M1 Cosmology - 8 - Probes of Cosmological Structure

Nick Kaiser

January 14, 2022

Contents

1 Microwave Background Anisotropies **2**
1.1 Recombination and the Cosmic Photosphere . . . 2
1.2 Large-Angle Anisotropies . . . 2
1.2.1 The Sachs-Wolfe effect . . . 2
1.2.2 The ISW or Rees-Sciama effect . . . 4
1.3 Small-Angle Anisotropies . . . 5
1.4 Polarization of the CMB . . . 6
1.4.1 Polarization from electron scattering . . . 6
1.4.2 Polarisation from gravitational waves . . . 6

2 Mapping, and clustering analysis of, the large-scale structure **8**
2.1 3D structure from redshift surveys . . . 8
2.1.1 The cosmic web . . . 8
2.2 Statistical analysis of redshift surveys . . . 8
2.2.1 The 2-point correlation function - definition . . . 8
2.2.2 Estimating the 2-point function . . . 10
2.2.3 The power spectrum of galaxy clustering . . . 11
2.2.4 Higher order n-point functions . . . 13
2.2.5 Biased galaxy clustering . . . 13
2.3 Peculiar velocities and redshift space distortion . . . 14
2.4 Baryon Acoustic Oscillations . . . 15

3 Cosmic flows as a probe of large-scale structure **16**
3.1 Measuring distances to galaxies . . . 16
3.1.1 The 'knee' in the luminosity function . . . 17
3.1.2 Applications: the infall to the Virgo supercluster and the Rubin-Ford effect . . . 17
3.2 Distance measurement using internal 'scaling relations' of galaxies . . . 18
3.2.1 The Tully-Fisher relation . . . 18
3.2.2 The fundamental plane for elliptical galazies . . . 18
3.3 The Hubble expansion rate and the 'great attractor' . . . 19

4 Gravitational Lensing **19**
4.1 Introduction . . . 19
4.1.1 History of gravitational lensing . . . 19
4.1.2 The 'lumpy glass' analogy . . . 21
4.2 Some applications of gravitational lensing in cosmology . . . 21
4.2.1 The Einstein radius and the critical surface density . . . 21
4.2.2 Micro-lensing of stars by compact halo objects . . . 23
4.2.3 Micro-lensing at cosmological distances . . . 24
4.3 Strong lensing by galaxies and galaxy clusters . . . 24
4.4 Gravitational lensing by galaxy clusters . . . 25
4.4.1 Cluster masses from giant arcs . . . 25
4.4.2 Caustics and critical curves . . . 25
4.4.3 The optical depth for strong lensing . . . 27

4.4.4 Amplification bias and quasar galaxy associations . . . 27
4.5 Weak Lensing . . . 29
4.5.1 Introduction . . . 29
4.5.2 Order of magnitude of the WL effect . . . 29
4.5.3 State of the art of cosmic shear measurements . . . 31

5 Cosmology with clusters of galaxies 31
5.1 Overview of cluster-cosmology . . . 31
5.2 George Abell's cluster catalog . . . 33
5.3 Cluster Masses from Galaxy Motions . . . 34
5.4 Clusters observed in X-rays . . . 35
5.4.1 Thermal bremsstrahlung . . . 35
5.4.2 Thermal bremsstrahlung from galaxy clusters . . . 37
5.4.3 Cooling flows in clusters . . . 38
5.5 The Sunyaev-Zel'dovich effect . . . 38
5.5.1 The thermal SZ effect . . . 39
5.5.2 The kinematic SZ effect . . . 39
5.5.3 Weak lensing and the bullet cluster . . . 40

1 Microwave Background Anisotropies

An extremely important probe of linear structures are *cosmic microwave background anisotropies*. This is rather different from galaxy clustering and bulk flows in that the anisotropy is generated at early times. Small angle anisotropy, for instance, probes the state of the universe at the redshift of recombination $z_{\rm dec} \simeq 1000$. This gives the benefit that the density fluctuations are well described by linear theory, which simplifies matters greatly.

1.1 Recombination and the Cosmic Photosphere

First we need to understand where the photons we are seeing originated. In the standard cosmology the universe was highly ionized prior to $z_{\rm dec}$ and then rather rapidly became neutral. Detailed calculations show that the 'visibility function' — this gives the distribution over redshift of last scattering for CMB photons — is correspondingly narrow, with width $\Delta z/z \sim 1/10$. In an Einstein - de Sitter cosmology, the comoving distance ω is related to redshift by

$$\omega = 1 - \frac{1}{\sqrt{1+z}}. \tag{1}$$

This means that the horizon (the surface of infinite redshift) is at $\omega = 1$, whereas the surface $z = z_{\rm dec}$ is at $\omega_{\rm dec} \simeq 1 - 1/\sqrt{1000} \simeq 0.97$. The *cosmic photosphere* is then a rather narrow fuzzy shell quite close to the horizon. If the universe is spatially flat then the angle subtended by the horizon size at decoupling is $\theta_{H,{\rm dec}} \simeq 1/\sqrt{z_{\rm dec}}$ or about 2 degrees. If the universe is open, this angular scale is reduced.

1.2 Large-Angle Anisotropies

1.2.1 The Sachs-Wolfe effect

Consider first the anisotropy generated by perturbations larger than the horizon size at decoupling. The situation is sketched in figure 1. The primary effect driving large-angle anisotropy was first calculated by Ray Sachs and Art Wolfe and is called the *Sachs-Wolfe effect*. Photons which emerge from an over-dense (under-dense) region sufffer what is effectively a gravitational redshift (blueshift). However, for subtle reasons, the fractional photon energy change — which is equal to the fractional change in temperature — is one third of the dimensionless potential perturbation at the point of emission. The Sachs-Wolfe temperature anisotropy is

$$\frac{\Delta T}{T} = \tfrac{1}{3}\Phi_{\rm em} \tag{2}$$

where $\Phi_{\rm em}$ is the dimensionless gravitational potential perturbation – the thing appearing in the weak-field metric perturbations – at the point of emission. The temperature anisotropy is therefore on the order of the density perturbation amplitude at horizon crossing $\delta_{\rm H}$.

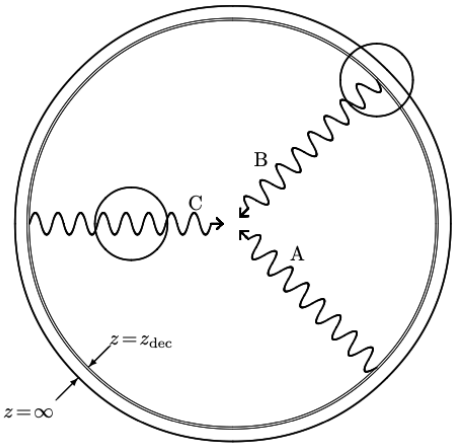

Figure 1: Schematic illustration of the generation of large angle CMB anisotropy. Photon A arrives from an unperturbed region of the universe. Photon B emerges from an over-dense region. It suffers the 'Sachs-Wolfe' effect — effectively a gravitational redshift — and has slightly lower energy than photon A when it reaches the observer. Photon C passes through an over-dense region on its way form the last scattering surface, and suffers what is known either as the the ISW or 'Rees-Sciama' effect.

