their lives. It is believed that there are two fundamental types of supernovae. The first type consists of "core collapse" supernovae. The progenitor of such supernovae are massive stars ($M > 5M_\odot$) in which the core has burned all the way to ^{56}Fe. As soon as the core has grown to a mass of one

Chandrasekhar mass[4] ($\sim 1.4 M_\odot$), the core collapses to form a neutron star. It is believed that this mechanism is responsible for supernovae classified spectroscopically as types SNII, SNIb, and SNIc. The negative gravitational energy of the resulting neutron star [5] is compensated by the radiation of neutrinos of total energy $\sim 10^{56}$J. Neutrinos are radiated rather than photons because their long scattering length makes it much easier for them to escape through the surrounding envelope of stellar matter than for photons. Since the neutron star mass is always close to the Chandrasekhar mass, these supernovae are "neutrino standard candles." This hypothesis was confirmed by the detection of neutrinos from SN1987a in the LMC. The total energy radiated as photons is at most $\sim 10^{43}$J, much less than the integrated neutrino luminosity. The photon luminosity is mostly powered by the radioactive decay of ^{56}Ni and other radioactive nuclei in the material outside the collapsed core. This is seen in the photon luminosity as a function of time for SN1987a, shown in Fig. 2.7.

The second and brighter type of supernova, SNIa, are the most useful for cosmology. They are believed to be the thermonuclear explosion of carbon–oxygen white dwarfs that are pushed beyond the Chandrasekhar mass by matter falling on them from a binary partner. When the limit is reached, the star starts to implode and the resulting temperature increase results in the explosive burning of the carbon/oxygen to nuclei near ^{56}Ni. Most of the $\sim 10^{44}$J of nuclear energy released is
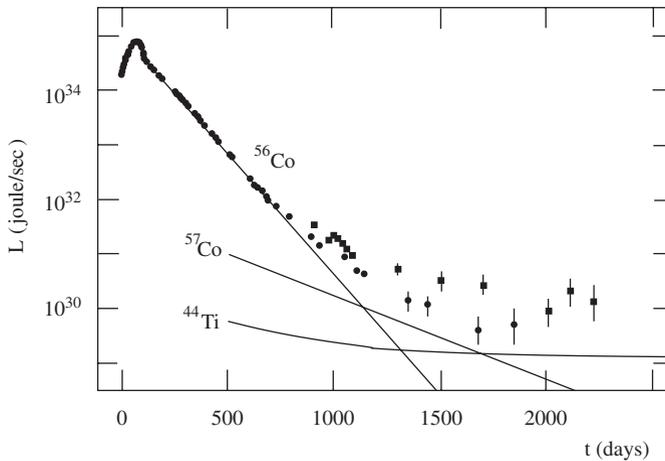

Fig. 2.7 The light curve (total luminosity versus time) for the type II supernova SN1987a in the Large Magellanic Cloud [47]. The *curves* show the energy released by the decay of the radioactive nuclei believed to be responsible for the photon luminosity

---

[4] The Chandrasekhar mass $\sim (\hbar c/G)^{3/2}/m_p^2$ is the largest mass object that can be supported by pressure of a degenerate electron gas against gravitational collapse.

[5] The negative of the gravitational binding energy of the neutron star is $\sim GM^2/r \sim 10^{56}$J for a Chandrasekhar mass neutron star of radius $R \sim 3$ km.

converted to the kinetic energy of the nuclei resulting in the complete disintegration of the star. As with core-collapse supernovae, the photon luminosity is mostly powered by the radioactive decay of ^{56}Ni to ^{56}Co and then to ^{56}Fe. Since the total mass of ^{56}Ni is generally between 40% and 70% of the Chandrasekhar mass, type Ia supernovae are nearly "photon standard candles" making them very useful for distance determinations. They are bright enough to be seen at redshifts near unity. Their only problem is that they are rare, roughly one explosion per galaxy per century remaining visible over a period of a month.

The "light curve" (flux versus time) of a typical SNIa is shown in Fig. 2.8. As illustrated in Fig. 2.9, the luminosities at maximum light are correlated with the color at maximum light and the event time duration. An empirical relation is

$$M_B \sim -19.2 - 1.52(s - 1) + 1.57c \tag{2.7}$$

where $c = (B - V) + 0.057$ at maximum light and the "stretch" s can be roughly defined as the event time duration relative to the mean time duration of SNIa. This relation is the "brighter-bluer" and "brighter-slower" relation. While not completely understood it has plausible explanations. The brighter-bluer relation could reflect absorption in the atmosphere of the supernova or in the interstellar medium of the host galaxy. The brighter-slower relation could reflect the amount of ^{56}Ni produced since increased production implies both increased luminosity and increased opacity, resulting in a longer time scale for photon escape.

Note that Fig. 2.9 includes 6 SNIa that occurred in galaxies or galaxy clusters with Cepheid distances. These six supernovae calibrate the luminosity of type Ia supernovae for their use a standard candles.
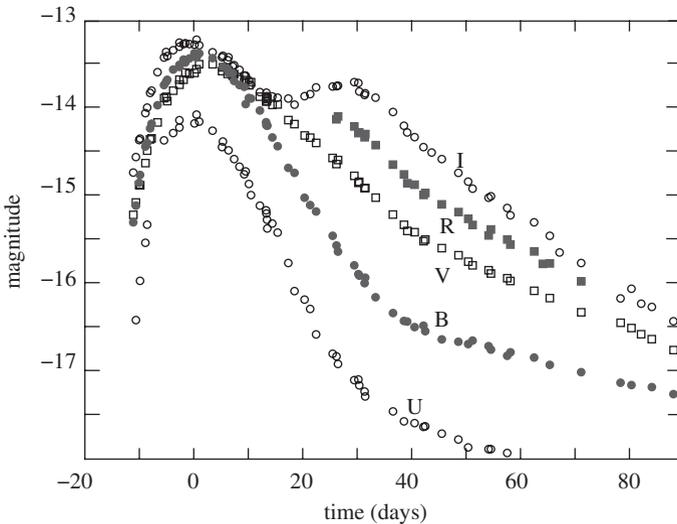

Fig. 2.8 The light curve (apparent magnitude versus time) for a typical type Ia supernova (sn2005cf [48]). The conventional (Vega-based) magnitudes in the Johnson–Cousin bands (UBVRI, as marked) have been offset by (0.70,-0.15,-0.01,-0.18,-0.43) to correspond to AB magnitudes reflecting the true flux. For clarity, the R-band curve has been shown only for $t > 30$ days
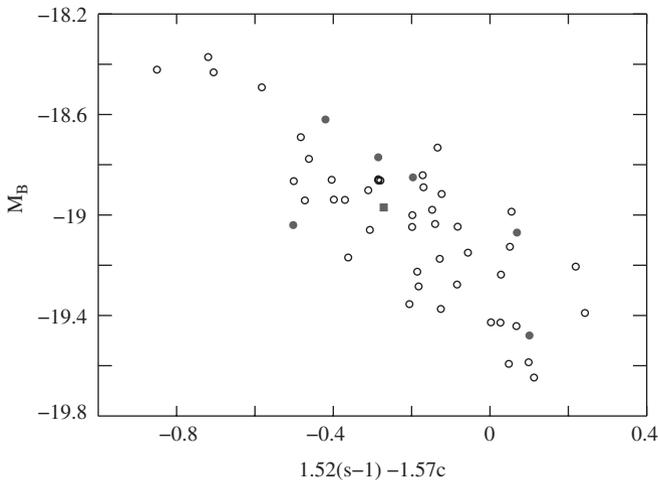
Fig. 2.9 The B-band absolute magnitude, M_B, vs. $1.52(s-1) - 1.57c$, where s is the "stretch" and $c = (B - V) + 0.057$ at maximum light. The *open circles* are a collection of nearby ($0.01 < z < 0.03$) supernovae [2] for which the distance needed to convert apparent magnitudes to absolute magnitudes was calculated from the redshift assuming $h_{70} = 1$. The *filled circles* are supernovae in galaxies with Cepheid-determined distances [49]. The *filled square* is sn2005cf (Fig. 2.8. The figure illustrates the "brighter-bluer" and "brighter-slower" relations for type Ia supernovae

The most luminous known objects are "quasars" or "QSOs" (quasi-stellar objects). The luminosities can be up to 100 times that of a bright galaxy. They are believed to consist of a massive black hole $M > 10^6 M_\odot$ surrounded by gas clouds. If this is true, their source of energy would be the accretion of the surrounding gas by the hole during which the increasing gravitational binding energy of the accreted matter is accompanied by radiation of photons. The spectrum of a typical quasar is shown in Fig. 6.8. The spectrum is non-thermal with a prominent peak due to "Lyman-α" emission by atomic hydrogen ($n = 2 \rightarrow n = 1$) in the surrounding gas. Being so bright, quasars can be found at very high redshifts, the highest to date being $z \sim 6$.

Quasars are important for cosmology for two reasons. First, being among the oldest identified objects, they provide important clues for theories of structure formation. In particular, structure formation must proceed sufficiently rapidly to have produced such objects by $z \sim 6$. The large quasar masses also suggest that they may have played an important role in the formation of the first galaxies, as also suggested by the fact that many galaxies now have central black holes of mass $\sim 10^6 - 10^7 M_\odot$ [50].

Second, quasar spectra show absorption lines due to elements in intervening clouds of gas so they serve as an important probe of the intergalactic medium. The spectrum in Fig. 6.8 exhibits a "Lyman-α" forest of absorption lines on the blueward side of the quasar's own Lyman-α emission. As photons travel to us from the quasar, their energies are degraded by the cosmological redshift. When passing through a cloud of redshift z_{cloud}, a photon will be resonant at Lyman-α if the photons original

energy E_1 satisfies $(z_{quasar}+1)/(z_{cloud}+1) = E_1/E_{Ly-\alpha}$, where $E_{Ly-\alpha}$ is the energy of the Lyman-α transition. The forest of absorption lines thus corresponds to a series of clouds between us and the quasar with each cloud absorbing just those photons whose energies have been degraded by the correct amount. The observed number density of such clouds places important constraints on structure formation theory and permits one to determine the quantity and chemical state of the intervening hydrogen. In particular, we will see in Chap. 5 that the fact that any of the photons blueward of-Lyman-α reach us indicates that most of the hydrogen between us and high-redshift quasars is ionized. This fact is confirmed by the polarization of the CMB (Chap. 7).

2.2 Galaxies

Galaxies are clusters of stars, gas, and dark matter, the largest of which contain 10^{11} order of stars. They have a variety of morphologies, loosely classified as "ellipticals," "spirals," and "irregulars."[6] The Milky Way is a typical spiral galaxy and the characteristics of the solar neighborhood (Table 2.2) give an indication of the environment of a typical galactic disk 8 kpc from its galactic center. We note especially

Table 2.2 The sources of mass and luminosity in the solar neighborhood, 8 kpc from the galactic center [51, 18]. The dark matter is thought to have two components. The first is confined to the galactic disk and has a density that is estimated from the movement of stars perpendicular to the galactic plane. The second, "halo," component is believed to be roughly spherical in shape. Its total mass is deduced from the galactic rotation curve

Component	Mass density $(M_\odot \, pc^{-3})$	Luminosity density $(L_\odot \, pc^{-3})$
Visible stars	0.044	0.067
Dead stars	0.028	0
Gas	0.042	0
Dark matter (disk)	< 0.07	0
Dark matter (halo)	0.003–0.017 (0.1–$0.7 \, GeV \, cm^{-3}$)	0

[6] Elliptical and spiral galaxies are also referred to as "early-type" and "late-type" galaxies. This historical classification is misleading because elliptical galaxies are believed to be in a *later* stage of evolution, being the result of mergers of smaller (spiral) galaxies. A handy way or remembering the correspondence is to note the "elliptical" and "early-type" both begin with the letter "e" while "spiral" and the German word for "late" both begin with the letter "s."

the mass-to-light ratio, $M/L \sim 2.5 M_\odot/L_\odot$, which is, within a factor two or so, typical for the visible parts of most galaxies.

As for a star, the only directly observable properties of a galaxy are its position in the sky, its photon flux, and the shape of its spectrum (Fig. 2.10). The position and widths of spectral lines in the spectrum and their variation over the surface of a galaxy can be used to determine the galactic redshift as well as velocity dispersion of the stars. Luminosities can be determined only if the distance to the galaxy is known, as is the case for a few galaxies nearer than ~ 50 Mpc through the observation of Cepheid variable stars in the galaxy. The resulting luminosities have calibrated empirical relations between luminosities and velocity dispersions, the Faber–Jackson or Fundamental Plane relation for elliptical galaxies and the Tully–Fisher relation for spiral galaxies (Fig. 2.11). Once calibrated, these relations can be used to measure distances out to $R \sim 200$ Mpc.

