2 Section II: large-scale structure of the Universe

2.1 Evidence for galaxy clustering: redshift surveys

As we survey the local Universe, we see that the spatial distribution of galaxies is not uniform: galaxies appear to be **clustered**. We see evidence of galaxy clustering in the projected (i.e., 2-dimensional) distribution of galaxies on the sky. Surveys such as the APM Galaxy Survey, showing galaxies visible towards our south galactic pole, reveal a tangled filamentary distribution of galaxies through the local Universe. Of course to get a better idea of the 3-dimensional distribution of these galaxies we also need to know their distances.

In 1936 Edwin Hubble plotted the observed (radial) velocities of nearby galaxies, deduced from the Doppler shift of their spectral lines, against their distances, derived from Cepheid variables within the galaxies. The Doppler shift of a line is just

$$z = \frac{\lambda_{\rm observed} - \lambda_{\rm emitted}}{\lambda_{\rm emitted}}, \tag{7}$$

where λ stands for wavelength. By assuming the observed Doppler shift was related to the velocity of the galaxy by $v = cz$, where c is the speed of light[c], he found that the galaxies were nearly all moving away from us (that is they were 'redshifted') and that their recession velocities, $v_{\rm rec}$, were approximately proportional to their distances, d, so that

(8)

This is **Hubble's Law**, and H_0 is known as the **Hubble constant**, usually measured in units[d] of km s^{-1} Mpc^{-1} (we will discuss the cosmological significance of Hubble's law, and methods of measuring H_0 in Section III). So provided Hubble's law holds for more distant galaxies, the measured recession velocity of a galaxy gives us an accurate estimate of its distance, assuming we know the value of the Hubble constant.

For example, suppose we observe a galaxy with a recession velocity $v_{\rm rec} = 13\,000\,{\rm km\,s^{-1}}$. If we assume $H_0 = 71\,{\rm km\,s^{-1}\,Mpc^{-1}}$, then the galaxy is at distance $d = 13\,000/71 = 183\,{\rm Mpc}$.

You could be worried that because we don't know H_0 exactly, using recession velocity (or equivalently redshift) as an indicator of a galaxy's distance is unreliable. In fact we believe we know H_0

---

[c] This is the standard Doppler shift formula for velocities much less than c. In cosmology, the exact formula depends on the cosmological model used, but all of them approximate to $v = cz$ at low redshift.

[d] Remember a megaparsec (Mpc) is a million parsecs, which is 3.3 million light years or about 3×10^{22} m.

rather well now, but *whatever* its value, Hubble's law tells us that $v_{\rm rec}$ at least provides an accurate measure of the *relative distances* of galaxies. For example, if galaxy A is found to have $v_{\rm rec} = 12\,000\,{\rm km\,s^{-1}}$ and galaxy B is found to have $v_{\rm rec} = 18\,000\,{\rm km\,s^{-1}}$ then, regardless of the value of H_0, we can say

$$\frac{d_B}{d_A} = \frac{18\,000}{12\,000} = 1.5\,. \tag{9}$$

We can therefore make accurate maps of the galaxy distribution on large scales using the measured redshift to indicate the relative separation of galaxies. We call such a map a **redshift survey**. Currently about a million galaxy redshifts have been measured, and in the next few years several new redshift surveys will significantly increase this number.

Redshift surveys reveal patterns in the galaxy distribution. In particular we see

- galaxy clusters,
- sheets and filamentary structure,
- voids (i.e., regions which are empty of galaxies).

Fig. 6 shows a famous 'slice' from the Harvard CfA redshift survey and results from the more recent, and deeper, Sloan survey. The largest of these features we can presently see is the Sloan Great Wall. This shows a recession velocity range of about $30\,000\,{\rm km\,s^{-1}}$ which, from Hubble's law, corresponds to a size of about 430 Mpc (1.4 billion light years). Cosmologists use various different statistical methods to quantify the degree of structure and clustering in redshift surveys; these methods lie beyond the scope of this course (and will be discussed in the honours cosmology course), but when applied to the most recent redshift surveys they show that on scales larger than about $30\,000\,{\rm km\,s^{-1}}$ the Universe begins to looks uniform and **homogeneous**.