The large-scale temperature anisotropy effectively provides – via the Sachs-Wolfe effect - a map of the Newtonian gravitational potential variation on the last scattering surface. The prediction of inflation (and also cosmic string models) is that the temperature fluctuation field should take the scale-invariant flicker-noise form corresponding to spectral index $n = 1$. This is just what was observed by the COBE satellite, as seen in figure 2. The resolution of the map is set by the angular resolution of the telescope, and the signal-to-noise ratio is relatively high, so COBE provided, in essence, the definitive map of the large-angle anisotropy. Subsequent missions have, however, added more information on smaller-scale anisotropies and have provided polarisation as will as temperature information.

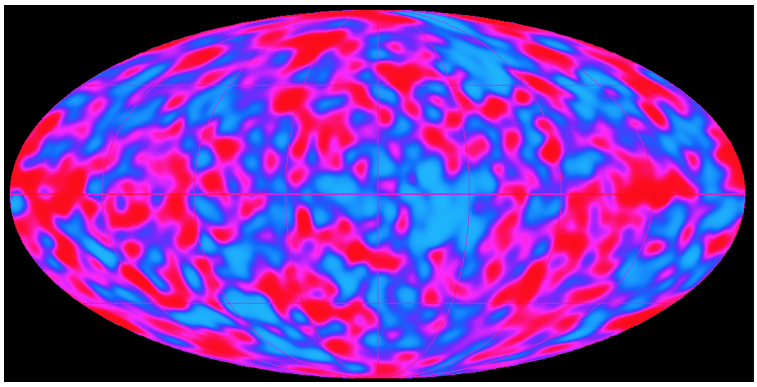

Figure 2: The map of the temperature anisotropy $\Delta T(\theta, \phi)$ of the CMB made by the COBE satellite mission. The fluctuations are on the order of one part in 10^5. A larger dipole anisotropy, thought to be caused primarily by our motion of a few hundred km/s with respect to the 'cosmic rest-frame' has been subtracted. This is an 'Aitoff projection'; one can think of this crudely as being what you get if you unpeel an orange and then squash it flat – and in the process distorting it somewhat – while maintaining constant area.

This sky map was analysed by decomposing the temperature field into spherical harmonics as

$$\Delta T(\theta, \phi) = \sum_l \sum_m a_{lm} Y_{lm}(\theta, \phi) \tag{3}$$

where the (real part of the) mode functions – which are given by Legendre polynomials $P_l^m(\cos\theta)$ times $e^{im\phi}$ – for $l = 0, 1, 2, 3$ are shown in figure 3.

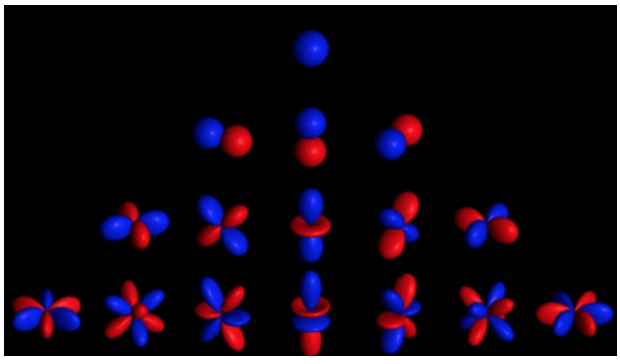

Figure 3: Spherical harmonic functions $Y_{lm}(\theta, \phi)$ for, from the top, $l = 0, 1, 2, 3$. These are functions are orthogonal to one another in that $\int d\Omega\, Y_{lm}\, Y^\star_{l'm'}$ vanishes unless $l = l'$ and $m = m'$, and are analogous to Fourier modes in that regard. One can say that the Y_l^m are the 'normal modes' of the sphere. The primary index l sets the angular scale. For each l, there are $2l+1$ different modes with m running from $-l$ to $+l$. A scalar function like $T(\theta, \phi)$ can be decomposed into a sum of $Y_{lm}(\theta, \phi)$ multiplied by complex amplitudes $a_{lm} = \int d\Omega\, T(\theta, \phi)\, Y^\star_{lm}(\theta, \phi)$.

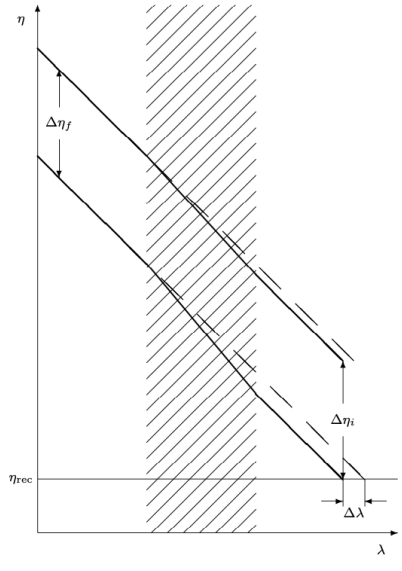

Figure 4: Schematic illustration of the generation of large angle CMB anisotropy from the ISW or 'Rees-Sciama' effect where we see CMB photons that have propagated through a structure along the line of sight; hatched region in this conformal space-time plot. The weak field metric – appropriate for the structures we see in our Universe – is $ds^2 = a^2(\eta)(-(1+2\Phi)d\eta^2+(1-2\Phi)|d\boldsymbol{\lambda}|^2)$. This means that the coordinate velocity is $|d\boldsymbol{\lambda}/d\eta| = 1+2\Phi$. If Φ is unchanging then the observed interval of conformal time $\Delta\eta_{\rm f}$ between reception of pulses is the same as the interval between their emission $\Delta\eta_{\rm i}$. And as the proper time intervals are given by $\Delta\tau = (a/c)\Delta\eta$, the observed and emitted periods of light are related by $\tau_{\rm f}/\tau_{\rm i} = a_{\rm f}/a_{\rm i}$ just as without any perturbation. But if the potential Φ is changing with time there is a non-vanishing effect. Since $\eta_{\rm f} = \eta_{\rm i} + \int d\lambda(1-2\Phi)$, the observed to emitted period ratio is $\tau_{\rm f}/\tau_{\rm i} = (a_{\rm f}/a_{\rm i})(d\eta_{\rm f}/d\eta_{\rm i}) = (a_{\rm i}/a_{\rm f})(1-2\int d\lambda\partial_\eta\Phi)$. Equivalently, the fractional period change is $\delta\tau_{\rm f}/\tau_{\rm f} = -2\int d\lambda\partial_\eta\Phi$. For a perturbation of size R with $\dot\Phi \sim H\Phi$ – i.e. changing on a cosmological timescale – $\delta\tau_{\rm f}/\tau_{\rm f} \sim (HR/c)\Phi$.

The average of the squared amplitudes for harmonic l gives the angular power-spectrum

$$C_l = \frac{1}{2l+1}\sum_m |a_{lm}|^2 \tag{4}$$

which gives the contribution to the variance $\langle(\Delta T)^2\rangle$ from the harmonic l. And just as, in 3D, $k^3P(k)$ gives the power per logarithmic interval of wave-number k, l^2C_l gives the power per logarithmic interval of l; so l is like an angular wave-number. Since inflation predicts scale-invariant potential fluctuations – equal contribution to the variance from each logarithmic interval of spatial frequency – the prediction is that l^2C_l should be independent of l also.