Of course, the distance to any galaxy can be deduced from the redshift of its spectrum if one assumes a value of H_0. Over the last decade, the large redshift surveys of galaxies, SDSS and 2dFRS, have produced $\sim 10^6$ redshifts, giving a rather complete picture of galaxies out to $z \sim 0.2$. Figure 2.12 shows the luminosity distribution of SDSS galaxies at $z \sim 0.1$. The number of galaxies per unit volume and per unit luminosity interval is well described by the "Schechter" distribution:

$$\frac{\mathrm{d}n_{\mathrm{gal}}}{\mathrm{d}V\mathrm{d}L} = \frac{\phi_*}{L_*}\left(\frac{L_*}{L}\right)^\alpha \exp(-L/L_*) \qquad \alpha \sim 1 , \tag{2.8}$$

where ϕ_*, L_*, and α are constants. The distribution is particularly well measured at $z \sim 0.1$ by SDSS (Fig. 2.12) who gives, for luminosities measured near $\lambda \sim 560$ nm

$$\phi_* = (0.511 \pm 0.016) \times 10^{-2} h_{70}^3 \mathrm{Mpc}^{-3}$$
$$L_* = (2.45 \pm 0.02) h_{70}^{-2} \times 10^{10} L_\odot . \qquad \alpha = 1.05 \pm 0.001 . \tag{2.9}$$

The factors of h_{70} comes from the use of Hubble's law to determine galactic distances, $R = H_0^{-1} z$. The logarithmic divergence in the luminosity distribution at small L makes the total number density of galaxies ill-defined. However, the low-luminosity galaxies produce little light so the total light output per unit volume is (fortunately) finite. For SDSS galaxies, the integral of the Schechter distribution gives:

$$J = \int_0^\infty \mathrm{d}L \phi_* \exp(-L/L_*) \sim L_* \phi_* = 1.29 \, h_{70} \times 10^8 \, L_\odot \mathrm{Mpc}^{-3} \tag{2.10}$$

While the number of galaxies is not well defined, 95% of the light comes from galaxies brighter than $\sim 10^9 L_\odot$. The integral then gives the number density of bright galaxies:

$$n_{\mathrm{gal}} = \int_{10^9}^\infty \mathrm{d}L \frac{\phi_*}{L} \exp(-L/L_*) = 0.015 h_{70}^3 \, \mathrm{Mpc}^{-3} . \tag{2.11}$$
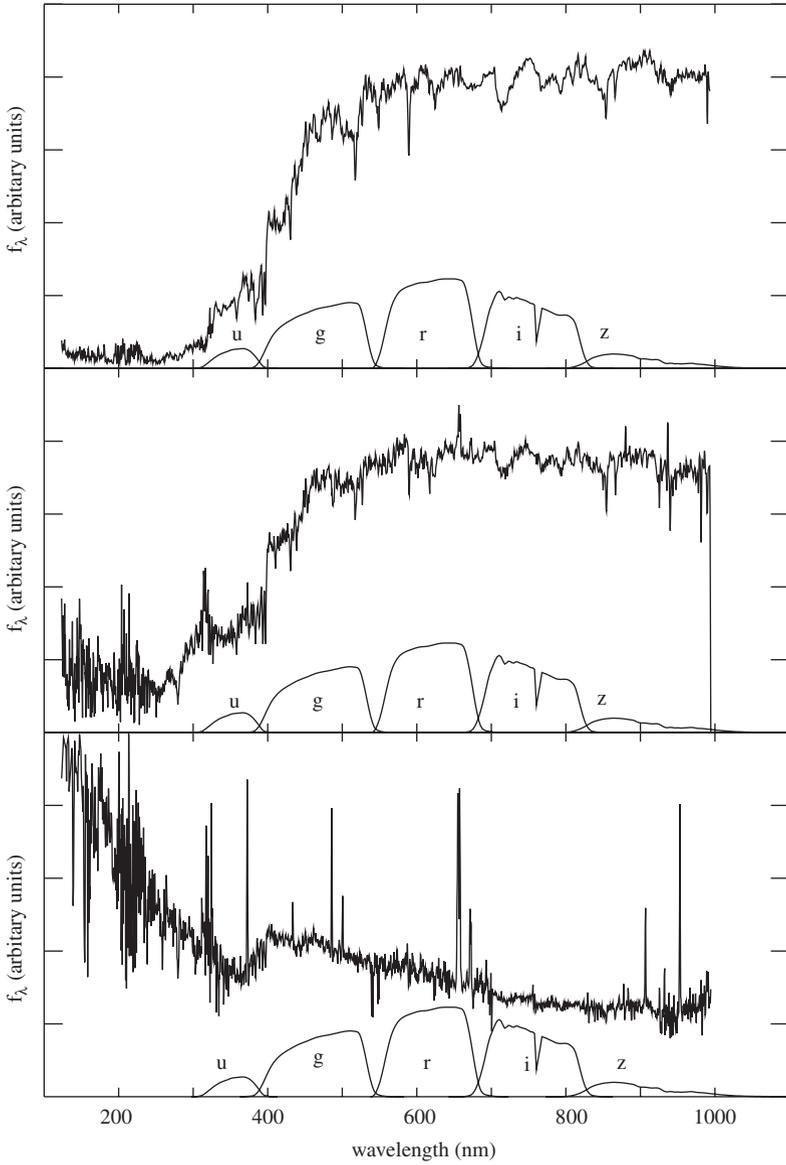
Fig. 2.10 Typical spectra (energy per unit wavelength) of elliptic (*top*), spiral (*middle*), and star-forming (*bottom*) galaxies [52]. Being the sum of its stellar spectra, galactic spectra show the same spectral lines as those in Fig. 2.1. Galactic emission lines (present in the lower spectrum) are due to hot interstellar gas that is excited by ultraviolet photons from massive stars or QSOs within the galaxy. Also shown are the bandpasses of the SDSS filters [39]
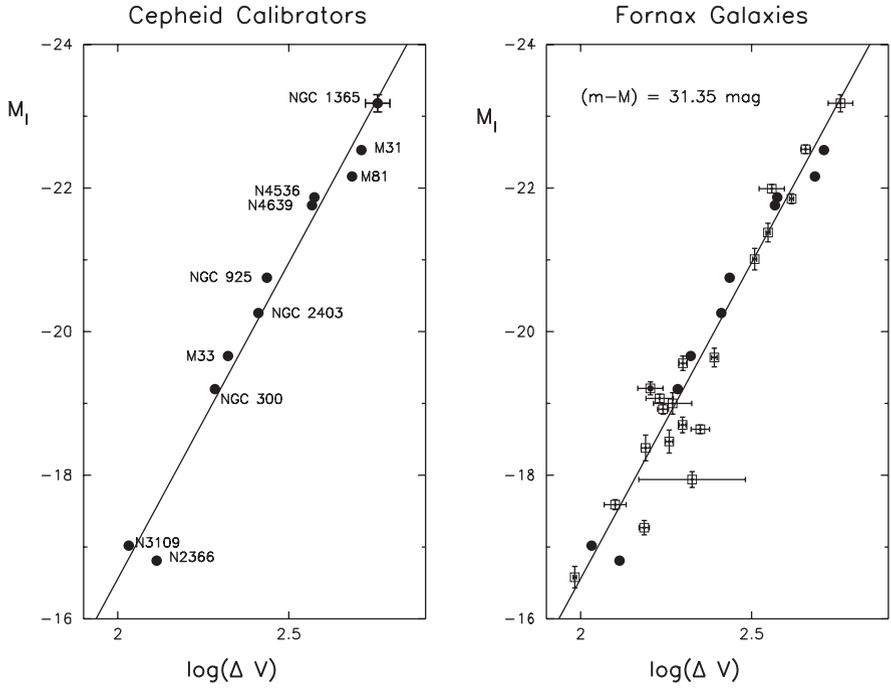
Fig. 2.11 The I-band absolute magnitude of spiral galaxies with Cepheid distances [53]. The magnitude is a linear function of $\log \Delta v$ where Δv is the velocity dispersion in $\mathrm{km\,s^{-1}}$
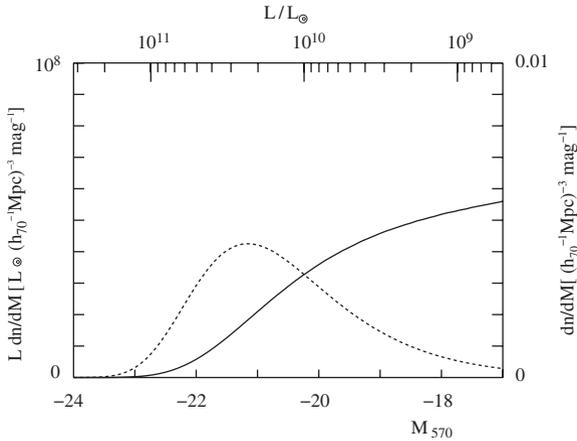

Fig. 2.12 The number density (*solid line*, right scale) and luminosity density (*dashed line*, left scale) of SDSS galaxies at redshift $z \sim 0.1$ [54]. The lower scale gives the absolute magnitude at $\lambda \sim 560\,\mathrm{nm}$ ($M_{570\odot} = 4.76$) while the upper scale give the galactic luminosity in units of $L_{\odot}$

The total mass in stars for a given galaxy can be deduced from the galaxy's spectrum with the help of stellar evolution models. For SDSS galaxies [55], a galaxy with $L = L_*$ has on average a total stellar mass of $6 \times 10^{10} M_\odot$ corresponding to $M/L = 6/2.45 = 2.45$. Multiplying this by J from (2.10) and dividing by the critical density gives the contribution of stars to the universal density

$$\Omega_{stars} = 0.0023 \ . \tag{2.12}$$

This is a factor 6 greater than the contribution of atomic hydrogen gas (HI) contained in galaxies deduced from galactic 21 cm emission [56]

$$\Omega_{HI} = 3.5 \times 10^{-4} \ . \tag{2.13}$$

There is a comparable amount of molecular hydrogen.

The visible parts of galaxies bathe in an extensive "halo" of dark matter. The density profile of these halos (and of clusters of galaxies) is often described by a "NFW" profile proposed by Navarro, Frenk and White [57] on the basis of n-body simulations of galaxy formation:

$$\rho(r) \ \propto \ \frac{1}{r(1 + r/r_s)^2} \ , \tag{2.14}$$

where r_s parameterizes the size of the halo. The density has a singularity at $r = 0$ though this does not lead to a singularity in the potential or mass. The density falls like r^{-3} for $r \gg r_s$ so the integrated mass of the halo rises logarithmically with r and the total halo mass is undefined by the density profile. However, the halo mass can be defined as the mass within a radius where objects are effectively bound. Structure formation theory informs us that within this radius the mean density is of ~ 200 times the mean universal density (Eq. 7.17).

Because of its simplicity, a distribution of mass that is often used as a first approximation for galaxy clusters and galaxy halos is that for an "isothermal sphere":

$$G\rho(r) = \frac{\sigma_v^2}{2\pi r^2} \ , \tag{2.15}$$

where σ_v is the one-dimensional velocity dispersion of objects bound in the potential. The r-independent rotational velocity for circular orbits is $\sqrt{(2)}\sigma_v$. Having an r dependence that is intermediate between the small and large r dependence of the NFW profile, the isothermal profile gives results that over a large range of r resemble those of more realistic profiles.

Galactic mass profiles can be derived from the velocities of stars or gas clouds orbiting a galaxy (Exercise 2.7). From Newtonian dynamics, the circular velocity $v(r)$ at a distance r from a galactic center is given by $v(r)^2 \sim GM(r)/r$ where $M(r)$

is the mass within r. Figure 2.13 shows the rotation curve, $v(r)$, of a spiral galaxy. The velocity is that of clouds of atomic hydrogen determined by the Doppler shift of the 21 cm hyperfine line of the hydrogen ground state. The rotation curve is flat for distances beyond the visible radius of the galaxy, indicating $M(r) \propto r$. Since the rotation curve stays flat out to the last measured point, $r \sim 30\, h_{70}^{-1}$ kpc, one can deduce only a lower limit on the galactic mass or mass-to-light ratio: $M/L > 20\, h_{70}\, M_\odot/L_\odot$. The factor of h_{70} comes from the use of Hubble's law to convert observed angular sizes $\Delta\theta$ to physical sizes $r = R/\Delta\theta = H_0^{-1} z/\Delta\theta$.

The study of weak gravitational lensing of background galaxies ($z > 0.5$) by foreground galaxies ($z < 0.2$) has overcome some of the limitations of rotation curve measurements. As discussed in detail in Sect. 3.8, the trajectories of light from
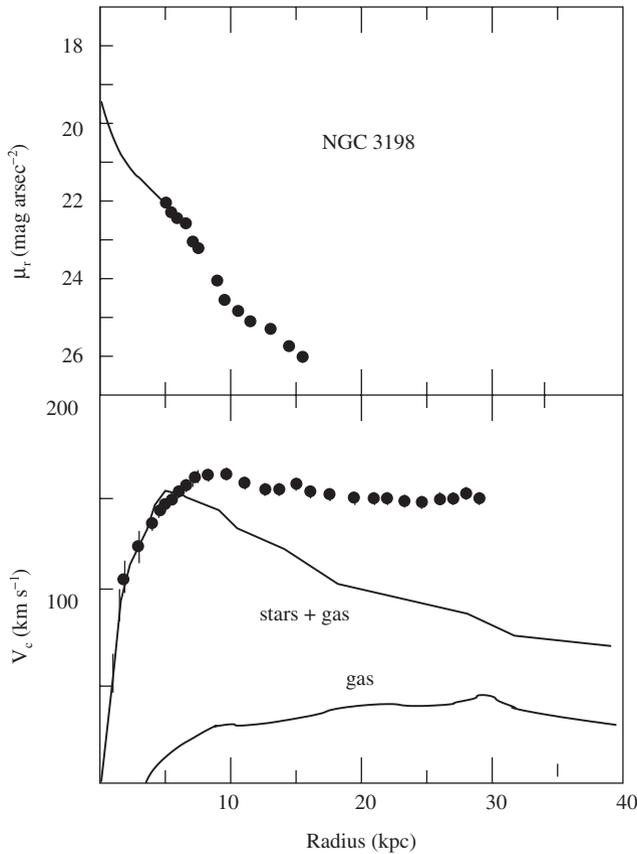