2.2 Galaxy clustering hierarchy: summary

On small scales galaxies are grouped together in **clusters**. Small clusters may contain about 10 galaxies; the largest rich clusters contain several thousand galaxies. Unlike constellations, galaxy clusters are not simply chance occurrences or 'line of sight' effects: the galaxies in a cluster are believed to have been formed together at the same epoch[e], and are gravitationally bound together – a property which allows the mass of a galaxy cluster to be estimated

[e]Cosmologists often use the word *epoch* to mean a common moment in time.

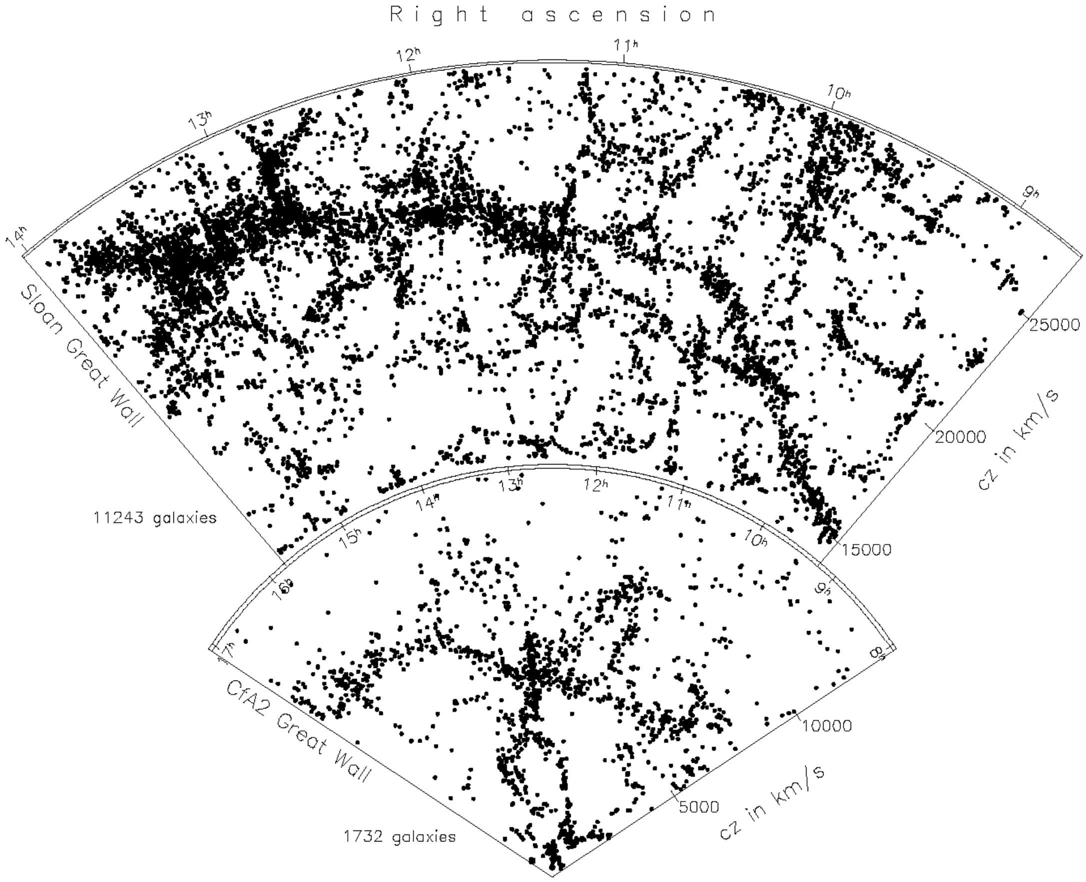

Figure 6: Slices from the CfA and Sloan galaxy redshift surveys. Note the 'walls' of galaxies in each, indicating strong clustering (from astro-ph/0310571).