The main result of the COBE measurement was:

- Direct measurement of the large scale gravitational potential field on the last-scattering surface on large ($\gtrsim 10^\circ$) scales.
- From which the power-spectrum was found to be essentially scale invariant: $l^2C_l \simeq$ constant.
- And the variance was in accord with that predicted by extrapolating from clustering measurements at much smaller scale assuming spectral index $n = 1$ – or very close thereto – as predicted by inflation.
- Convincing evidence that, on very large scales approaching the current Hubble scale $ct \sim c/H$, the universe extremely homogeneous: $(\delta\rho/\rho)_\lambda \lesssim 10^{-4}(c/H\lambda)^2$
  - in accord with the 'cosmological principle'
  - and this, one might think, would have put paid to the long standing speculation that our universe might be 'fractal-like' on large scale

1.2.2 The ISW or Rees-Sciama effect

In addition to the temperature anisotropy induced as the photons 'climb out' of the potential well where they originate, there is an additional source of temperature anisotropy that shows up at large angular scales. This is generated as the photons pass through intervening inhomogeneity (photon C in figure 1). This effect was discussed in linear theory by Sachs and Wolfe and is usually called the integrates Sachs-Wolfe (ISW) effect . It was also analysed by Martin Rees and Dennis Sciama around the same time, and they also discussed how non-linear structures would cause this effect. How the effect arises is illustrated in figure 4.

This effect, however, is expected to be important, if at all, only on very large scales. First, in a flat universe, and in linear theory, the potential is unchanging and the effect vanishes. In an open universe, or one with dark energy, there is a non-vanishing effect, but the amplitude is on the order $\Delta T/T \sim (HR/c)\Phi$, where R is the size of the perturbation; this is smaller than the Sachs-Wolfe effect by the factor $\sim HR/c$ which is the size of the perturbation in units of the horizon size. Now the effect of multiple perturbations

along the line of sight will add in quadrature, resulting in a net anisotropy $\Delta T/T \sim \sqrt{HR/c}\Phi$, but this is still small compared to the Sachs-Wolfe effect unless we are dealing with perturbations with $HR/c \sim 1$.

It may be possible to measure the effect statistically by cross-correlating the temperature fluctuations with the projected density of galaxies[1] and thereby probe the rate at which the potential is changing. People have tried to do this and have claimed varying degrees of success. There is also a somewhat anomalous 'cold-spot' on the microwave sky, and it appears that there is a deficit of galaxies along this line of sight (which would give an effect of the right sign). But this explanation is challenged by the fact that the cold-spot is on the order of a few degrees across, which is hard to reconcile with the ISW effect, whose nature is to be be sub-dominant compared to the primary (i.e. SW) anisotropy for any individual line of sight[2].

1.3 Small-Angle Anisotropies

Since the COBE measurement, the angular resolution of observations has increased dramatically, as illustrated in figure 5. This shows the results from the major space missions WMAP and Planck, though it there were also pioneering results from the ground and from balloon borne missions.

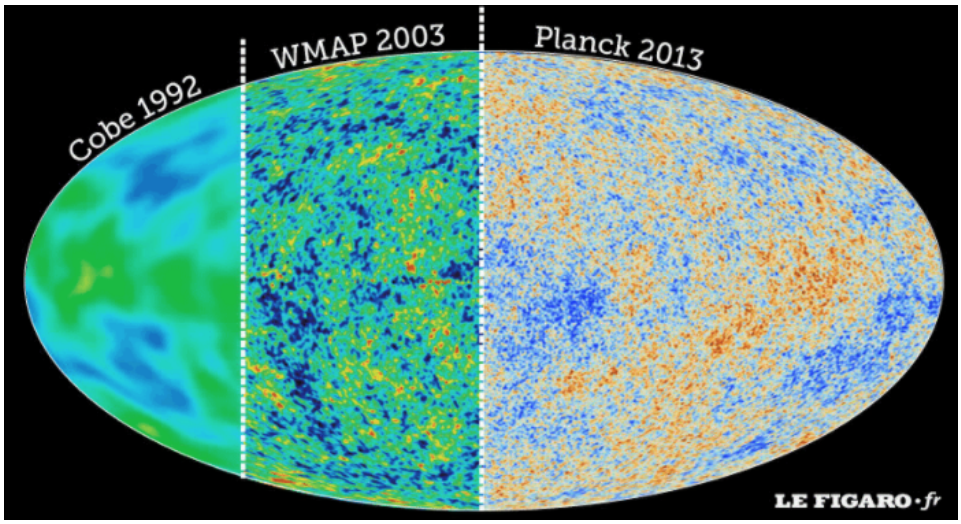

Figure 5: Maps of the temperature anisotropy of the CMB. On the left is part of the map from the COBE satellite; part of the full-sky map shown above. In the middle is shown part of the sky as measured by the WMAP satellite and on the right is the more recent map from Planck.

These later missions were important as they were able to measure the power spectrum on angular scales comparable to the maximum comoving Jeans length – i.e. the comoving sound horizon at around decoupling – at which scale we expect to see features in the spectrum that are strongly dependent on the matter content of the universe. The current state of the art of measurements of the power spectrum of temperature fluctuations is shown in figure 7 produved by the Planck collaboration.

As discussed in the last lecture, prior to de-coupling, and on scales small compared to the sound horizon, the baryon-photon fluid undergoes acoustic oscillations. In such sound waves, there are adiabatic variations of the temperature $\Delta T/T \sim \Delta\rho/\rho$ and there are also peculiar velocities $v \sim c_s \Delta\rho/\rho$. Both of these give rise to temperature anisotropy, the latter through the Doppler effect (the effect of the gravitational potential being less important). Detailed calculation of small-scale anisotropy is quite involved as it involves solving the collisional Boltzmann equation for the photons undergoing scattering off electrons and being coupled to gravity and must be performed numerically. An example of such a calculation is shown as the red curve in figure 7. As one can see, the agreement with the theory – admittedly with several parameters adjusted to fit the observation – is impressive.

The sensitivity of these measurements to cosmological parameters is illustrated in figure 7 from Hu and Dodelson (2002). As we might have anticipated, the amplitude of the wiggles in the power spectrum are

[1] A back of the envelope calculation suggests this is hard, but doable. Imagine one has a deep all sky redshift and one divides the sky up into $N_{\rm p}$ patches of angular size θ, corresponding to a physical scale $L \sim \theta c/H$, at which scale we should have $N \sim 1/\theta$ independent density perturbations along the line of sight and therefore an RMS ISW signal $\Delta_{\rm ISW} \sim \sqrt{HR/c}\Phi \sim \Delta_{\rm SW}/\sqrt{N}$. The expectation is that the $N_{\rm p}/2$ patches with an excess (deficit) of galaxies will be hotter (colder) by $\Delta T/T \sim \Delta_{\rm ISW}$. The primary temperature fluctuations $\Delta_{\rm SW}$ provide a 'cosmic variance', whose square root will be on the order of $\Delta_{\rm SW}/\sqrt{N_{\rm p}}$. But $N_{\rm p} \sim 4\pi/\theta^2 \sim 1/N^2$, so the cosmic variance is smaller than the expected signal.

[2] It has been argued that, in selecting the most extreme cold-spot on the entire sky – out of thousands of patches of this size – one would be selecting the line of sight for which the ISW effect is particularly large. But one would also be picking that patch where the primary anisotropies are particularly large. So one would have to appeal to non-gaussianity of the ISW effect for this argument to have any weight.