Fig. 2.13 The rotation curve of the galaxy NGC3198 [58]. The *upper panel* shows the luminosity distribution indicating that most of the luminous matter is concentrated at distances less than 5 kpc from the center. The *lower panel* shows the measured rotation curve. The curve remains flat far beyond the luminous radius indicating the presence of dark matter far from the center. The curve labeled "stars and gas" shows the rotation curve that would be expected if the mass distribution followed the luminosity distribution with a mass-to-light ratio of $3.8 M_\odot/L_\odot$. The horizontal scale supposes $H_0 = 75\,\mathrm{km\,s^{-1} Mpc^{-1}}$

background galaxies are bent in the gravitational field of a foreground galaxy. This results is a small distortion of the shape of the background galaxies that can be used to deduce the mass of the foreground galaxy (Eq. 3.113). For a given background galaxy, the effect is not measurable because the unperturbed shape is not known. However, if the shapes of many background galaxies are measured, one can find systematic stretching of the galaxies in the direction tangent to circles centered on the foreground galaxy. Since the effect of generally less than 1% for a given galaxy, it can only be seen by averaging over many foreground galaxies. This has been done with SDSS galaxies [55] yielding the density profiles shown in Fig. 2.14. The total galactic mass can be found by integrating the profile out to a radius where the mean density is a factor ~ 200 greater than the mean universal density. The data in Fig. 2.14 imply that a galaxy with $L = L_*$ is surrounded by a halo of mass $1.4 \times 10^{12} M_\odot$ [55], corresponding to a mass-to-light ratio

$$M/L = 140/2.45 \sim 60 . \tag{2.16}$$

Multiplying this by J (2.10) and dividing by the critical density gives the contribution of galaxies to the universal density

$$\Omega_{\text{gal}} = 0.054 . \tag{2.17}$$

Since $\Omega_{\text{M}} \sim 0.27$, this means that $\sim 80\%$ of the mass is in intergalactic space.
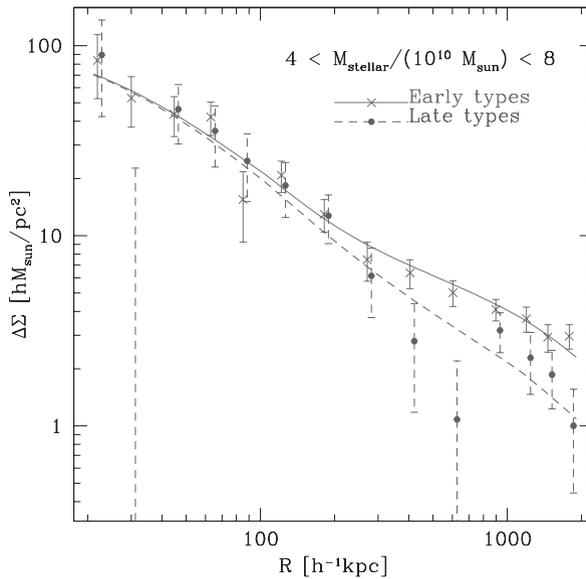

Fig. 2.14 Galactic mass surface densities (density integrated along the line of sight) as a function of distance from the galactic center as deduced from the weak lensing by SDSS galaxies of background galaxies [55]. The lenses are SDSS galaxies with $4 < M_{\text{stars}}/10^{10} M_\odot < 8$. The surface densities are deduced from the tangential shear given by Eq. (3.113). Figure courtesy of Rachel Mandelbaum

Finally, we note that one would expect that the ratio of total galactic baryonic mass to total galactic mass would be roughly the universal value, $\Omega_b/\Omega_M \sim 0.17$. For an L_* of mass $1.4 \times 10^{12} M_\odot$ this would imply a baryonic mass of $2 \times 10^{11} M_\odot$ to be compared with the estimated stellar mass $6 \times 10^{10} M_\odot$. This means that within galaxies, about 25% of baryonic material has been transformed into stars.

2.3 Galaxy Clusters

Galaxies are often gravitationally bound in groups. Our own galaxy is a member of the "Local Group" containing ~ 30 small galaxies plus three large spirals, the Milky Way, M31 (Andromeda), and M101. The largest known bound structures are "rich clusters" of galaxies that can contain thousands of galaxies and up to $10^{15} \, h_{70}^{-1} M_\odot$ in volumes of a few Mpc^3. This mass corresponds to the mean mass contained in a sphere of radius $\sim 10 h_{70}^{-1} Mpc$ so regions within $1 \, Mpc$ of the centers of rich clusters have over-densities of order 1000. The nearest rich cluster is the Virgo Cluster at a distance of $\sim 20 \, Mpc$. The number density of clusters as a function of their mass is shown in Fig. 2.15. There are $\sim 10^{-5}$ clusters per $(h^{-1} Mpc)^3$ with
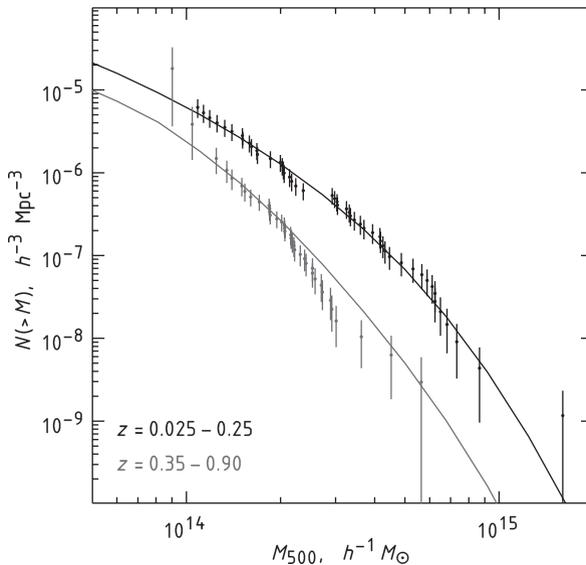

Fig. 2.15 The number density of clusters with masses greater than M_{500} as a function of M_{500} [100] as determined by Chandra X-ray observations. (M_{500} is the mass contained within the radius where the mean cluster density is a factor 500 greater than the universal density.) The data is shown for two redshift ranges. There are a factor ~ 3 more clusters at low redshift than at high redshift meaning that clusters are still in the construction process. The evolution of the number of clusters with redshift is consistent with that expected for a ΛCDM universe with $\Omega_M \sim 0.0.34 \pm 0.08$. (The ratio between the high- and low-redshift densities would be ~ 20 in a $\Omega_M = 1$ universe.)

a mean mass $\sim 10^{14} h^{-1} M_\odot$ which allows us to estimate the total mass contained in such large clusters:

$$\Omega_{\text{cluster}} \sim \frac{10^9 h^2 M_\odot \text{Mpc}^{-3}}{3H_0^2/8\pi G} \sim 0.003 \tag{2.18}$$

Comparing with Ω_{gal} (2.17), we see that $\sim 5\%$ of galaxies are in large clusters.

Clusters like that in Fig. 2.16 are rather ill-defined when viewed as simple collections of galaxies. They become much more distinct when observed through their X-ray emission. The X-rays are produced through bremsstrahlung by electrons in the ionized intergalactic gas. It is believed that this gas constitutes the majority of the baryonic mass of clusters.

Cluster masses can be most easily estimated by measuring the velocity dispersion of the member galaxies and then applying the virial theorem (Exercise 2.6). More modern techniques use the X-ray temperature (Exercise 2.10) or the shapes of background galaxies that are deformed by the gravitational lensing action of the cluster (Sect. 3.8). The most massive clusters have mass-to-light ratios of order 200
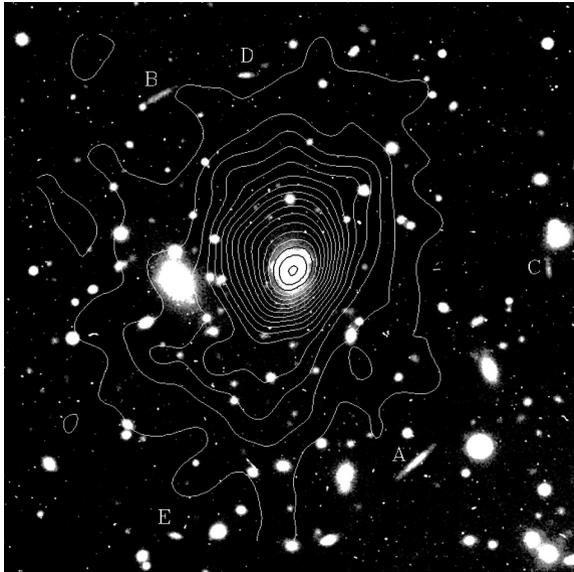

Fig. 2.16 The galaxy cluster RXJ1347.5-1145 [59]. The contours show the level of emission of X-rays by the ionized intergalactic gas in the cluster. The cluster is a "gravitational lens" for galaxies behind it. The images of distant lensed galaxies are in the arc-shaped objects marked with the letters A, B, C, D, and E. Courtesy of S. Schindler

$$\frac{M}{L} \sim 200 \, h_{70} \, M_\odot / L_\odot \; . \tag{2.19}$$

The mass of the gas in a cluster can be estimated from the observed X-ray flux (Exercise 2.10). It is an order of magnitude less than the total mass [99]:

$$\frac{M_{\text{gas}}}{M_{\text{total}}} \sim 0.12 \, h_{70}^{-3/2} \Rightarrow \frac{M_{\text{baryons}}}{M_{\text{total}}} \sim 0.16 \, h_{70}^{-3/2} \; . \tag{2.20}$$

where the second form includes the mass of stars and gas in individual cluster galaxies. This factor of 6 between total mass and baryonic mass is perhaps the best evidence for non-baryonic dark matter.

An important prediction of CDM models is that the dark matter is made up of particles that have very weak non-gravitational interactions. This was confirmed by observations of the "Bullet cluster" [60] shown in Fig. 2.17. Superimposed on the optical and X-ray images are mass contours as derived from gravitational lensing of background galaxies. One sees that the Bullet cluster actually consists of two clusters that recently passed through each other. The CDM and the galaxies have only gravitational interactions so these two components more or less retain their form after passing through the gravitational potential of the other cluster. On the of other hand, the ionized gas consists of particles that scatter on each other. This causes "frictional" forces on the gas that slows it down with respect to the CDM. The result is a displacement of the gas and dark matter, as clearly demonstrated in the figure. This constitutes "visual proof" of the existence of non-collisional dark matter.

2.4 Large-Scale Structure

At scales above that of galaxy clusters, large redshift surveys have shown that galaxies appear to be grouped along filamentary "walls" sometimes surrounding great "voids" containing few visible galaxies. Figure 2.18 shows a "slice" of the sky as seen in by SDSS [61]. The largest walls or voids have sizes of order $100 \, h_{70}^{-1}$ Mpc. These structures are not bound and, at the present epoch, are still participating in the universal expansion though at a reduced rate because of their self-gravity. As discussed in Sect. 1.2.6, these density fluctuation are unlikely to form bound structures in the future because of the acceleration of the expansion caused by dark energy.

At scales > 100 Mpc the density is relatively uniform (Fig. 7.4). The characterization of the small fluctuations above these scales gives important clues about the matter content of the universe. This will be discussed in Chap. 7.
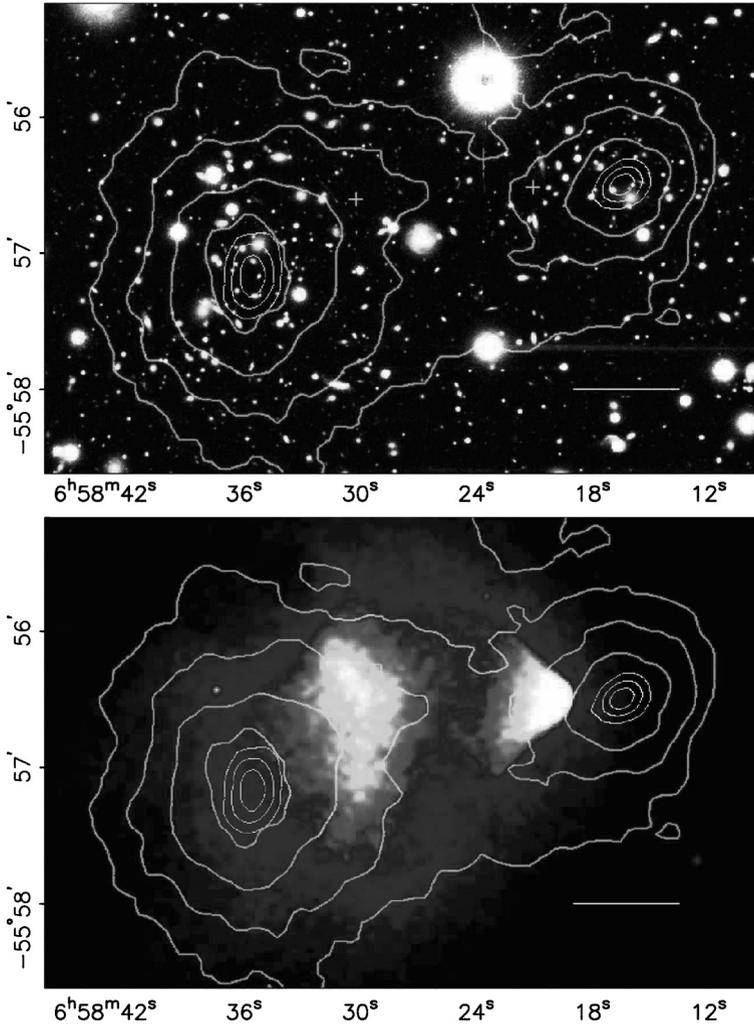
Fig. 2.17 Optical (*top*) and X-ray (*bottom*) images of the "Bullet Cluster" [60]. Mass contours as derived from a lensing analysis are shown. The cluster actually consists of two mass concentrations (shown by the lensing mass contours) with two concentrations of gas (shown by the X-ray image) displaced from the centers of mass. The interpretation is that the two clusters recently passed through each other causing the collisional gas to be displaced with respect to the non-collisional dark matter and galaxies. This provides "visual proof" of the existence of weakly interacting dark matter
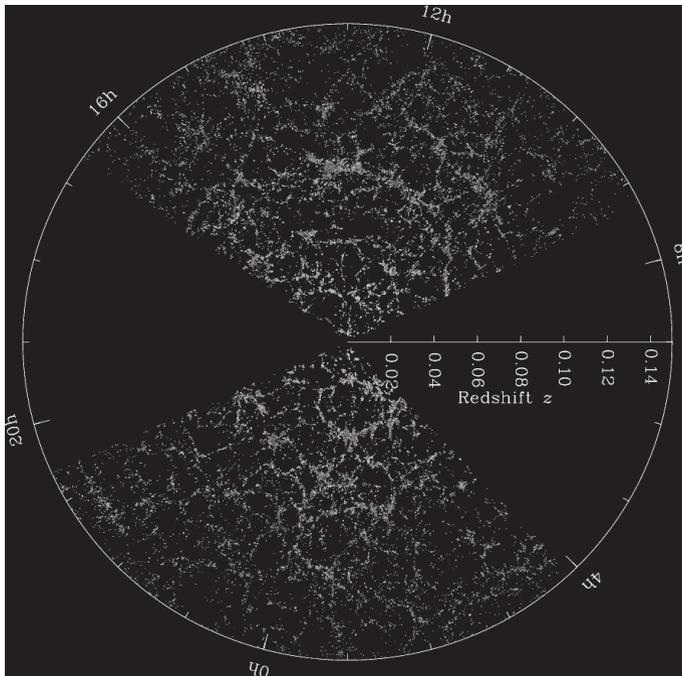
Fig. 2.18 Slices through the SDSS three-dimensional map of the distribution of galaxies. Each slice covers about 7 hours (105 deg.)in right ascension and 2.5 deg. in declination. Our galaxy is positioned at the intersection of the northern and southern slices. The radial coordinate is given as the redshift, z. "Walls" and "voids" are seen up to sizes of $\Delta z \sim 0.02 \Rightarrow \Delta R \sim\sim 86\, h_{70}^{-1}$ Mpc. Courtesy of the Sloan Digital Sky Survey

2.5 Dark Matter

It is clear that there will always be doubts about the ΛCDM cosmological model as long the dark matter has not been identified. Here, we review efforts to detect the two favored non-baryonic candidates, WIMPs and axions. We also review limits on the numbers of dark astrophysical objects (MACHOs) in the galactic halo, which appear to have eliminated most such objects as candidates for galactic dark matter. Finally, we discuss the possibility that dark matter consists of cold gas.