(see Section III). *Within* galaxy clusters, galaxies may have a large **peculiar motion**,or speed, that differs slightly from their recession velocity given by Hubble's law. This is caused by their gravitational interaction with the other cluster members. The effects of this are most pronounced for galaxies that are reasonably close, and have relatively low recession velocities due to the Hubble flow.

The distribution of galaxy clusters is also non-uniform. Galaxy clusters are themselves clustered, and are organised into larger-scale structures which we refer to as **superclusters**. The filamentary structures observed in redshift surveys (such as CfA, Las Campanas or 2dF) delineate superclusters.

The Milky Way is part of a small cluster of about 30 galaxies which is called the **Local Group**. The Local Group is roughly disc-shaped and about 2 Mpc in diameter. The dominant members of the Local Group are the Milky Way and the Andromeda galaxy, M31, which lies at a distance of about 2.2 million light years, or 0.67 Mpc. These two spirals dominate the dynamics of the Local Group and contain the majority of the luminous matter. The remaining galaxies are mainly **dwarf ellipticals** and **irregulars**. The irregular galaxies include the **Large** and **Small Magellanic Clouds**, which are satellite galaxies of the Milky Way at a distance of about 50 kpc.

The nearest rich cluster of galaxies is the **Virgo cluster**, which contains ~2 500 galaxies. Both the mean distance and the mean recession velocity of Virgo cluster galaxies are somewhat uncertain, due to the large peculiar motions of galaxies in the cluster (this has contributed to the uncertainty in the value of H_0 measured using Cepheids). This problem is compounded by the fact that the Virgo cluster is thought to be highly elongated along the line of sight. Recent Hubble Space Telescope data suggests that the core of the Virgo cluster lies at a distance of about 18 Mpc.

The Local Group and Virgo cluster are part of the **Local Supercluster**, which is a planar concentration of many rich galaxy clusters within about 5 000 km s^{-1}. The scale of the Local Supercluster is of the order of the largest scales over which structure is observed in galaxy redshift surveys.

type	*typical scale*	*typical $N_{\rm gal}$*	*examples*
galaxy group, small cluster			Local Group, Fornax cluster
rich cluster			Virgo and Coma clusters
supercluster			Local Supercluster

2.3 Morphological segregation

Statistical analysis of galaxy redshift surveys reveals that elliptical galaxies are preferentially found in the cores of rich clusters, while spirals are generally *not* found there. This *morphological segregation* is thought to be a consequence of the galaxy formation process. It is believed that spirals existed briefly in galaxy clusters shortly after the clusters formed, but their discs could not survive the strong gravitational tidal forces in the cores of clusters. This early population of cluster spirals was, therefore, torn apart and many may have been 'cannibalised' by the giant elliptical (cD) galaxies in the centre of the clusters.

2.4 Redshift-independent galaxy distance indicators

As we have seen, for the closest galaxies the Hubble expansion law is *distorted* by peculiar motions, due to the gravitational pull of nearby galaxies, i.e.,

$$v_{\rm obs} = H_0 d + v_{\rm pec}. \qquad (10)$$

Typical magnitudes for $v_{\rm pec}$ are about 300 km s^{-1}, although in rich clusters some galaxy peculiar motions may be as much as several thousand km s^{-1}. Note that if $v_{\rm pec} = 300$ km s^{-1} and $H_0 = 71$ km s^{-1} Mpc^{-1}, then for a galaxy at distance, $d > 100$ Mpc, then $v_{\rm pec}$ is less than 5% of the **cosmic expansion velocity**, $H_0 d$ (i.e., the recession velocity predicted by Hubble's Law). Therefore:

- We can expect that Hubble's law will hold to within a few percent provided we are *not* considering galaxies in our immediate neighbourhood, where peculiar velocities significantly affect the observed recession velocities.