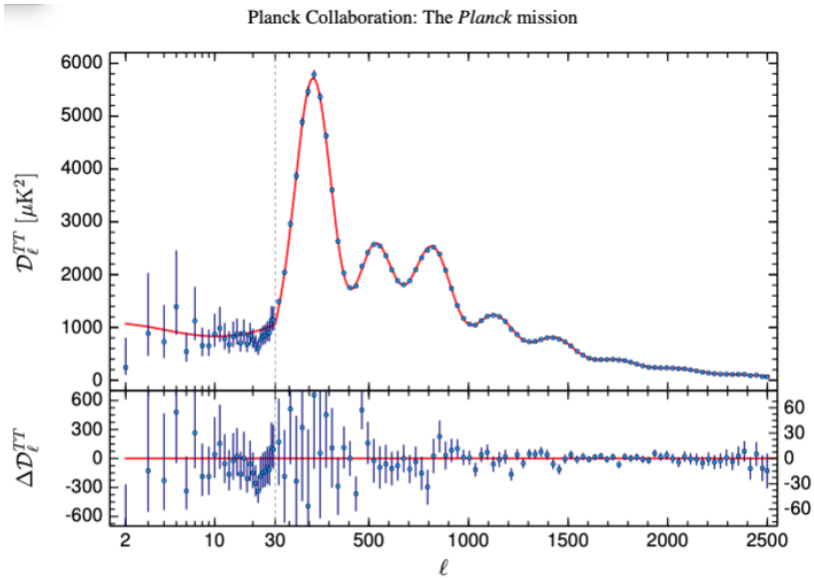

Figure 6: Planck temperature power spectrum. The quantity plotted in the upper panel is $l(l+1)C_l$, which gives the variance per log-l. At low l the power for each angular frequency is shown. To the right of the dashed line, groups of C_l's have been averaged together (and the frequency scale becomes linear). On the left we see the Sachs-Wolfe, essentially as seen by COBE. The first peak at $l \sim 100$ was well measured by WMAP, and Planck has added the extra peaks. The amplitudes of these provide a sensitive measurement of $\Omega_{\rm m}$ and $\Omega_{\rm b}$ while the location of the peaks is very sensitive to curvature.

larger the larger is the baryon content. In the upper panels one can see that the ISW fluctuations are enhancing the temperature fluctuations at very low l. In the upper left panel we see, on the right, that the effect of curvature is essentially to shift the spectrum right or left. Since the angular frequency is on a logarithmic scale that means that the effect is just a change of the angular scale. This is illustrated in figure 8.

1.4 Polarization of the CMB

1.4.1 Polarization from electron scattering

In addition to the temperature anisotropy $\Delta T/T$, the CMB is predicted to display polarisation. Polarisation of the CMB arises primarily by virtue of the anisotropic nature of Thomson scattering; if an electron is illuminated with radiation which is natural – i.e. unpolarised – but anisotropic then the radiation it scatters will be polarised. This is illustrated in figure 9. For example, if we lie along the z-axis and observer an electron at the spatial origin which is being illuminated by a lamp sited out along the y-axis then the radiation we see will be linearly polarised in the x-direction. In general, the degree of polarisation is proportional to the quadrupole moment of the incident radiation. Measurement of the CMB polaris ation therefore provide us with a kind of 'remote-sensing' of the anisotropy of the radiation field as it was at the time of last scattering.

This effect was first calculated in 1983, and it was shown that the primary CMB radiation is weakly linearly polarised on angular scales corresponding to the Sakharov oscillations; the degree of polarisation being roughly 5-10% of the fractional temperature anisotropy.

There is also 'secondary' polarisation that is generate by electron scattering when the hydrogen becomes re-ionised at the end of the 'dark-ages' at $z \sim 10$ when the first stars and galaxies form. Both of these effects have been measured by Planck.

1.4.2 Polarisation from gravitational waves

This type of polarisation has a very interesting characteristic. While linear polarisation has, in general, two degrees of freedom, the polarisation pattern produced by electron scattering has only one. If you have electrons around a light source, the polarisation you see will be purely tangential. An alternative pattern, in which the polarisation vector is rotated by 45 degrees cannot be generated this way. This means that the a polarisation map can be decomposed into two components – known as the 'electric' and 'magnetic' components – and the latter, for electron scattering, should vanish.

One way to see this is to think about a planar sound wave; the quadrupole moment of the radiation an electron sees is aligned with the wave-vector $\mathbf{k}$. That means that the linear polarisation of the scattered radiation an observer sees will have the electric field aligned with the projection of $\mathbf{k}$ 'onto the sky'. The component of the polarisation at 45 degrees to the projection of $\mathbf{k}$ vanishes.

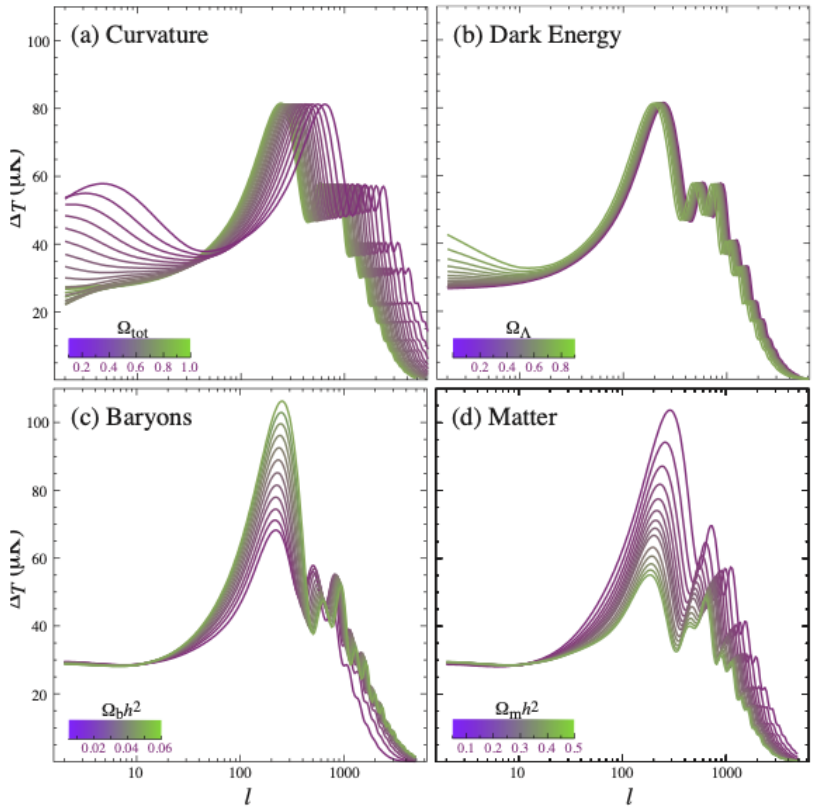

Figure 7: The dependence of C_l on the four most important parameters in the ΛCDM model by Hu and Dodelson (2002). The fiducial model had $\Omega_k = 0$, $\Omega_\Lambda = 0.65$, $\Omega_{\rm m}h^2 = 0.147$ and $\Omega_{\rm b}h^2 = 0.02$. At the upper left one can see that, at high l, introducing curvature shifts all frequencies (as discussed more below) while at low l is seen the ISW contribution if there is negative curvature. The ISW effect is also apparent in models with stronger dark energy. The strengths of the peaks provide constraints on the matter density parameters. Other important parameters are the present day Hubble parameter and the optical depth to scattering when the universe became reionised around $z \sim 10$ and which acts to decrease the amplitude of the primary temperature perturbations.