2.5.1 WIMPs

Since the 1980s, WIMPs (weakly interacting massive particles) have been the standard cold dark matter (CDM) candidate. Supersymmetric extensions of the standard model of particle physics (invented to solve problems unrelated to dark matter) predict the existence of such particles. The fact that they have not yet been seen at accelerators suggests that they must have a mass $m_\chi > 30$ GeV [62]. The particles would

have been thermally produced in the early universe (Chap. 6) yielding a cosmological abundance inversely proportional to their annihilation cross-section. Supersymmetric models contain many free parameters yielding relic densities within a few orders of magnitude on either side of the critical density.

Today, WIMPs would be expected to inhabit the halos of spiral galaxies like our own. From the galactic rotation velocity, one can estimate the local density to be about $0.3 \, \text{GeV} \, \text{cm}^{-3}$ [17]. The orbital velocities of objects trapped in the Galaxy are of order $250 \, \text{km} \, \text{s}^{-1}$ so the local WIMP flux is of order $10^7 \times (1 \, \text{GeV}/m_\chi) \, \text{cm}^{-2} \, \text{s}^{-1}$.

Goodman and Witten [63] suggested that these WIMPs could be detected via the observation of nuclei recoiling from WIMP-nucleus elastic scatters. Galactic WIMPs with masses in the GeV range have kinetic energies in the keV range so we can also expect nuclear recoils in the keV range. The rate is proportional to the elastic WIMP-nucleus scattering cross-section which depends on the parameters of the particle physics model. Typical values of the supersymmetric WIMP-nucleon cross-section are of order $10^{-44} \, \text{cm}^2$, corresponding to a very weak interaction.

WIMP scatters can be observed with "calorimetric" techniques (Fig. 2.19). Unfortunately, it is difficult to distinguish WIMP events from events due to beta or gamma radioactivity (also shown in the figure). Statistically, a signal from WIMPs can be isolated through the expected $\sim 5\%$ seasonal modulation of the event rate [17, 64]. This modulation is due to the fact that while the Solar System moves through the (isotropic) WIMP gas, the Earth's motion around the Sun alternately adds or subtracts from the WIMP-detector velocity. Alternatively, certain detectors can distinguish nuclear recoils from the Compton-electron background, e.g. hybrid cryogenic detectors that detect two types of excitations, phonons and ionization or
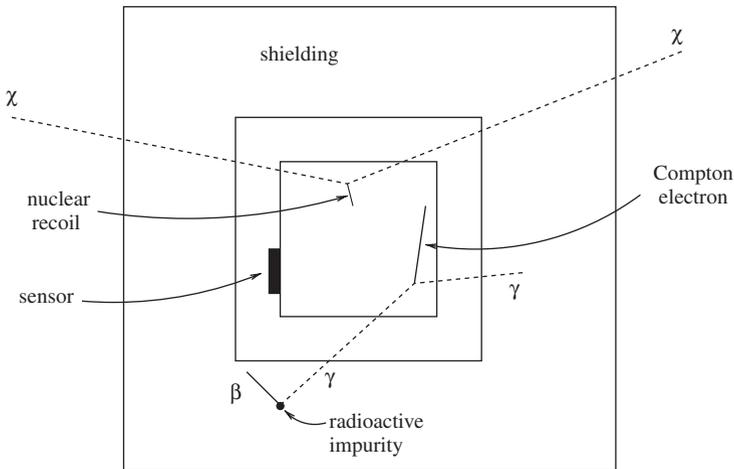

Fig. 2.19 A generic "calorimetric" WIMP detector surrounded by its shielding. The galactic WIMP χ enters the detector, scatters off a nucleus, and leaves the detector. The recoiling nucleus creates secondary excitations (e.g., scintillation light, ionization, phonons) that can be detected by the sensor. Also shown is a background event due to the ambient radioactivity yielding a Compton electron in the detector

phonons and scintillation. The ratio of the two signals is different for nuclear recoils and Compton electrons, allowing background rejection.

The present generation of experiments using hybrid cryogenic calorimeters [65] and xenon ionization–scintillator calorimeters [66] have backgrounds that make them sensitive to the supersymmetric WIMP candidates with the highest scattering cross-sections. The present limits are shown in Fig. 2.20. Particles with $m_\chi \sim 100\,\text{GeV}$ must have cross-sections on nucleons less than $\sim 10^{-43}\,\text{cm}^2$ to have escaped detection. The cross-section limits are very weak for particles with $m < 1\,\text{GeV}$ because such light particles create nuclear recoils with energies too small to be detected.

We note that the limits shown in Fig. 2.20 are in nominal conflict with an experiment using NaI detectors [67]. This experiment reported an annual modulation of the event rate similar to what one would expect for WIMPs. The source of the disagreement is not understood.

Besides direct detection, it is possible to detect WIMPs "indirectly" through the detection of particles produced in present-day WIMP-antiWIMP annihilation. While WIMP annihilation ceased in the early universe because of the universal expansion, it started up again once the WIMPs became gravitationally bound in galactic halos. These annihilations are a source of cosmic-ray photons, electron/positrons and proton/antiprotons and the observed fluxes are sometimes interpreted as being due to WIMP annihilation [68] but it has so far proved impossible to rule out astrophysical sources for the particles.
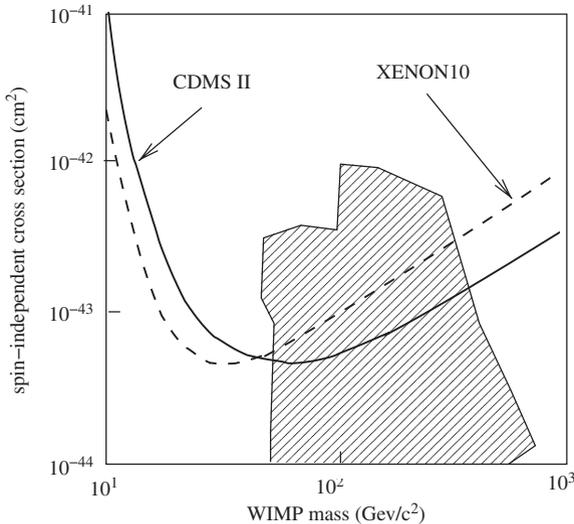

Fig. 2.20 Current limits on WIMP mass and WIMP-nucleon scattering cross-section from direct-detection experiments using xenon [66] (*dashed line*) and germanium [65] (*solid line*). The region above and to the right of the curves are excluded. The shaded regions show the region of mass-cross-section space favored by supersymmetric models
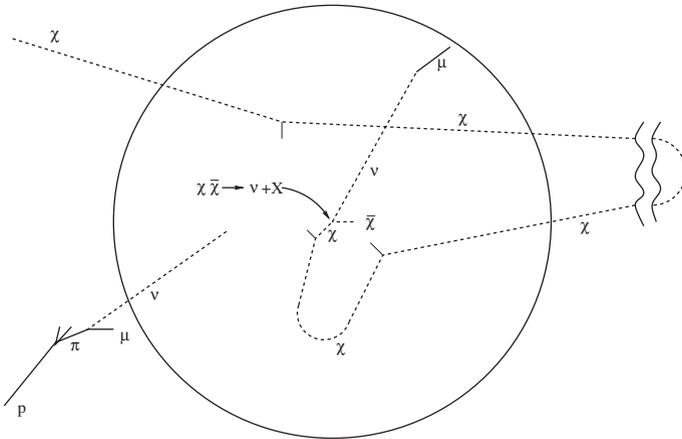
Fig. 2.21 The capture of a WIMP χ in the Earth. If the WIMP loses sufficient energy in a collision with a nucleus, the WIMP's velocity will drop below the escape velocity placing it in an orbit intersecting the Earth. The subsequent collisions will eventually thermalize the WIMP in the center of the Earth. An annihilation with a thermalized anti-WIMP may lead to the production of neutrinos that can be detected at the surface. As shown, the background for such events comes from neutrinos produced by decays of pions and kaons produced by cosmic rays in the Earth's atmosphere

The annihilation rate is further enhanced inside material objects like the Sun or Earth. This is because it is possible for WIMPs to be trapped in such objects if, while traversing the object, the WIMP suffers an elastic collision with a nucleus (Fig. 2.21). If the scatter results in a WIMP velocity below the object's escape velocity, the WIMP will find itself in an orbit that passes through the object. After repeated collisions the WIMP will be thermalized in the core. In the case of supersymmetric dark matter, the trapping rate in the Sun is sufficiently high that the concentration of WIMPs reaches a steady state where trapping is balanced by either annihilation (for high-mass WIMPs) or by evaporation (for low-mass WIMPs). For the Sun, the dividing line between low mass and high mass is $\sim 3\,\text{GeV}$ so accellerator limits ($m > 30\,\text{GeV}$) suggest that WIMPS should be annihilating inside the Sun.

The only annihilation products that can be seen emerging from the Sun or Earth are, of course, neutrinos. The flux of such neutrinos can be calculated for a given WIMP candidate and the flux compared with that observed in underground detectors. The observed flux is entirely understood as being due to the decay of cosmic-ray pions and kaons in the Earth's atmosphere (Fig. 2.21). Certain supersymmetric WIMPs would give a higher flux and are thus excluded [69, 70].

Because the observed flux of neutrinos is due to an unavoidable background, the only improvements in the limits from these techniques would come from the observation of a small excess of neutrinos coming from the direction of the Sun or center of the Earth. The most reasonable possibility is to search for upward-going muons coming from ν_μ interactions in the rock below a detector. Calculations [71] indicate that a 1 km^2 detector with a muon energy threshold of $\sim 10\,\text{GeV}$ would be needed

to observe a statistically significant solar signal for typical supersymmetric dark matter. Efforts in this direction are underway by instrumenting the Mediterranean [72] or the Antarctic Continental Glacier [73] to observe Cerenkov light produced by muons.

2.5.2 Axions

Axions [74] are hypothetical light scalar particles invented to prevent CP violation in the strong interactions.[7] They would have been produced in the early universe via both thermal and non-thermal mechanisms and might produce near-critical relic densities if they have masses in the range $m_a \sim 10^{-5}$ eV to $\sim 10^{-3}$ eV. It was also recently emphasized [29] that axions with smaller masses, $m_a \sim 10^{-8}$ eV, could give a near critical density in some models that require the application of anthropic selection.

Axions act as cold dark matter and should be present in the galactic halo. The most popular detection scheme for galactic axions is based on the expectation [75] that axions can "convert" to a photon of frequency $\nu = m_a c^2 / h$ in the presence of a magnetic field. If a microwave cavity is tuned to this frequency, the axions will cause an excess power to be absorbed (compared to neighboring frequencies). If the halo is dominated by axions, the predicted power is small, about 10^{-21} W for a cavity of volume 3 m^3 and a magnetic field of 10 T. Since the axion mass is not known, it is necessary to scan over the range of interesting frequencies. Pilot experiments [76, 77] have produced limits on the local axion density about a factor 30 above the expected density. Experiments are now in progress to search for axions of masses in the range $m_a \sim 10^{-5}$ eV to $\sim 10^{-3}$ eV at the required level of sensitivity [78].

2.5.3 MACHOs

Unless the current estimates of the baryon density, $\Omega_b = 0.0456 \pm 0.0015$ are incorrect, baryons cannot account for all of the dark matter. Nevertheless, baryons could account for *galactic* dark matter if they are in a form that neither absorbs nor emits light in significant quantities. The various possibilities have been reviewed in [79]. The simplest way to hide baryons is to place them in compact objects that either do not burn (e.g., brown dwarfs) or have ceased to burn (e.g., white dwarfs, neutron stars, black holes). Such dark objects in a galactic halo are called MACHOs for MAssive Compact Halo Objects.

Brown dwarfs have masses $< 0.07 M_\odot$ making them too cool to burn hydrogen. They were originally the favored MACHO candidates because they completely avoid constraints based on the production of background light or pollution

---

[7] Such a violation would produce a permanent nucleon electric dipole moment in violation of experimental limits [18].
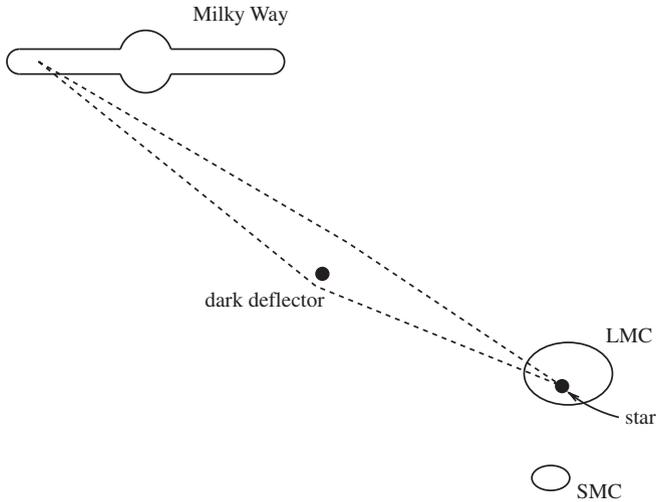
Fig. 2.22 A schematic of the lensing of a star in the Large Magellanic Cloud (LMC) by an unseen object in the galactic halo. While the two images cannot be easily resolved, the combined light from the two images gives a transient amplification of the light from the star as the unseen object passes near the line-of-sight. The light curve for a point source is shown in Fig. 2.23

of the interstellar medium with heavy elements through supernova explosions [80]. MACHOs in the form of black holes could also avoid the Ω_b constraint if they are "primordial" black holes that were produced before the epoch of nucleosynthesis.