- If we want to measure the distance to nearby galaxies, we cannot rely on using Hubble's law to do so, i.e., we need to use galaxy distance indicators which are *independent* of redshift.

- Such redshift-independent distance indicators can also be *combined* with measured recession velocities of more distant galaxies (where Hubble's law holds accurately) in order to estimate the value of H_0.

2.5 Standard candle distance indicators

We introduced the idea of a **standard candle** in Section I, and many galaxy distance indicators are based on the standard candle principle. To recap, a standard candle is any class of object, such as a type of galaxy or star, whose luminosity (or equivalently absolute magnitude) can be assumed to be constant.

Suppose we observe a standard candle, of absolute magnitude $M = M_*$, to have an apparent magnitude, $m = m_{\rm obs}$. Then (neglecting absorption)

$$m_{\rm obs} = M_* + 5\log r + 25, \qquad (11)$$

where the distance, r, of the standard candle is measured in Mpc.[f] Hence an *estimate* of the distance to the object is

$$r_{\rm est} = 10^{0.2(m_{\rm obs} - M_* - 25)}. \qquad (12)$$

[f] You can prove this if you remember the definition of apparent magnitude and absolute magnitude. Two stars of fluxes S_1 and S_2 at Earth have apparent magnitudes related by $m_1 - m_2 = -2.5\log(S_1/S_2)$ by definition. For the same star at two distances, its flux is just proportional to one over the distance squared, so

$$m_1 - m_2 = -2.5\log(r_2^2/r_1^2) = -5\log(r_2/r_1).$$

The absolute magnitude of a star is its apparent magnitude at a distance of 10 pc, or 10^{-5} Mpc. So working in Mpc we have

$$M_* - m_{\rm obs} = -5\log(r/10^{-5}) = -5\log r - 25.$$

Of course in practice the standard candle assumption is only an approximation (not all our standard candle objects have absolute magnitude exactly equal to M_*) but the standard candle is still useful as a distance indicator provided that the *spread* in absolute magnitude from object to object is small.

If the difference between the *assumed* absolute magnitude and the *true* absolute magnitude is ΔM, then the *fractional error* in the estimated distance is

$$\epsilon = \frac{r_{\rm est} - r_{\rm true}}{r_{\rm true}} = 10^{0.2\Delta M} - 1. \tag{13}$$

So to be practically useful the standard candle should have $\Delta M \ll 1$, giving $\epsilon \ll 1$. For example, suppose $m_{\rm obs} = 15$ and M_* is assumed to be -20, then

$$r_{\rm est} = \underline{\hspace{5cm}} \tag{14}$$

$$= \underline{\hspace{3cm}} \tag{15}$$

But if M_* is *actually* -21, then

$$r_{\rm true} = \underline{\hspace{5cm}} \tag{16}$$

$$= \underline{\hspace{3cm}} \tag{17}$$

which represents a significant error ($\epsilon \simeq -0.37$).

In summary, a good standard candle distance indicator should

- have a small spread in absolute magnitude,
- be observable to large distances.

Some examples of standard candles in common use are

1. Sc spiral galaxies,
2. brightest cluster elliptical (cD) galaxies,
3. type Ia supernovae (see Section III).

2.6 Primary and secondary distance indicators

Distance indicators can also be divided into two categories:

Primary indicators can be calibrated from theory or from distances measured within our immediate neighbourhood, the Local Group. These include Cepheid variable and RR Lyrae stars (via their respective period-luminosity relations), measurements of annular stellar parallax and main sequence fitting (see stellar astrophysics course)

Secondary indicators must be calibrated using a sample of galaxies beyond the Local Group whose distances have been determined by other methods (using primary indicators). These include type Ia supernovae (as standard candles), the Tully-Fisher relation (see later) and a variety of other indicators such as galaxy luminosity class.