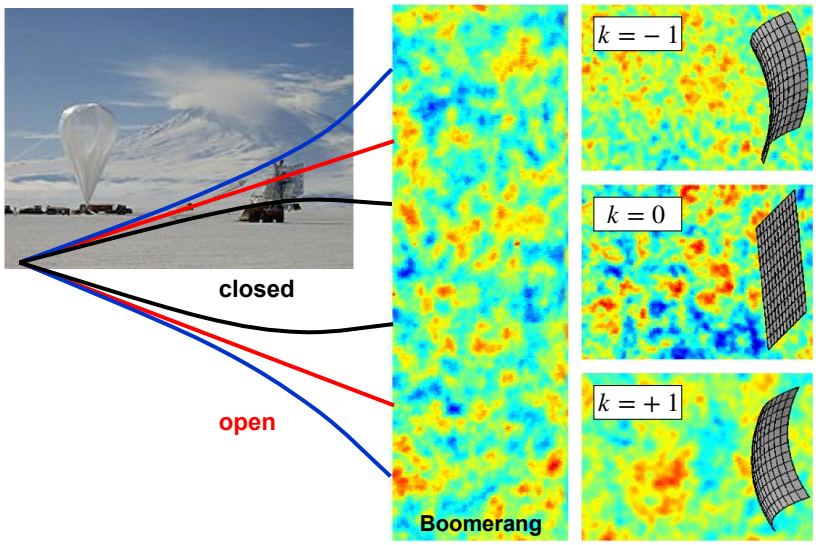

Figure 8: The BOOMERanG experiment (Balloon Observations Of Millimetric Extragalactic Radiation ANd Geophysics) gave the first convincing measurement of the Sakharov peaks in the CMB spectrum. The results, published in 2000, showed that the angular scale of these requires that our universe be spatially flat. This was more or less contemporaneous with the measurements, using Type a supernovae as standard candles, that showed the expansion of the universe is beginning to accelerate. It was the combination of these that argued convincingly for the existence of dark energy.

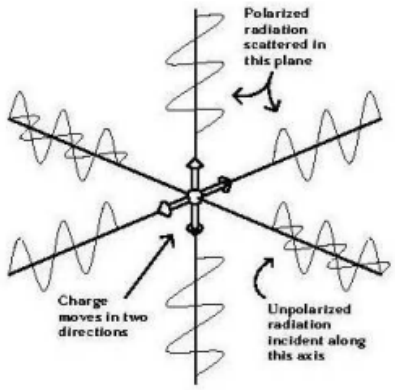

Figure 9: Polarisation of the CMB is generated primarily by electron scattering of anisotropic radiation. As illustrated, if one has natural radiation incident from a single direction $\hat{\mathbf{n}}$, it will cause the electrons to oscillate in the plane perpendicular to $\hat{\mathbf{n}}$. The scattered radiation can be calculated classically, considering the electron to be a small oscillating current. An observer in that plane will see radiation that is fully linearly polarised as the electrons only move in one direction perpendicular to the line of sight. Observers looking along other lines of sight will see partially linearly polarised. In general the degree of linear polarisation is proportional to the quadrupole anisotropy seen by the electron.

The smallness of the magnetic component is seen in the Planck data. But it is not, in general, exactly zero. That is because polarisation can also be generated by gravitational waves. A gravitational wave, unlike a longitudinal sound wave, causes the electrons to see radiation with a quadrupole moment that is perpendicular to the wave-vector, and there are two components; the '+' and × components. And these will, in general excite both the E and B modes.

Gravitational waves are produced in inflation in much the same way as zero-point fluctuations of the modes of the scalar field arise. The energy of a Newtonian gravitating system can be written as $E = (8\pi G)^{-1} \int d^3r \boldsymbol{\nabla}\varphi \cdot \boldsymbol{\nabla}\varphi$, or as $E = \int d^3r \mathcal{E}$ with energy density $\mathcal{E} = \boldsymbol{\nabla}\varphi \cdot \boldsymbol{\nabla}\varphi / 8\pi G$. The energy density of gravitational waves is, analogously, given by the squared gradient of the metric perturbations. This leads to a prediction for the minimal zero-point background of gravitational radiation at horizon exit during inflation. These are frozen in until the modes re-enter the horizon, at which time they can excite polarisation. The prediction is simpler than that of the scalar density perturbation modes (which depend on $\dot{\phi}$) as their strength $h_{\mu\nu}$ is essentially just the square of the energy scale of inflation in natural units. So perhaps around 10^{-6} for GUT-scale inflation. In 2014 there was a claim by the BICEP2 collaboration to have measured B-mode polarisation in the CMB. This generated considerable excitement in the community. Unfortunately, it was later realised that what was seen can be attributed to radiation from dust in the galaxy; the dust grains being aligned by the magnetic field. There is now considerable effort to try to find ways to remove this contaminating 'foreground' and reveal the cosmological B-modes, but this is very challenging.

2 Mapping, and clustering analysis of, the large-scale structure

2.1 3D structure from redshift surveys

2.1.1 The cosmic web

- In framework of FRW(L) models redshift $\lambda_{\rm obs}/\lambda_{\rm em} - 1 = z \quad \Rightarrow \quad$ distance
  - linear for $z \ll 1$: $cz \simeq H_0 D \quad \Rightarrow \quad D = cz/H_0$
  - significant cosmology dependent modifications for $z \sim 1$ or greater
- allows 3D maps of the universe
  - galaxies are clustered
  - reveals complex *cosmic web*:
    - there are 'voids' surrounded by 'sheets' of galaxies
    - the sheets are bounded by (and flowing, or 'draining', into) 'filaments'
    - and the filaments are flowing into clusters

2.2 Statistical analysis of redshift surveys

- the observed and simulated distributions – based on predictions from inflationary models – appear very similar
- this can be quantified, and turned into a powerful test of theory, by extracting so-called 'low-order' statistics of the spatial distribution
  - this was pioneered by Jim Peebles, who worked with Marc Davis, Margaret Geller who along with John Huchra made the first redshift surveys of sufficient size to make this analysis possible

2.2.1 The 2-point correlation function - definition

- the lowest order statistic is called the *'2-point correlation function'* denoted by $\xi(r)$:
  - it may be defined as the *fractional excess probability* to find a galaxy at a distance r from another galaxy as compared to what one would find for a unclustered random 'Poissonian' distribution
- it is also equivalent to the *autocorrelation function* of the *galaxy density contrast field*