Paczyński [81] suggested that MACHOs could be detected through their gravitational lensing of visible background stars in the Large Magellanic Cloud (LMC) (Fig. 2.22). This small galaxy is at a distance of 50 kpc from Earth.

The theory of gravitational lensing will be presented in Sect. 3.8. As the MACHO approaches the line-of-sight to the background star, two images are formed. It turns out that in the case of lensing by stellar objects in the galactic halo, the angle separating the two images is small (<1 milliarcsec). This type of gravitational lensing is therefore referred to as "microlensing." Earth-bound telescopes cannot resolve the two images because atmospheric turbulence smears images so that stellar objects have angular sizes of order 1 arcsec. The only observable effect is therefore a transient increase of the total observed light as the MACHO moves toward and then away from the line-of-sight. The amplification is

$$A = \frac{u^2 + 2}{u\sqrt{u^2 + 4}} \, , \qquad (2.21)$$

where u is the distance of closest approach of the (undeflected) line-of-sight to the deflector in units of the "Einstein radius" $R_E = \sqrt{4GMLx(1-x)/c^2}$ where L is the observer–source distance, Lx is the observer–deflector distance, and M is the MACHO mass.

The amplification is greater than 1.34 when the distance to the line-of-sight is less than R_E. This amplification corresponds to a reasonable observational threshold since photometry can "easily" be done to better than 10% accuracy. At a given moment, the probability, P, of a given star being amplified by more than a factor 1.34 is just the probability that its undeflected light passes within one Einstein radius of a MACHO:

$$P \sim n_{\mathrm{MACHO}} \, L \, \pi \, R_E^2 \, , \tag{2.22}$$

where n_{MACHO} is the mean number density of MACHOs between us and the LMC and L is the distance to the LMC. If the entire halo is comprised of MACHOs, the density of MACHOs is roughly $n_{\mathrm{MACHO}} \sim M_{\mathrm{halo}}/(ML^3)$ where M_{halo} is the total halo mass out to the position of the LMC. Using the expression for the Einstein radius, we find that P is independent of M and determined only by the velocity of the LMC:

$$P \sim \frac{G M_{\mathrm{halo}}}{L c^2} \sim \frac{v_{\mathrm{LMC}}^2}{c^2} \, . \tag{2.23}$$

The LMC is believed to orbit the galaxy with $v_{\mathrm{LMC}} \sim 200 \, \mathrm{km \, s^{-1}}$ (corresponding to a flat rotation curve out to the position of the LMC). In this case, P is of order 10^{-6}. More detailed calculations give $P = 0.5 \times 10^{-6}$ [82].

Since the observer, star, and deflector are in relative motion, a sizable amplification lasts only as long as the undeflected light beam remains within the Einstein radius. The light curve for a star near the center of the Milky Way lensed by a star in the Milky Way disk is shown in Fig. 2.23. The time scale of the amplification is the time t_E for the deflecting object to cross one Einstein radius with respect to the observer and source. For the lensing of stars in the LMC by objects in our halo, the relative speeds are of order $200 \, \mathrm{km \, s^{-1}}$ and the position of the deflector is roughly midway between the observer and the source ($x \sim 0.5$). The mean t_E is then

$$t_E \sim \frac{R_E}{200 \, \mathrm{km \, s^{-1}}} \sim 75 \, \mathrm{days} \sqrt{\frac{M}{M_\odot}}. \tag{2.24}$$

The observed t_E distribution can therefore be used to estimate the mass of the MACHOs if one assumes that they are in the galactic halo.

Three groups, the MACHO, EROS and OGLE collaborations have published results of searches for events in the directions of the LMC and the SMC (the neighboring Small Magellanic Cloud). The limits are shown in Fig. 2.24. The lack of events with $t_E < 15$ days allowed the two groups to exclude as the dominant halo component objects with masses in the range $10^{-7} M_\odot < M < 10^{-1} M_\odot$ [84]. These limits exclude as a major halo component brown dwarfs of masses $\sim 0.07 M_\odot$. Furthermore, the EROS collaboration observed no events with $t_E < 400$ days and this excludes MACHOs with masses less than $\sim 30 M_\odot$. The MACHO collaboration has, however, observed 13 events of mean duration ~ 50 days [85]. If interpreted
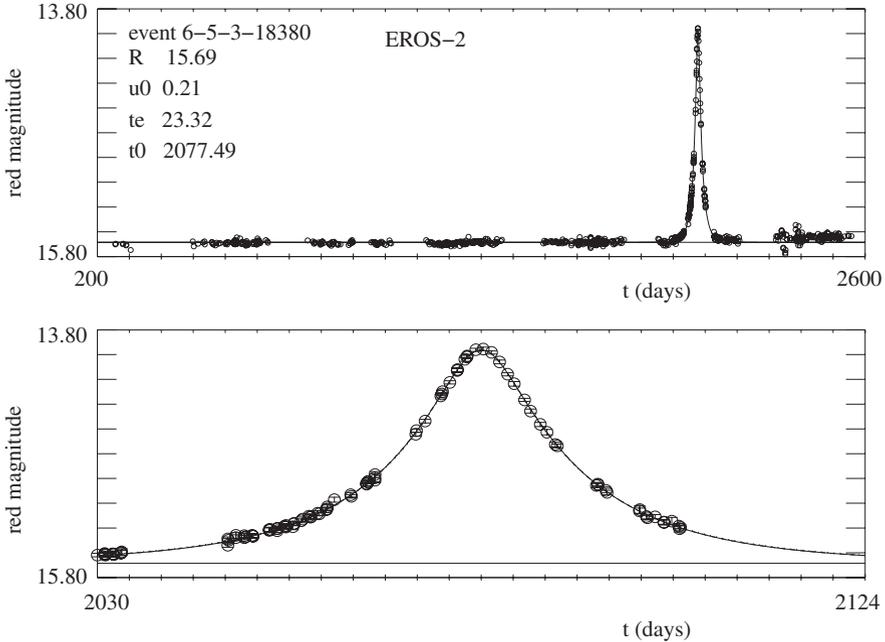
Fig. 2.23 A microlensing event observed by the EROS collaboration [83]. The lensed star is near the center of the Milky Way and the lens is a faint star in the disk of the Milky Way

as being due to dark lenses in the galactic halo, the rate corresponds to a fraction $f = 0.16$ of the total halo mass being comprised of MACHOs. The range of observed t_E correspond to halo objects of mass $\sim 0.4 M_\odot$. However, the results of the EROS group appears to rule out this possibility, suggesting that the events observed by the MACHO collaboration are due to lensing by stars in the LMC itself. (The MACHO group monitored mostly very dense regions of the LMC where such so-called "self-lensing" may have a high rate.)

Microlensing searches for dark objects are also being performed for the nearby spiral galaxy M31 [92]. Events have been observed but there is no consensus on the amount of MACHO dark matter that they represent.

The microlensing limits toward the Magellanic clouds appear to rule out MACHOs as the dominant dark matter for MACHO masses less than $\sim 30 M_\odot$. MACHOs with higher masses are mostly ruled out as dark matter candidates because they would disrupt bound systems as they pass through the disk of the Milky Way. The various limits are shown in Fig. 2.24.

2.5.4 Cold Gas

A second way to hide baryons in galactic halos is to place them in small clouds of cold gas comprised of primordial helium and *molecular* hydrogen [12]. The
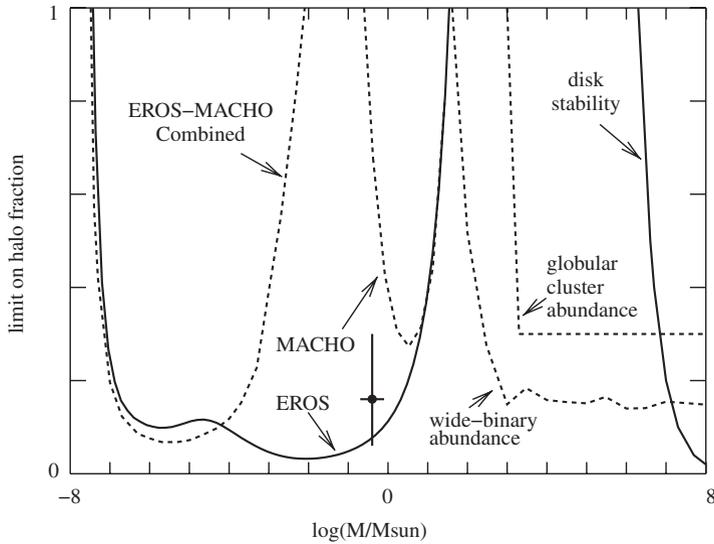
Fig. 2.24 Upper limits on the contribution of MACHOs to the halo of the Milky way as a function of MACHO mass. The curves labeled EROS [86], EROS-MACHO [84], and MACHO [87] are limits deduced from the lack of microlensing events toward the LMC. The curves labeled "wide-binary abundance" (controversial) [88], "globular cluster abundance" [89], and "disk stability" [90] are deduced from the lack of disruption of bound structures by passing MACHOs. The cross at a halo fraction of ~ 0.16 for a MACHO mass of $\sim 0.4 M_{\odot}$ corresponds to the microlensing events seen by the MACHO collaboration [85] as corrected for variable star contamination [91]. The strict EROS limit at this mass suggests that the MACHO collaboration events are not due to lensing by MACHOs in the Milky way halo but, rather, by faint stars in the LMC itself

hydrogen must be molecular in order to escape detection via 21 cm emission by atomic hydrogen. The gas must be in clouds because a spatially uniform gas would lead to unobserved absorption of extragalactic sources at molecular transitions [93]. Clouds of a sufficiently high density would be sufficiently rare that most lines-of-sight would have no such absorption.

While this proposal is very efficient in hiding the gas, the plausibility of producing such quantities of molecules is controversial. In galactic disks, molecules are believed to be produced primarily on the surfaces of dust grains and this would not be possible in a primordial mixture of gas.

Limits of the quantity of cold molecular gas clouds near the Milky Way disk have been obtained from limits on the flux of high-energy photons that would be produced by cosmic-ray interactions in the clouds [94]. Limits on the amount in the halo are more difficult to obtain. Under certain conditions, molecular clouds should be observable in microlensing surveys, either in our Galaxy by using the Magellanic Clouds [95] or in galaxy clusters by using background quasars [96].

2.6 The Cosmological Parameters

2.6.1 H_0

The current universal expansion rate, H_0, plays several roles in cosmology. Its inverse, the Hubble time t_H, is the time scale of the expansion giving the order of magnitude of the time elapsed since the beginning of the present epoch of classical expansion after the end of the inflationary epoch:

$$t_0 = H_0^{-1} f(\Omega_M, \Omega_\Lambda) , \qquad (2.25)$$

where the function $f(\Omega_M, \Omega_\Lambda)$ takes into account acceleration or deceleration and will be calculated in Chap. 5. The Hubble distance, $d_H = c\, H_0^{-1}$, gives the scale of the present classical horizon and the relation between the distances and the redshifts of nearby galaxies

$$R = c\, H_0^{-1} z \qquad z \ll 1 . \qquad (2.26)$$

H_0 determines the critical density

$$\rho_c = \frac{3 H_0^2}{8\pi G} . \qquad (2.27)$$

In a universe with $\Omega_\Lambda = 0$, the critical density determines the dividing line between universes that will continue to expand eternally and universes that will eventually contract.

Finally, as we will see in the next section, H_0 enters into measurements of universal densities in various ways, and will therefore be needed to compare densities. For example, the density of photons is directly measured giving $\rho_\gamma \propto \Omega_\gamma H_0^2$. The theory of primordial nucleosynthesis allows us to determine the baryon-to-photon ratio, $\eta \equiv n_b/n_\gamma$, yielding an estimate of $\rho_b \sim m_p \eta \rho_\gamma \propto \Omega_b H_0^2$. On the other hand, observations of BAO and high-redshift supernovae give directly an estimate of Ω_M and Ω_Λ. Conclusions about the relative quantities of photons, baryons, dark matter, and dark energy therefore depend on H_0. This fact makes the expected anisotropies of the CMB depend on the value of H_0. The precise value given in Table 1.1, $H_0 = 70.5 \pm 1.3\ \mathrm{km\,s^{-1} Mpc^{-1}}$ is based on CMB studies discussed in Chap. 7

While less precise than the estimates based on CMB anisotropies, the value of H_0 measured locally from Hubble's law remains an essential consistency test of cosmology. Hubble's law is

$$v = H_0 R , \qquad (2.28)$$

where v is the recession velocity of a galaxy (easily determined from the redshift of the galaxy's spectral lines) and R is the distance of the galaxy. The law applies in the velocity range $10^{-2}c < v < 10^{-1}c$, the lower limit necessary to neglect random peculiar velocities and the upper limit to neglect relativistic corrections that depend on q_0. Galaxies in this redshift range are called "Hubble flow" galaxies.

The Hubble Key project [9] on the Hubble Space Telescope (HST) provided a breakthrough in the measurement of H_0 giving a value with a precision of $\sim 10\%$:

$$H_0 = (72 \pm 8)\,\mathrm{km\ sec^{-1}\,Mpc^{-1}}\ . \tag{2.29}$$

Reported values of H_0 over the previous decade had spanned a range over nearly a factor 2 from $H_0 \sim 50\,\mathrm{km\,s^{-1}Mpc^{-1}}$ to $H_0 \sim 90\,\mathrm{km\,s^{-1}Mpc^{-1}}$. The Hubble Key project "Hubble diagram" on log-log scale is shown in Fig. 2.25

The primary difficulty in determining H_0 comes from the difficulty in determining the distances to galaxies. Very loosely speaking the various methods can be separated into "astrophysical methods" that are calibrated with the "distance ladder" and "physical methods" giving directly galactic distances.