Here is a brief list of some good distance indicators. We will look at some of them more carefully in Section III (see also Fig. 7).

method	*type*	*range*	*calibration*
MS-fitting	primary		Pleiades
parallax	primary		absolute (Hipparcos satellite)
RR Lyraes	primary		LMC, Milky Way RR Lyraes
Cepheids	primary		LMC, Milky Way Cepheids
GCLF [a]	secondary		ellipticals in nearby rich clusters
Tully-Fisher [b]	secondary		spiral galaxies in local supercluster
standard galaxies	secondary		galaxies in nearby rich clusters
supernovae	secondary		SNIa host galaxies in local supercluster

[a] Globular clusters luminosity functions: GCs have standard candle-like luminosities.
[b] A relationship between galactic rotation speed and luminosity.

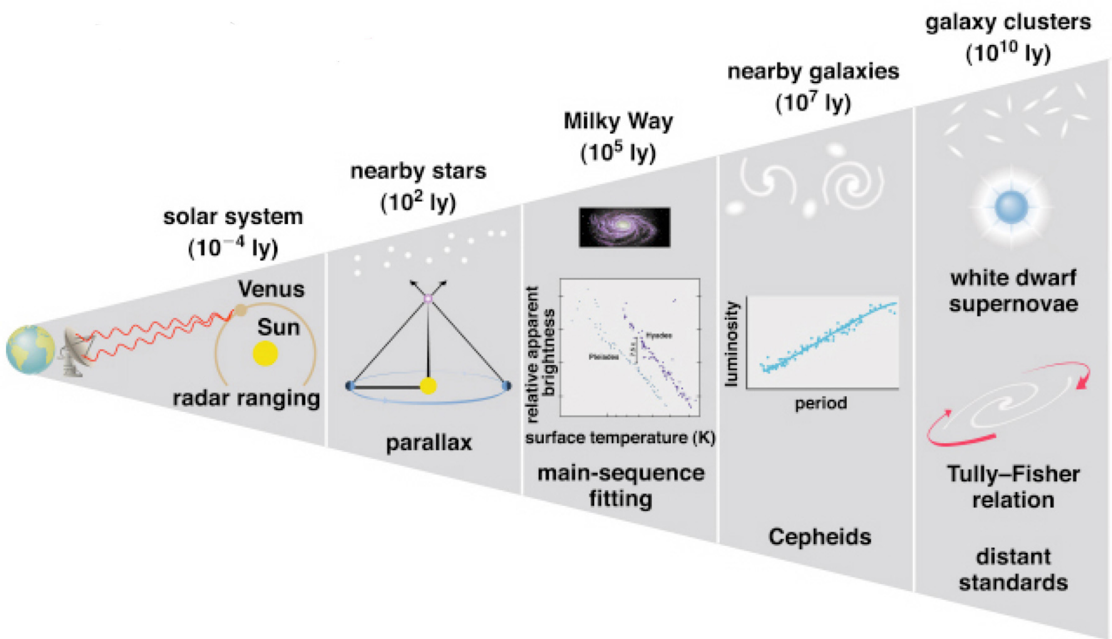

Figure 7: The cosmological distance ladder, showing how different overlapping measuring techniques allow us to measure distances out to Gpc scales (1 Gpc = 1 000 Mpc).

2.7 Improving the standard candle assumption

For some distance indicators we do not need to assume that they are standard candles, we can get an excellent indication of their luminosity by using some other directly measurable quantity which is **correlated** with absolute magnitude (which of course we cannot measure directly).

The best-known examples of such an indicator are **Cepheid variable stars**, introduced earlier. For Cepheids, a linear relationship exists between the mean absolute magnitude (averaged over the pulsation cycle of the star) and the logarithm of the pulsation period. We call this the **period-luminosity relation** (See also A1 stellar astrophysics). Therefore, if we can measure the apparent magnitude and the pulsation period of a Cepheid in a distant galaxy, we can use the PL relation to estimate the absolute magnitude of the Cepheid, and hence deduce an estimate of its distance.