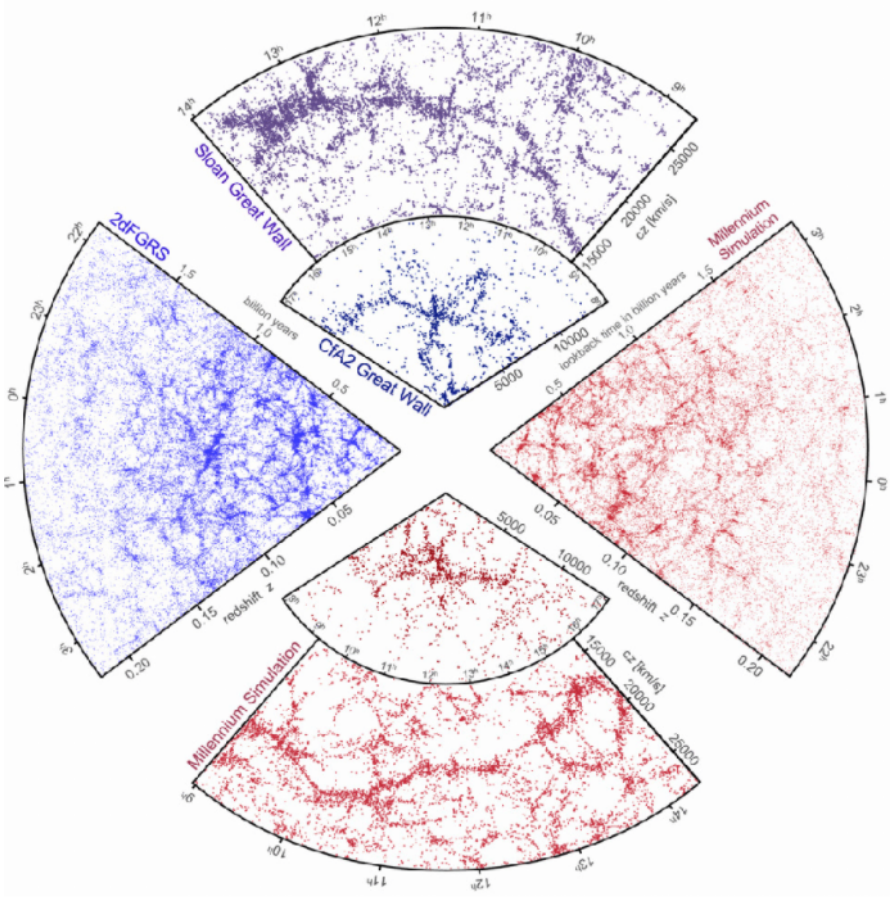

Figure 10: Redshift surveys in the real world and in simulations. North and west sectors show real data. South and east are from 'N-body' simulations. To make these, galaxy 'targets' are first selected using imaging surveys. Selection is usually based on apparent magnitude, though some surveys use colour selection either to target specific types of galaxies or particular ranges of redshift. The small blue 'pie slice' is the CfA survey performed in the '80s by John Huchra and colleagues. The redshifts here were collected painstakingly object by object. Technological advances followed: multi-object spectroscopy became possible; an example being the 2df (2 degree field) multi-fibre system on the 4m AAT telescope in Australia. Another leap forward came with the Sloan Digital Sky Survey (SDSS) which used a custom built telescope.

- in the 'Poisson sample' model, we imagine space to be divided into infinitesimal volumes or 'cells' δV_i labelled by an index i and with 'occupation numbers' n_i
- the probability that the cell at position $\mathbf{r}_i$ is occupied is
  - $P(n_i = 1) = n(\mathbf{r}_i)\delta V_i$
- and the probability that it is empty is (to 1st order in δV)
  - $P(n_i = 0) = 1 - n(\mathbf{r}_i)\delta V_i$
- where the continuous field $n(\mathbf{r})$ is the *'galaxy number density field'* from which the galaxies are assumed to be a 'sample'
  - underlying this model is the idea that there is a very strong 'stochastic' element to galaxy formation
    - i.e. we do not assume that there was some initial smooth distribution of stars that got assembled somehow into galaxy aggregations
    - rather one might think that galaxies are, to some approximation, assigned luminosities in a random manner drawn from the luminosity function
    - but it needs to be kept in mind that this is only a model and reality is, no doubt, more complicated
- the mean occupation number for a randomly chosen cell can be defined operationally as an average over the cells in our survey volume
  - $\langle n_i \rangle = \sum_i (1 \times P(n_i = 1) + 0 \times P(n_i = 0)) / \sum_i = \delta V \sum_i n(\mathbf{r}_i) / \sum_i = \overline{n}\delta V$
  - where
  - $\overline{n} = \sum_i n(\mathbf{r}_i) / \sum_i \Rightarrow \int d^3r\, n(\mathbf{r}) / \int d^3r.$
- or as an 'ensemble average' $\langle n_i \rangle = \langle P(n_i = 1) \rangle$ over realisations of the density field (and realisations of the samples of galaxies for each density field realisation)
  - and we say that a survey volume is large enough to be a 'representative sample' if these give effectively the same answer

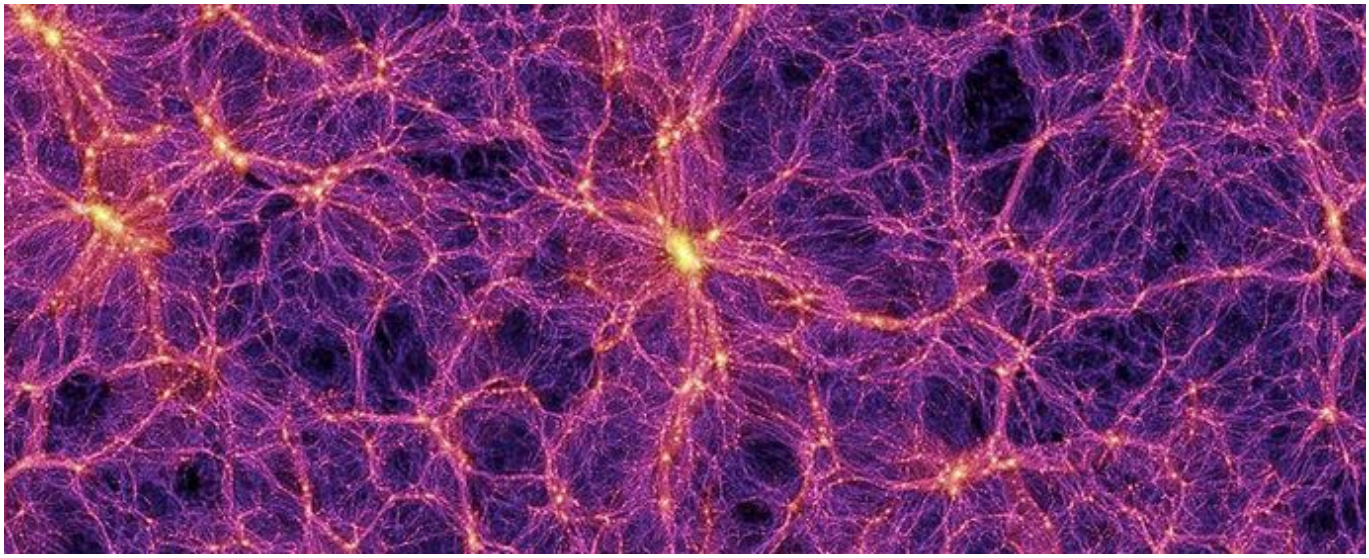

Figure 11: The cosmic web. A slice through a state-of-the-art massive N-body simulation of cosmic structure. In these calculations, billions of particles are laid down on a grid and then perturbed slightly off the grid positions in order to generate a realisation of the density perturbations predicted by inflation. These take the form of a Gaussian random field, with a power-spectrum which is an initial power-law $P(k) \propto k^n$ with $n \simeq 1.0$ multiplied by a 'transfer function' computed to model the pre-decoupling physics. The particles subsequently move according to Newtonian dynamics, which is an accurate model for structures of this kind as they are well described by 'weak-field' gravity.