Most of the points in Fig. 2.25 use the Hubble Key Project distance ladder summarized in Table 2.3. There are four steps on this ladder: the Milky Way, the large Magellanic Cloud, other nearby galaxies ($100\,\mathrm{kpc} < R < 30\,\mathrm{Mpc}$), and Hubble flow galaxies ($60\,\mathrm{Mpc} < R < 400\,\mathrm{Mpc}$). Only objects on the last step can be used to measure H_0. Generally speaking, the distance, R, to an object on a given step is deduced from its observed light flux, $f = L/4\pi R^2$, where L is the *known*
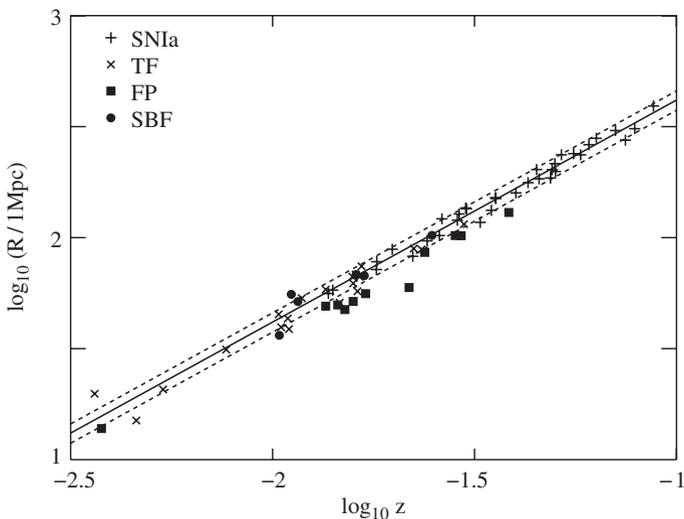

Fig. 2.25 The Hubble diagram of the Hubble Key Project [9]. The plot shows distances to galaxies vs. their redshift with distances determined by the techniques as marked

Table 2.3 The distance ladder used by the Hubble Key Project [9]. The four steps are (1) nearby Milky Way stars measured by parallax; (2) the Large Magellanic Cloud measured mostly by standard-candle stars calibrated locally by parallax (Fig. 2.3); (3) nearby galaxies measured by Cepheids variable stars calibrated in the Large Magellanic Cloud (Fig. 2.5); and (4) "secondary" distance indicators calibrated in nearby galaxies

Objects	Method
Milky Way stars ($R < 200$ pc)	Parallax (Hipparcos)
Large Magellanic Cloud (∼ 1000 Cepheids)	Cepheid variables......$\mu = 18.57 \pm 0.14$ Eclipsing variables...........18.33 ± 0.05 SN1987a rings.................18.47 ± 0.08 Red Giants.....................18.64 ± 0.05 Red Clump.....................18.27 ± 0.11 RR Lyrae variables.........18.30 ± 0.13 Mira variables................18.54 ± 0.04 average...................$\mu = 18.5 \pm 0.1$ $R_{LMC} = 10$pc $10^{\mu/5} = (50.1 \pm 2.5)$ kpc
31 galaxies (0.1Mpc $< R < 30$ Mpc) (6 SNIa hosts)	Cepheid variables

Galaxies (60Mpc $< R < 400$ Mpc)
  36 galaxies
  21 galaxy clusters
  11 galaxy clusters
  6 galaxy clusters
  4 galaxies
	
	SNIa............................$H_0 = 71 \pm 2 \pm 6$
	Tully-Fisher................$H_0 = 71 \pm 3 \pm 7$
	Fundamental Plane.....$H_0 = 82 \pm 6 \pm 9$
	Surface Fluctuations...$H_0 = 70 \pm 5 \pm 6$
	SNII............................$H_0 = 72 \pm 9 \pm 7$
	Average......................$H_0 = 72 \pm 8$

luminosity of the object. The luminosity is derived from similar objects at a lower step of *known distance*. The Hubble Key Project ladder proceeds as follows:

1: Distances to Milky Way objects are derived from parallax measurements of the Hipparcos satellite (Fig. 2.3) allowing one to deduce the luminosities of various classes of stars.
2: Observation of these classes of stars in the Large Magellanic Cloud (LMC) allows one to deduce the distance to the LMC. This LMC contains many bright Cepheid periodic variable stars that have a well-defined luminosity as a function of their period (Fig. 2.5) and the known distance to the LMC allows one to measure this period-luminosity relation.
3: Because of its excellent angular resolution, allowing it to resolve individual stars in nearby galaxies, the HST can detect Cepheid variable stars in galaxies within ∼ 50 Mpc, thus determining their distance. The galaxies on the third step are then used to calibrate various "secondary" distance indicators, the most important being the luminosity of type Ia supernovae (Fig. 2.9) and

the luminosity of spiral galaxies as a function of their velocity dispersion (Tully-Fisher relation, Fig 2.11).

4: The known luminosities of supernovae or galaxies are then used to determine the distances to galaxies more distant than 50 Mpc, thus determining H_0.

The uncertainty in the quoted value (2.29) includes the cumulative uncertainties in each step of the ladder, e.g., the 5% uncertainty in the distance to the Large Magellanic Cloud.

Three "physical" methods permit one, in principle, to bypass the distance ladder. The first is the "expanding photosphere method" (EPM) applied to type II supernovae (Sect. 2.1). After the initial explosion, the surface (photosphere) of the remnant expands at a velocity $v \sim 10^{-2}c$, as illustrated in Fig. 2.26. The velocity can be determined empirically from the Doppler shifts of the lines in the supernova spectrum. Knowledge of the moment of the explosion t_{exp} then allows one to calculate the physical diameter of the photosphere $D = 2v(t - t_{\text{exp}})$. If the angular size $\Delta\theta$ were known, the distance to the supernova could be determined directly via $R = 2v(t - t_{\text{exp}})/\Delta\theta$. For extragalactic supernovae, the angular size is too small to be directly measured but it can be estimated from the measured luminous flux by using Stefan's law (2.1), appropriately modified for the non-blackness of the photosphere. This technique has given a value of $H_0 = 73 \pm 15$ km s^{-1}Mpc^{-1} [97]. The Hubble Key project calibrated this method in their distance ladder (Table 2.3) and deduced a value of $H_0 = 72 \pm 9 \pm 7$.

The second physical method uses the time delay between two images of quasars that are gravitationally lensed by foreground galaxy clusters. Several examples of gravitationally lensed quasars have been found where intrinsic variability of the quasar permits the measurement of the time delay. The light-curve (flux vs. time) for one is shown in Fig. 3.23. Since the optical paths are proportional to the distance scale and therefore to c/H_0 it is not surprising that the time delay between the two images is proportional to c/H_0. The constant of proportionality will be calculated in Sect. 3.8. It depends on the angular separation of the two images and on the mass distribution of the lens. The mass distribution can be estimated from the distribution of gravitational arcs due to lensed background galaxies. Knowledge of the distribu-
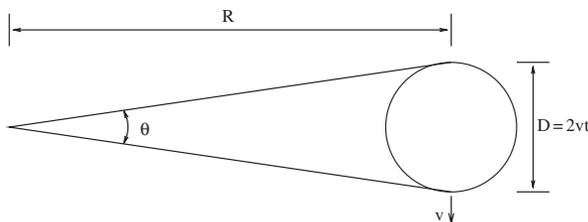

Fig. 2.26 The photosphere of a supernova expanding at a velocity v. The velocity can be deduced from the position and shape of the spectral lines. Knowledge of the time of explosion t_{exp} allows one to deduce the diameter $D = 2v(t - t_{exp})$. The angle θ can be estimated using Stefan's law, allowing one to estimate the distance R
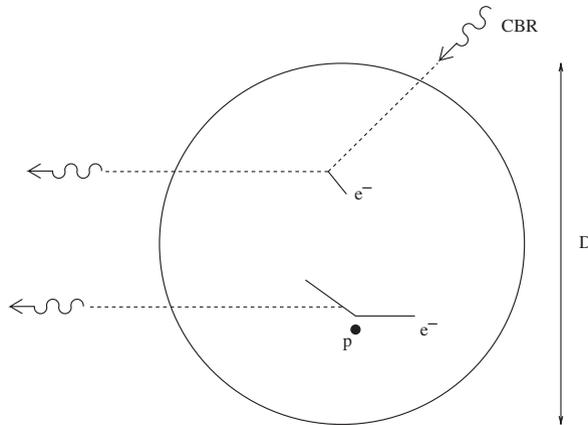
Fig. 2.27 A cluster of galaxies of diameter D. CMB photons can scatter on the free electrons in the hot ionized cluster gas increasing the mean energy of the CMB photons coming from the direction of the cluster (SZ effect). The ionized gas also emits X-rays via thermal bremsstrahlung of electrons scattering on protons and nuclei

tion limits the precision of this method. The Hubble Key Project [9] suggests that the measurements appear to be "converging" to about $H_0 \sim 65 \, \text{km s}^{-1} \text{Mpc}^{-1}$.

The last physical method uses the Sunyaev–Zel'dovich (SZ) effect by which the CMB is heated as it Compton scatters on the hot ionized gas of a galaxy cluster (Fig. 2.27). The CMB spectrum is thus deformed in the direction of the cluster in proportion to the probability to scatter in the cluster:

$$P \sim \langle n_e \rangle \sigma_T D = \langle n_e \rangle \sigma_T R \Delta \theta \;, \tag{2.30}$$

where $\langle n_e \rangle$ is the mean electron density in the cluster, σ_T is the Thomson cross-section, and $D = R \Delta \theta$ is the diameter of the cluster at a distance R and subtending an angle $\Delta \theta$. For a cluster, $\Delta \theta$ can be directly observed so if P is measured from the SZ effect, we need only an estimation of $\langle n_e \rangle$ to determine R.

The electron density can be estimated from the cluster X-ray luminosity which is (Exercise 2.10)

$$L_X \sim c^2 \langle n_e^2 \rangle \alpha \sigma_T \sqrt{m_e T} D^3 \;, \tag{2.31}$$

where α is the fine structure constant. It follows that

$$\langle n_e^2 \rangle \sim \frac{4\pi f_x}{c\alpha\sigma_T \Delta\theta^3 R^3} \;. \tag{2.32}$$

Combining (2.30) and (2.32), we find the distance to the cluster:

$$R \sim \frac{P^2 \, c\alpha \, \Delta\theta \, \sqrt{m_e T}}{4\pi \, f_x \sigma_T} \, \frac{\langle n_e^2 \rangle}{\langle n_e \rangle^2} \, . \tag{2.33}$$

Everything on the right can be measured except for $\langle n_e^2 \rangle / \langle n_e \rangle^2$. This last factor is of order unity if the gas is relatively uniform. Any non-uniformity would result in an overestimation of R and, therefore, an underestimation of H_0. It is also necessary to suppose that the cluster is spherical, though this hypothesis can be eliminated by averaging measurements of several clusters. The compilation of the Hubble Key Project [9] lists measured values in the range $40 \, \text{km s}^{-1} \text{Mpc}^{-1} < H_0 < 80 \, \text{km s}^{-1} \text{Mpc}^{-1}$. An increase in the number of studied clusters will enable astronomers to better estimate and control the uncertainties associated with this promising method.

2.6.2 ρs and Ωs

The densities listed in Table 1.1 were derived from WMAP measurements of CMB anisotropies combined with measurements of type Ia supernovae and of baryon acoustic oscillations (BAO). In this section, we give a summary of these methods and a few others that have somewhat lower precision but serve as important consistency checks. The major methods are listed in Table 2.4.

The only directly measured cosmological density is the photon density with the COBE measurement [5] of the CMB temperature giving:

$$\rho_\gamma = (2.61 \pm 0.01) \times 10^5 \, \text{eV m}^{-3} \Rightarrow \Omega_\gamma = 5.16 \, h_{70}^{-2} \times 10^{-5} \, . \tag{2.34}$$

Theoretical arguments based on the thermodynamics of the early universe allow one to deduce the densities of neutrinos and baryons:

- Neutrinos. The thermodynamic calculations of Chap. 6 will give $n_\nu = (3/11)n_\gamma$ for each species. The present neutrino density then depends on the neutrino masses with (1.36) for effectively massless neutrinos ($m_\nu \ll T_\gamma(t_0) = 2.349 \times 10^{-4} \, \text{eV}$) and (1.37) for massive neutrinos. We note that the same type of calculation will give the number density of any species of weakly interacting massive particle if there is no particle–antiparticle asymmetry.
- Baryons. The theory of primordial nucleosynthesis predicts the abundances of the light elements as a function of the baryon–photon ratio, $\eta = n_b/n_\gamma$. The observed abundances [10] imply $\eta = (5 \pm 1) \times 10^{-10}$ which gives $\Omega_b = (0.04 \pm 0.01) \, h_{70}^{-2}$.

Knowing Ω_b one can deduce Ω_M by measuring the total mass to baryonic mass in objects that are believed to have a representative mix of baryons and CDM. This is expected to be nearly true for large clusters of galaxies. The baryons in these galaxies is mostly in the form of hot ionized gas and the mass in this form can be

Table 2.4 Some methods of determining the cosmological density parameters. Only ρ_γ is directly measured. Knowledge of H_0 allows one to deduce Ω_γ followed by Ω_V, Ω_b, and Ω_M. So-called "geometrical" methods using standard candles or rulers allow one to directly deduce Ω_M (BAO), $\Omega_M - \Omega_\Lambda$ (SNIa) and Ω_T (peak positions in CMB power spectrum). The details of the CMB power spectrum allows one to deduce Ω_Λ, Ω_M, and Ω_b if one knows H_0. Finally, measurements of the abundance of galaxy clusters and weak gravitational lensing can be used to determine the product of Ω_M and σ_8, the relative mass fluctuations of the scale of $8h^{-1}$Mpc

Quantity	Method	Reference
$\rho_\gamma \propto \Omega_\gamma H_0^2$	CMB density	
Ω_γ	CMB density $+H_0$	
Ω_V/Ω_γ	primordial thermal equilibrium $+ m_V$	Sect. 6.4
Ω_b/Ω_γ	primordial nucleosynthesis	Sect. 6.5
Ω_b/Ω_M	galaxy cluster baryon fraction	Exercise 2.6
$\Omega_M H_0$	Large Scale Structure power spectrum	Sect. 7.2
Ω_M	Baryon Acoustic Oscillations (BAO)	Sect. 5.3
	cluster number evolution	Sect. 2.3
$\Omega_M - \Omega_\Lambda$	SNIa Hubble diagram	Sect. 5.2
Ω_T	CMB anisotropies + BAO	Sect. 7.4.2
Ω_Λ, Ω_M, Ω_b	CMB anisotropies $+ H_0$	Sect. 7.4.2
$\Omega_M \sigma_8$	galaxy cluster abundance	
	large-scale weak lensing	Sect. 7.2

deduced from the X-ray luminosity. The fraction of the total mass of the largest galaxy clusters that is comprised of hot gas is measured to be $f_{gas} \sim 0.12 \, h_{70}^{-3/2}$ [98, 99]. This gives $\Omega_M \sim \Omega_b/0.12 \sim 0.37$. Corrections taking into account the amount of baryons in stars bring Ω_M down to $\Omega_M \sim 0.28 \pm 0.06$ [99].