The Cepheid PL relations were discovered by Henrietta Leavitt, from a plot of mean apparent magnitude against log(period) of Cepheids in the Large Magellanic Cloud. Because these Cepheids were approximately all equidistant from us, her plot translated directly into a linear relationship between absolute magnitude and log(period). Cepheid PL relations exist at all wavebands [g] between U and K. All are well fitted by the linear form

$$\overline{M} = a \log P + b, \tag{18}$$

where a and b are constants. Since the relations may be calibrated using Cepheids in the LMC and SMC (see Fig. 8) Cepheids are *primary* distance indicators.

For example, LMC data show that in the V band the PL relation is

$$\overline{M}_{\rm V} = -2.76 \log P - 1.40.$$

If the HST observes a Cepheid in another galaxy with $\overline{m}_{\rm V} = 25$ and $\log P = 1.5$, then the prediction for the star's absolute magnitude

[g] The letters U, B, V, R, I, J, H, K denote well-defined wavelength ranges, called the *Johnson wavebands*, used in manufacturing astronomical filters. The sequence follows increasing wavelength: U is in the ultra-violet and K is in the infra-red part of the spectrum. Stars and galaxies will give out different amounts of light at different wavelengths, so that measuring their apparent magnitudes through different Johnson filters will give different results. See A1 stellar astrophysics and observational astrophysics for more details on apparent magnitude filters.

is

$$
\begin{aligned}
\underline{\qquad} &= \underline{\qquad\qquad\qquad\qquad\qquad} \\
&= \underline{\qquad\qquad} \\
\underline{\qquad\qquad\qquad} &= \underline{\qquad\qquad\qquad\qquad\qquad\qquad} \\
&= \underline{\qquad\qquad\quad} \\
&= \underline{\qquad\qquad\quad}
\end{aligned}
$$

The Cepheid PL relations are not perfect, but have some statistical scatter (the true absolute magnitude at a given period does not always lie exactly on best-fit straight line). Consequently, Cepheid distance estimates are also not perfect, but are generally subject to a random error of about 10%. A more serious problem with Cepheids is the presence of *systematic* errors due to extinction (the stars can look dimmer than they should, because their light is absorbed slightly by the gas and dust along the line of sight to us). Cepheids are found in spiral arms, where extinction due to dust may be considerable. Ignoring extinction leads to an over-estimate of distance. Recent HST and ground-based observations have overcome this problem by using multicolour observations giving B, V, R, I Cepheid PL relations. The amount of extinction varies with wavelength, and so may be estimated and corrected.

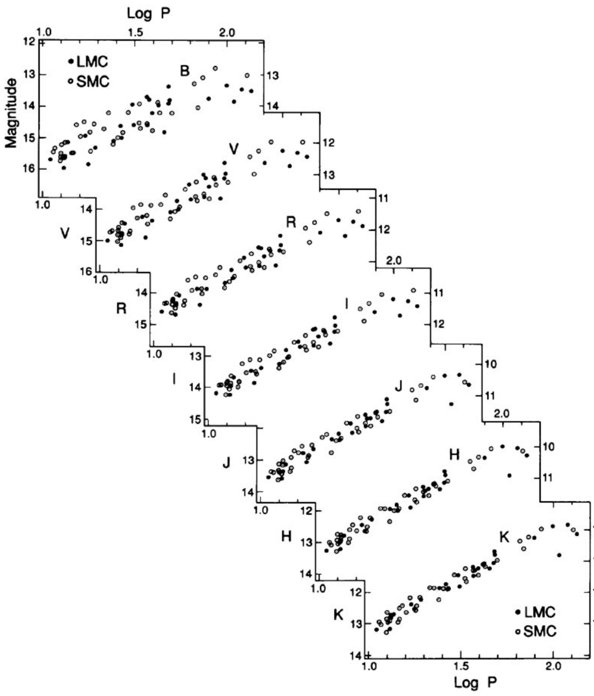

Figure 8: Cepheid period-luminosity relations in the LMC and SMC at different wavebands. Note the small variations in apparent magnitude between the different bands.