The cluster number density measurements (Fig. 2.15) can give Ω_M if the density can be measured as a function of time or equivalently of redshift. The evolution of the number of clusters depends on Ω_M because structure formation slows at $\hat{a} = \Omega_M/(1 - \Omega_M)$ for $\Omega_\Lambda = 0$ and at $\hat{a} = (\Omega_M/(1 - \Omega_M))^{1/3}$ for $\Omega_T = 1$. The small amount of evolution shown in Fig. 2.15 indicates [100] $\Omega_M = 0.34 \pm 0.08$.

At super-cluster scales, Ω_M is one of the parameters that determine the shape of the "power spectrum" of density fluctuations (Fig. 7.4). The large-scale structure of the universe can be characterized by Fourier decomposing the density contrast into modes of comoving wavelength $\lambda = 2\pi\hat{a}(t)/k$ where k is the fixed wavenumber. In structure formation theories with cold dark matter and primordial adiabatic density

fluctuations, short-wavelength modes have amplitudes that are suppressed because these modes oscillated as acoustic waves during the radiation epoch whereas the amplitude of long-wavelength modes grew during both radiation and matter epochs. The separation between short and long wavelengths corresponds to the Hubble distance at the moment of matter-radiation equality, $a(t_{eq}) = a_0 1.68 \Omega_\gamma / \Omega_M$ and to a present wavelength of $\lambda_{eq} \sim 600\, h_{70}^{-1} \mathrm{Mpc} \times 0.27/(\Omega_M h_{70})$.

The most important methods for measuring Ω_M and Ω_Λ are "geometrical" methods that use type Ia supernovae as standard candles and the sound horizon at recombination as a standard ruler seen in the matter and CMB fluctuation spectra. These techniques are described in Chap. 5.

Finally, we note that Ω_M can be deduced from the abundance of galaxy clusters and from the gravitational lensing galaxies by large-scale structure if one uses information of the mass fluctuation spectrum (Chap. 7).

Exercises

2.1 The luminosity of a typical galaxy is $\sim 2 h_{70}^{-2} \times 10^{10}\, L_\odot$ and the mean energy of stellar photons is $\sim 2\,\mathrm{eV}$. What is the photon flux (in $\mathrm{m^{-2}\,s^{-1}}$) of a galaxy of redshift z ($z \ll 1 \Rightarrow R \sim z d_H$)?

Compare the photon flux from the nearest large galaxies ($R \sim 1$ Mpc) with the photon flux from the nearest stars ($R \sim 1$ pc). (This calculation should explain why most objects visible to the naked eye are stars.)

2.2 The luminosity density of the universe is $\sim 1.2\, h_{70} 10^8 L_\odot\, \mathrm{Mpc^{-3}}$. Supposing that stellar light output has been relatively constant since the formation of the first stars about one Hubble time ago, estimate the number of photons ($E \sim 2$ eV) that have been produced by stars. Compare the number of stellar photons with the number of CMB photons. (This problem will be treated more rigorously in Exercise 5.12.)

Stellar energy is mostly produced by the fusion of hydrogen to helium $4p \rightarrow^4$ He $+ 2e^+ + 2\nu_e$. This transformation occurs through a series of reactions in stellar cores that liberate a total of ~ 25 MeV. After thermalization, the energy emerges from stellar surfaces in the form of starlight. Estimate the number of protons (per $\mathrm{Mpc^3}$) that have been transformed into helium over the last Hubble time. Compare this number with the number of protons available $n_b \sim \Omega_b \rho_c / m_p$.

2.3 Estimate the contribution to the universal photon mean free path of the following processes:

- Thomson scattering of photons on free electrons of number density $n_e \sim n_b$.
- Absorption by stars of number density $n_{stars} \sim \Omega_{stars} \rho_c / M_\odot$, $\Omega_{stars} \sim 0.0023$ and cross-section $\sim \pi R_\odot^2$.

- Absorption by dust in galaxies with $n_{gal} \sim 0.005$ Mpc^{-3} and cross-section $\sim \epsilon \pi R_{gal}^2$ where $R_{gal} \sim 10$ kpc and the fraction of visible light absorbed when passing through a galaxy is $\epsilon \sim 0.1$.

Compare these distances with d_H ($\sim$ the distance of the most distant visible objects). Is the universe "transparent"? (Section 5.7 will treat this problem more rigorously.)

2.4 Supposing that we can only measure redshifts, angles, and photon fluxes, explain the factors of h_{70} in (2.9), (2.10), and (2.11) as well as the absence of such factors in (2.12).

2.5 By comparing the apparent magnitudes of LMC Cepheids (Fig. 2.5) to the apparent magnitudes of Cepheids in the galaxy NGC 1365 (Fig. 2.28), estimate the ratio of the NGC 1365 distance to the LMC distance. If the LMC distance is taken to be 50 ± 5 kpc, what is the distance to NGC 1365.

NGC 1365 is a member of the Fornax galaxy cluster. The recession velocity of this cluster is 1441 km s^{-1}. Estimate H_0.
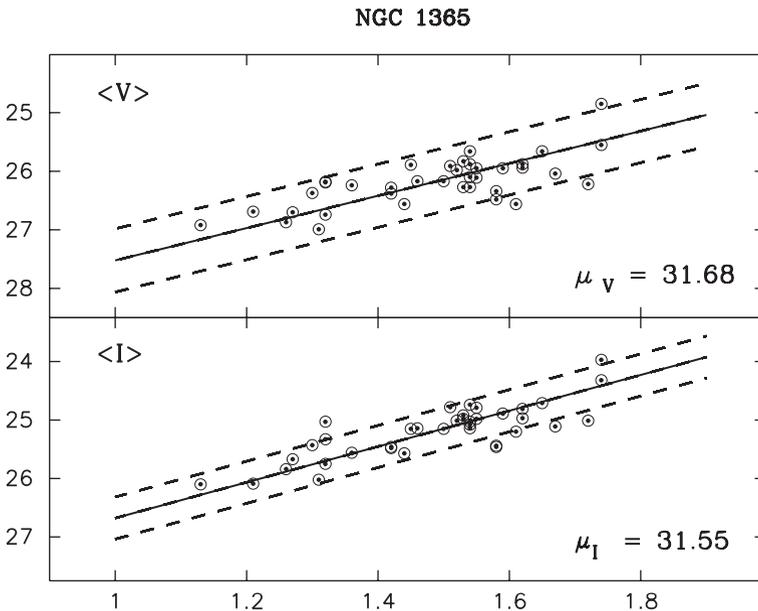

Fig. 2.28 The apparent magnitudes in the I and V bands of Cepheids in the galaxy NGC 1365 [53]. The magnitude is a linear function of $\log P$ where P is the Cepheid period in days

2.6 Abell-496 is a galaxy cluster whose properties were studied in [101].

(a) The recession velocity of A496 is 9885 km s^{-1}. What is its distance as a function of h_{70}?

(b) Figure 2.29 shows the angular distribution of the brightest galaxies in A496. Estimate the radius of the cluster.

(c) Figure 2.30 shows the distribution of recession velocities in the direction of A496. The accumulation near 9885 km s^{-1} corresponds to the cluster members. The width of this accumulation suggests that the line-of-sight velocity dispersion of A496 is $\Delta v \sim 715$ km s^{-1}. Use the virial theorem to estimate the cluster mass: $(GM/\Delta r \sim \Delta v^2)$. A detailed study in [101] gives $M_{vir} = (5.1 \pm 0.8) h_{70}^{-1} 10^{14} M_\odot$.

(d) The flux of visible light from A496 indicates a total luminosity of $L = 2.0\, h_{70}^{-2} \times 10^{12}\, L_\odot$. By assuming that M_{vir}/L is equal to the universal value ρ_M/J_0 with the universal luminosity density given by (2.10), estimate ρ_M and Ω_M.

(e) The X-ray spectrum indicates that the temperature of the intergalactic gas in A496 is 4 ± 1 keV. The X-ray flux allows one (Exercise 2.10) to estimate the total mass and the total mass of intergalactic gas. For A496, the total mass within $0.7\, h_{70}^{-1}$ Mpc of the cluster center is $(1.7 \pm 0.4) h_{70}^{-1} \times 10^{14} M_\odot$ and the total gas mass within the same radius is $(1.4 \pm 0.5) h_{70}^{-5/2} \times 10^{13} M_\odot$. Assuming that the ratio between baryonic mass and total mass of A496 is equal to the universal value, estimate Ω_b/Ω_M.
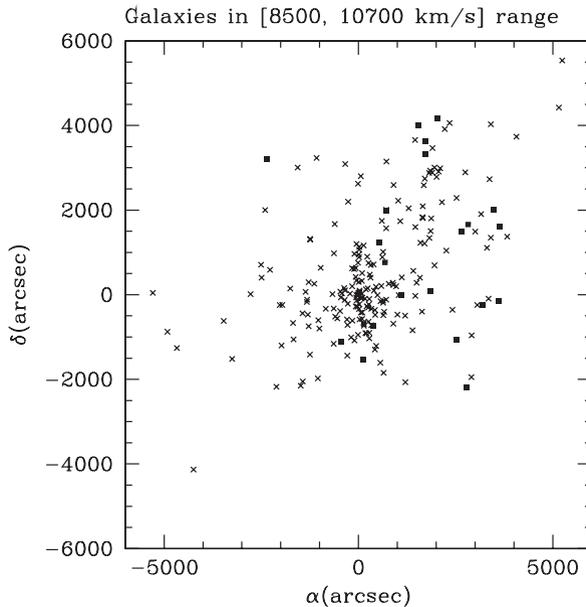

Fig. 2.29 The angular distribution of bright galaxies in A496 [101]
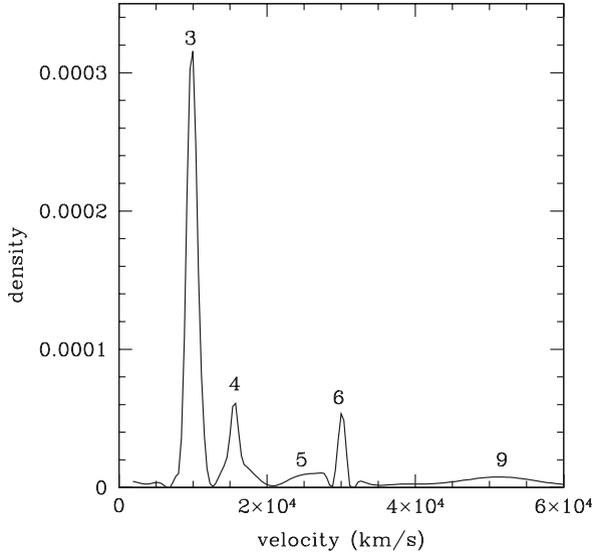
Fig. 2.30 The distribution of recession velocities in the direction of the A496 [101]. The accumulation of 274 galaxies around 9885 km s^{-1} corresponds to the cluster members

2.7 Figure 2.31 shows the iso-recession velocity curves of the galaxy NGC 5033 deduced from the Doppler shift of the 21 cm line of atomic hydrogen. The curves are superimposed on an optical image of the galaxy.

(a) What is the redshift of NGC 5033. By neglecting its peculiar velocity, estimate its distance as a function of h_{70}.
(b) The visible angular radius of NGC 5033 is about 3 arcmin. What is the visible radius as a function of h_{70}?
(c) What is the rotation velocity far from the galactic center? Take into account the galaxy inclination by supposing that the galaxy would appear to be circular if viewed face-on.
(d) Estimate the mass of NGC 5033 that is within 6 arcmin of the galactic center (in units of $M_\odot$ and as a function of h_{70}).
(e) NGC 5033 has an apparent magnitude in the V band of 10.1. What is its absolute magnitude and its luminosity (in units of $L_{\odot V}$) as a function of h_{70}? What is its mass-to-light ratio?

2.8 The most reliable distance indicators out to $\sim$ 50 Mpc are Cepheid variable stars. The most reliable method of calibrating the luminosity-period relation of Cepheids is to use the large number of Cepheids observed in the Large Magellanic Cloud (LMC). This method of calibration obviously requires a knowledge of the LMC distance.

One of the most direct measurements of the LMC distance uses "eclipsing binaries." Such systems consist of two orbiting stars whose orbital plane is oriented such
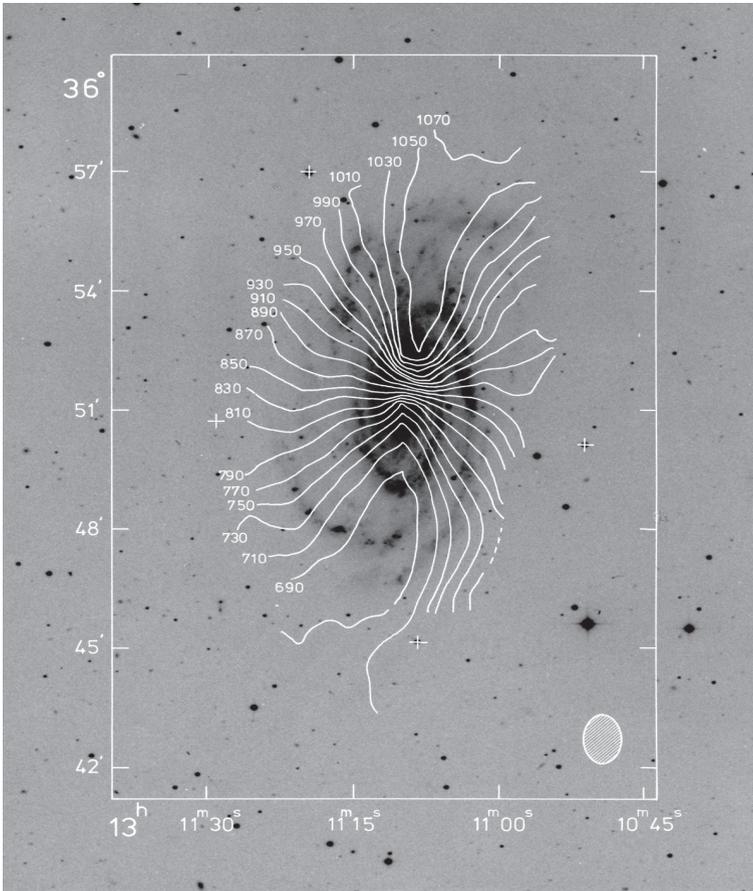
Fig. 2.31 The iso-recession-velocity curves of the galaxy NGC 5033 deduced from the Doppler shift of the 21 cm line of atomic hydrogen [102]. The curves are superimposed on an optical image of the galaxy. The angular scale of the greater dimension is in arcmin. Courtesy of A. Bosma

that, viewed from Earth, the two stars periodically eclipse each other. For eclipsing binaries at the distance of the LMC, the two stars generally have an angular separation that is so small that the two stars cannot be optically resolved. Rather, they appear as a single star with a periodic luminosity due to the periodic eclipses.

Figure 2.32 shows the "light curve" (apparent magnitude versus time) of the binary system HV2274 in the LMC [103]. Two eclipses are present with a period of 5.726 days. The magnitude change of 0.75 during the eclipses corresponds to a factor of two in flux indicating a total eclipse of two stars of equal luminosities and radii.

The spectral lines of the two stars do not coincide because of the Doppler shift due to their orbital motion. It is therefore possible to determine independently the
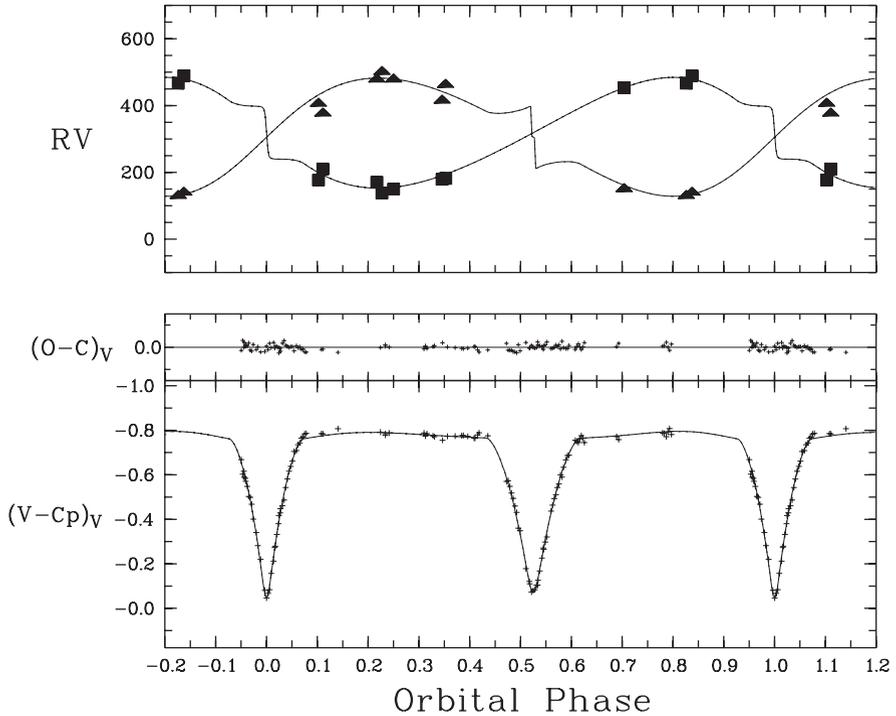
Fig. 2.32 The binary system HV2274 in the LMC [103]. The *upper panel* shows the recession velocity of the two stellar components as a function of the orbital phase (period = 5.726 days). The *lower panel* shows the light curve (apparent magnitude versus time)

recession velocities of the two stars. The two velocities as a function of time are also shown in Fig. 2.32.

(a) Estimate the orbital velocity of the two stars and the radius of the (circular) orbit.

(b) Supposing that the two stars have equal masses, estimate their mass (in units of $M_\odot$).

(c) Use the durations of the eclipses to estimate the common radius of the two stars (in units of $R_\odot$).

(d) The surface temperature of the two stars is $\sim 23\,000$ K. The measured flux indicates, via (2.1), an angular size of $D/R = 9.48 \times 10^{-12}$. Estimate the distance R to HV2274. After a small correction for the relative distance between HV2274 and the center of the LMC, the authors of [103] give an LMC distance of 45.77 ± 1.6 kpc.

(e) Figure 2.5 shows the apparent magnitudes of LMC Cepheids as a function of their periods. Using the distance to the LMC, transform the apparent magnitudes into absolute magnitudes. Compare these magnitudes with those of Cepheids with distances determined by parallax [44], $M_V = -2.81 \log P - 1.43 \pm 0.16$

(period P in days). Compare the value of H_0 that would be estimated using Hipparcos Cepheids with that using LMC Cepheids.

2.9 It is perhaps surprising that the luminosity of a star can be estimated theoretically without knowing the nuclear reactions that power the star. To see how this can be done, we consider a sphere of radius R containing N_p protons and N_p electrons in the form of an ideal ionized gas.

(a) if the sphere has a uniform density and is in hydrostatic equilibrium with a mean pressure P and volume V, show that

$$3PV = -E_g \sim (3/5)\frac{Gm_p^2 N_p^2}{R} , \tag{2.35}$$

where E_g is the total gravitational energy of the sphere and m_p is the proton mass.

The numerical factor $(3/5)$ in (2.35) applies only to a sphere of uniform density. This is not the case for a star but a nonuniform distribution would simply give a different numerical factor. For the rest of this exercise we will ignore all numerical factors.

Applying the ideal gas law to (2.35), we can estimate the mean temperature T in the star:

$$T \sim \frac{Gm_p^2 N_p}{R} . \tag{2.36}$$

(b) Supposing that the sphere contains photons in thermal equilibrium at the temperature T, show that the total number of photons inside the star is

$$N_\gamma \sim N_p^3 \left(\frac{m_p}{m_{pl}}\right)^6 , \tag{2.37}$$

where $m_{pl} = \sqrt{\hbar c/G} = 1.2 \times 10^{19}$GeV is the Planck mass. Compare N_γ with N_p for the sun ($N_p \sim 10^{57}$).

The photons diffuse in the star before escaping at the surface. The number of collisions in this random walk is of order

$$N_{col} \sim \left(\frac{R}{\lambda}\right)^2 , \tag{2.38}$$

where λ is the mean free path of a photon in the star.

(c) Show that the mean escape time for a photon is

$$\tau \sim \frac{N_p \sigma}{Rc} , \tag{2.39}$$

where σ is the mean photon-particle cross-section in the star. From this, argue that the stellar luminosity is

$$L \sim N_{\mathrm{p}}^3 \left(\frac{m_{\mathrm{p}}}{m_{\mathrm{pl}}} \right)^8 \frac{\hbar c^2}{\sigma} . \tag{2.40}$$

For a star like the Sun, the atoms are nearly all ionized except near the surface. It follows that $\sigma \sim \sigma_{\mathrm{T}}$ (the Thomson cross-section). For $N = 10^{57}$, compare the luminosity from (2.40) with $L_\odot$.

A more careful management of the numerical factors multiplies the above result by $\pi^4/(5 \times 3^8) \sim 3 \times 10^{-3}$ [104]. This gives an agreement with the observed solar luminosity that is reasonable considering the approximations involved in the calculation.

We note that (2.40) predicts that a stellar luminosity is proportional to the third power of its mass, in good agreement with observations.

If the luminosity of a star depends only on its mass, one can ask what is the role of the nuclear reactions that power the star. The answer is that they allow the star to burn *longer* at a stable radius. A star begins its life as a diffuse cloud that is too cold to initiate nuclear reactions. It nevertheless radiates photons as required by (2.40). In so doing, it radiates energy and total energy conservation requires the star's radius to diminish. As the star becomes smaller, its temperature rises until nuclear reactions are ignited. At this point, a stable regime is reached where the energy radiated is replaced by the energy liberated by the nuclear reactions.

2.10 Roughly 10% of the mass of large galactic clusters is contained in ionized intergalactic gas in hydrostatic equilibrium. This gas produces photons via bremsstrahlung:

$$\mathrm{e}^- \mathrm{p} \rightarrow \mathrm{e}^- \mathrm{p} \gamma . \tag{2.41}$$

Theoretical and observed spectra are shown in Figs. 2.33 and 2.34.

Unlike the photons produced in stars, these photons escape directly from the cluster:

(a) The largest clusters contain $\sim 10^{14} M_\odot$ of ionized hydrogen in a radius of ~ 1 Mpc. Verify that the mean free path of photons due to Thomson scattering in the cluster is greater than the cluster radius.

(b) The large angle Rutherford scattering cross-section is $\sim \sigma_{\mathrm{T}}/(v/c)^4$. Verify that the effective mean free path of electrons (scattering on protons) in the cluster is less than the cluster radius. This justifies the assumption that the electrons and protons form a thermal gas in hydrostatic equilibrium.

For non-relativistic electrons, the differential cross-section for bremsstrahlung production of photons of energy E_γ is approximately [106]
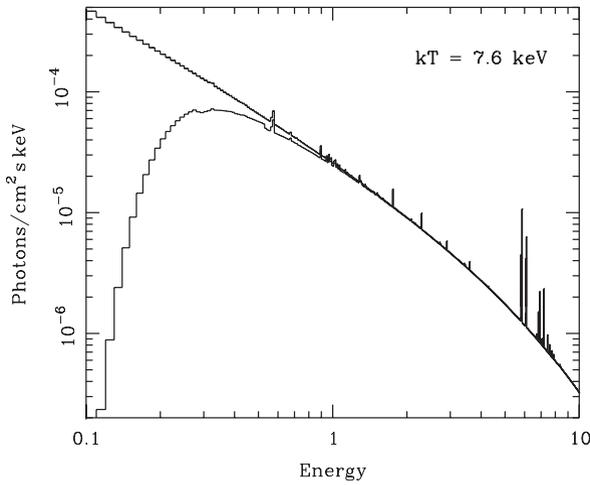
Fig. 2.33 The theoretical X-ray spectrum from a galactic cluster of temperature 7 keV. The spectrum follow the $1/E$ bremsstrahlung cross-section at low energy and then is exponentially cutoff at energies above the temperature. Recombination lines for iron are seen around 6 keV. The second *curve* that has a suppressed flux at low energy shows the effect of absorption in the Milky Way. Courtesy of Monique Arnaud
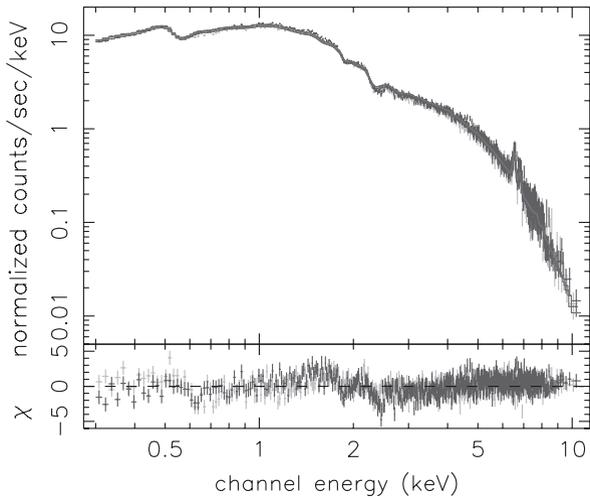

Fig. 2.34 The observed X-ray spectrum from the Coma galaxy cluster as observed by the XMM satellite [105]. The structures in the spectrum around 2 keV and 0.5 keV are due to the varying efficiency of the detection system. The fitted cluster temperature is $kT = 8.25$ keV. Courtesy of Monique Arnaud

$$\frac{d\sigma}{dE_\gamma} \sim \alpha \, \frac{c^2}{v^2} \frac{\sigma_T}{E_\gamma} E_\gamma \ll (1/2) m_e v^2 \;, \tag{2.42}$$

where $\alpha \sim 1/137$ is the fine-structure constant, $v \ll c$ is the electron–proton relative velocity, and σ_T is the Thomson cross-section.

Using a line of reasoning that will be justified in Chap. 6, we know that the production rate per unit volume of photons is proportional to the differential cross-section (2.42), to the electron density n_e, to the proton density $n_p \sim n_e$, and to the mean electron–proton velocity:

$$\frac{dN_\gamma}{dt\,dV} \sim n_p^2 \, v \frac{d\sigma}{dE_\gamma} \sim \frac{c}{v} n_p^2 \alpha c \frac{\sigma_T}{E_\gamma} \;. \tag{2.43}$$

(c) Integrate this expression up to a photon energy cutoff given by the temperature of the cluster T to find the total X-ray luminosity (energy/time):

$$L_x \sim n_p^2 \alpha c \sigma_T \sqrt{m_e c^2 T} \, D^3 \;, \tag{2.44}$$

where D is the diameter of the cluster.

(d) Show that the total number of baryons, N_b, in the cluster can be estimated from the observed X-ray flux, f_x:

$$N_b^2 \sim \frac{f_x R^5 \theta^3}{\alpha c \sigma_T \sqrt{m_e c^2 T}} \;, \tag{2.45}$$

where θ is the observed angular diameter of the cluster and R is the distance to the cluster. This formula shows that if a cluster redshift is used to estimate the cluster's distance, the measured total baryonic mass in the cluster scales as $h_{70}^{-5/2}$.

(e) Modify (2.35) so that the thermal pressure supports only the baryonic mass of the cluster and thereby show that the total cluster mass in terms of the X-ray temperature is

$$M_{\text{tot}} \sim \frac{6kTR}{Gm_p} \tag{2.46}$$