- the ensemble average probability that two cells at $\mathbf{r}_i$ and $\mathbf{r}_j$ are occupied is $\langle P(n_i, n_j)\rangle \equiv \langle P(n_i = 1 \text{ and } n_j = 1)\rangle = \langle n_i n_j\rangle = \delta V_i \delta V_j \langle n(\mathbf{r}_i) n(\mathbf{r}_j)\rangle$
- and if we define the *galaxy density contrast* or the fractional density enhancement by
- $\delta(\mathbf{r}) \equiv (n(\mathbf{r}) - \overline{n})/\overline{n}$ (which implies $\langle\delta\rangle = 0$) we have
- $\langle P(n_i, n_j)\rangle = (\overline{n}\delta V)^2(1 + \langle\delta(\mathbf{r}_i)\delta(\mathbf{r}_j)\rangle)$
- and using the law for conditional probabilities $P(A, B) = P(A|B)P(B)$ we have
- $\boxed{\langle P(n_j|n_i)\rangle = \overline{n}\delta V(1 + \langle\delta(\mathbf{r}_i)\delta(\mathbf{r}_j)\rangle)}$
- which with the unconditional probability $\langle P(n_j)\rangle = \overline{n}\delta V$, and assuming that $\delta(\mathbf{r})$ is a statistically homogeneous random process, so $\langle\delta(\mathbf{r}_i)\delta(\mathbf{r}_j)\rangle$ is only a function of $|\mathbf{r}_j - \mathbf{r}_i|$, establishes that the fractional excess probability that a galaxy has a neighbour at distance $\mathbf{r}$ is $\xi(\mathbf{r}) = \langle\delta(\mathbf{r}')\delta(\mathbf{r}' + \mathbf{r})\rangle_{\mathbf{r}'}$

2.2.2 Estimating the 2-point function

- to estimate $\xi(\mathbf{r})$ from a real redshift survey is a little more involved than the model above suggests as we usually deal with flux-density limited surveys in which, in addition to spatial structure, there is an overall decreasing trend of density with distance – called the selection function $\varphi(z)$ – which is the number density of galaxies above the absolute magnitude corresponding to the flux density limit at distance z. To account for this
  - generate an identical number of 'random' points (Poisson distribution) which have the same angular and radial selection functions
  - count the number of pairs with separation in range $\{r, r + dr\}$ for real and random objects
  - the simplest estimator is: $1 + \hat{\xi}(r) = n_{\rm real}/n_{\rm random}$
- from the CfA redshift survey performed in the '8os (and subsequently confirmed by more extensive surveys – see figure 15 above) $\xi(r)$ was found to have a roughly power-law form $\xi(r) \simeq (r/r_0)^{-\gamma}$
  - with *correlation length* $r_0 \simeq 5h^{-1}$ Mpc
  - which depends somewhat on the type of galaxy (E-gals more strongly clustered)
  - and with power-law *index* $\gamma \simeq 1.8$

- We infer a great deal from this
  - for $r \ll r_0$ (say Mpc scales) $\xi \gg 1$
    - this is the regime of bound and virialised structures like clusters and groups of galaxies
  - r_0 marks the *'transition to homogeneity'*
    - on larger scales the fractional fluctuations in the numbers of galaxies are small
    - prior to this measurement there was much speculation as to whether the distribution of galaxies was 'fractal like' with strong structures on much larger scales
    - but that was no longer tenable once $\xi(r)$ was measured on scales $\gg r_0$.
  - the power-law dependence extends to at least 10's of Mpc
  - there are various numbers one can extract from this:
    - if we compute $4\pi \int_0^r dr\, r^2 \xi(r)$ – dubbed J_3 by Peebles – we find this volume is increasing roughly linearly with r
    - and if we multiply by n, the mean number density of galaxies
      - approximately $n \simeq 10^{-2}\mathrm{Mpc}^{-3}$ for $L_\star$ galaxies
    - we get a growing number for the mean excess number of neighbours that a randomly chosen galaxy has within distance r
      - reaching 100-300 for $r_{\rm max} \simeq 30 Mpc$
    - similarly, if we multiply J_3 by the mean mass density $\rho_{\rm m} = \Omega_{\rm m}\rho_{\rm crit}$ we get the mean excess mass $\Delta M(r)$ within distance r – also growing roughly like r
    - and if we compute $G\Delta M(r)/r$ (essentially J_2 times $\rho_{\rm m}$) we get an estimate of the binding energy per unit mass associated with large-scale structure
    - and this has a significant contribution from scales of order tens of Mpc

2.2.3 The power spectrum of galaxy clustering

- the Wiener-Khinchin theorem states that the power spectrum $P_\delta(\mathbf{k}) \equiv |\tilde{\delta}(\mathbf{k})|^2/V$ is the Fourier transform of the two point function $\xi(\mathbf{r}) = \langle \delta(\mathbf{r}')\delta(\mathbf{r}'+\mathbf{r})\rangle_{\mathbf{r}'}$:
  - $\boxed{P_\delta(k) = \int d^3 r \xi(r) e^{i\mathbf{k}\cdot\mathbf{r}}}$
- with inverse transform
  - $\boxed{\xi(r) = \int \frac{d^3k}{(2\pi)^3} P_\delta(k) e^{-i\mathbf{k}\cdot\mathbf{r}}}$
- this may be used to extract the power spectrum from an estimate of $\xi(r)$, but it is preferable to extract it from the data
- for a pure power-law $\xi(r) \propto r^{-\gamma}$ with $\gamma \simeq 1.8$ we would get – from dimensional considerations – a power law power spectrum with $k^3 P_\delta(k) \simeq \xi(r \sim 1/k)$ or $P(k) \propto k^{\gamma-3} \sim k^{-1.2}$
  - figure 12 confirms this to be true, down to $k \sim 1/30 Mpc$ at least, but the growth of $P(k)$ with decreasing k or increasing $\lambda = 2\pi/k$ (which a volume and essentially the same as $J_3(r)$) tapers off
- The power spectrum an autocorrelation function contain equivalent information. The power spectrum is somewhat more useful for relating to other quantities
  - one thing that can be calculated quite easily is the mean square (galaxy) density fluctuations averaged with some 'smoothing window' $W(\mathbf{r})$.
  - this is a convolution in real space
    - one can define a 'smoothed field', smoothed with a window of size R:

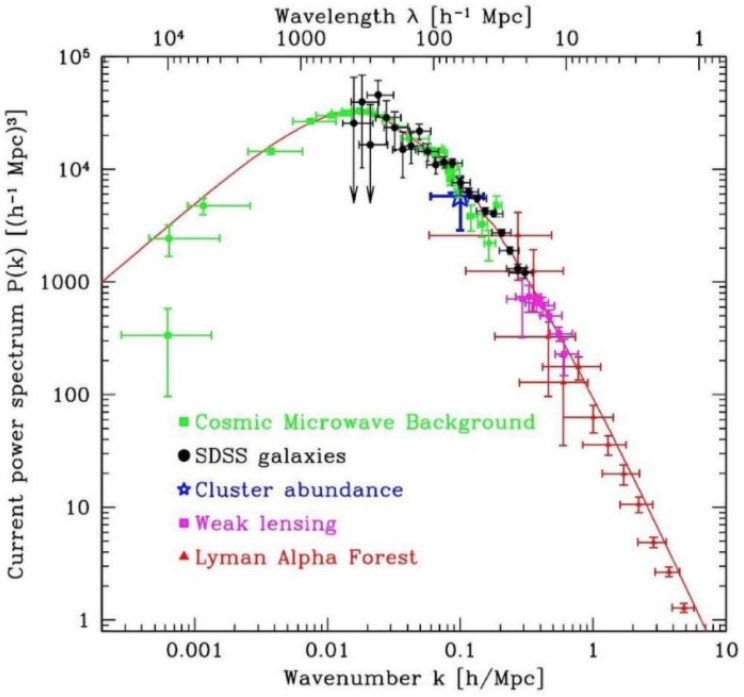

Figure 12: The power spectrum of galaxy clustering $P(k)$ from the Sloan Digital Sky Survey (SDSS) (dark points). This shows that, on the scales between $k \sim 0.02$/Mpc an $k \sim 0.5$/Mpc – wavelengths the power-spectrum is roughly power-law like with $P(k) \sim k^{-1.2}$ (as expected from the roughly power-law shape of $\xi(r)$ on the relevant scales ($r \sim 1/k$)). At the longest wavelengths probed by this survey, the slope flattens. Including the CMB measurements indicates that the power spectrum turns over to the 'primordial slope' $P(k) \propto k$ at larger scales. In the conventional model, what we are seeing is the primordial power-law spectrum modulated by a 'transfer function' $T(k)$ expressing the relative lack of growth of modes that were of order the horizon size $\lambda \sim c\tau$ during the period when the universe transitioned from radiation to matter domination.

    - $\delta_R(\mathbf{r}) = \delta \otimes W_R = \int d^3r' \delta(\mathbf{r} - \mathbf{r}') W_R(\mathbf{r}')$
  - so in Fourier space
    - $\tilde{\delta}_R(\mathbf{k}) = \tilde{\delta}(\mathbf{k}) \tilde{W}_R(\mathbf{k})$
  - so the mean squared smoothed field is, from Parseval's theorem,
    - $\boxed{\langle \delta_R^2 \rangle = \int \frac{d^3k}{(2\pi)^3} P_\delta(\mathbf{k}) \tilde{W}_R^2(\mathbf{k})}$
  - it turns out, that the RMS fluctuation is unity when averaged within a spherical window with radius of $R = 8\text{Mpc}/h$
    - this is a very common and useful way to express the overall amplitude of the power spectum
    - it is usually referred to as σ_8

- this can be generalised to compute other quantities of interest
- this is facilitated by the fact that, in the Fourier domain, relation between the gravitational potential ϕ and the density (i.e. Poisson's equation: $\nabla^2 \phi = 4\pi G \rho$) becomes algebraic:
  - $\tilde{\phi}(\mathbf{k}) = -4\pi G \tilde{\rho}(\mathbf{k})/k^2$, so
  - $\tilde{\phi}(\mathbf{k}) = -4\pi G \overline{\rho} \tilde{\delta}(\mathbf{k})/k^2$
- as does therefore the relation between the gravity vector $\mathbf{g}(\mathbf{r}) = -\boldsymbol{\nabla}\phi(\mathbf{r})$ and the density:
  - $\tilde{\mathbf{g}}(\mathbf{k}) = -4\pi G \overline{\rho} \tilde{\delta}(\mathbf{k}) i\mathbf{k}/k^2$
- and this allows one to compute, for example, the mean squared gravity averaged with some smoothing window – this reveals that the large scale structure, while of small amplitude, drives interestingly large 'streaming flows' or 'peculiar velocities'
  - this can be compared to – and is in general accord with – the fact that the local group has a net velocity of about 600 km/s with respect to the frame defined by the microwave background
  - this velocity is generated by the gravitational pull of super-cluster scale ($\sim 30 - 100$Mpc) structures in the local universe
- a related application is measurement of the anisotropy of the clustering pattern caused by these peculiar velocities (see below)

2.2.4 Higher order n-point functions

- the initial density perturbations generated in inflation are believed to take the form of a statistically and homogeneous *'Gaussian random field'*
  - such a field can be realised as a sum of modes (Fourier synthesis) with amplitudes given by the square root of the power spectrum and with random phases
  - as such the field is completely described at a statistical level by $P(k)$ (or $\xi(r)$)
- the actual structures on small scales are, however, highly *non-Gaussian*
- to describe this, the 2-point function can be generalized to so-called 'higher order' statistics such as the 3-point function $\zeta(\mathbf{r}, \mathbf{r}', \mathbf{r}'') = \zeta(\mathbf{r} - \mathbf{r}', \mathbf{r}' - \mathbf{r}'')$, 4-point function etc...
  - crudely speaking, while the two-point function describes the variance of the density field, the higher order statistics describe things like the skewness of the distribution of densities
  - 'higher order' statistics become important once structures become 'non-linear' ($\Delta\rho/\rho \gtrsim 1$)
  - early measurements showed that the 3-point function, for instance, which is a function of two separations $\mathbf{r} - \mathbf{r}'$ and $\mathbf{r}' - \mathbf{r}''$ can be modelled as the product of 2-point functions
    - this ruled out, for instance, the idea that the structure we see is simply the superposition of randomly placed clusters with some universal profile
  - another useful early application of the 3-point function was the *'cosmic virial theorem'* (Davis and Peebles, 1983) which is, in essence, the Jeans equation, and which relates the observed two-point function of the velocity (the velocity dispersion tensor) to an integral over the 3-point function and provides a way to measure the density parameter
- considerable effort is currently being directed towards measurement of higher-order statistics on large scales
  - the goal is to detect *'primordial non-Gaussianity'* that is predicted to be generated at interesting levels in some inflationary models
  - one challenge in this is to distinguish the relics of early universe physics from non-Gaussianity that develops as the structure starts to become non-linear
- like the 2-point function these statistics can also be measured 'in Fourier space'
  - e.g. the *power spectrum* $P(k)$ (square of the Fourier transform) is the Fourier transform of $\xi(r)$
  - spatial 3-point function $\quad\leftrightarrow\quad$ the *bispectrum*

2.2.5 Biased galaxy clustering

- Inflationary models make quantitative predictions for $P(\mathbf{k})$ - massive redshift surveys can test these theories
  - however, the theories predict the clustering of *mass*
  - while galaxies are thought to be *biased tracers* of the mass
  - usually modelled as $n_{\rm gal}(\mathbf{r}) = \overline{n}(1 + b\delta\rho/\rho)$ – expected to be a good model for large-scales where $\delta\rho/\rho \ll 1$
  - but not on small scales – where there is much information, in principle, to be extracted
  - and where we know empirically – from the 'morphological segregation' seen between early- and late-type galaxies – that the clustering bias is rather complex

2.3 Peculiar velocities and redshift space distortion

- redshift is not exactly proportional to distance (as in FRW models)
  - because of structure there are *peculiar motions*
  - similar to problems mapping the disk of the galaxy with HI maps
- in dense clusters, groups there are 'incoherent' (or 'random') motions
  - gives rise to the 'fingers of god'
  - the elongation of virialised systems like clusters along the line of sight directions
- outside of clusters, groups there are 'coherent' streaming motions
  - associated with the growth of structure
  - e.g. a spherical over-dense region will be expanding slower than the universe as a whole
  - and will appear *flattened* along the line of sight
  - this distorts the 2-point function $\xi(r) \rightarrow \xi(\mathbf{r})$
- *'redshift-space distortions'* (RSD) provide a way to measure the growth rate of cosmic structure
  - simplest to analyse on large scales where random motions are negligible
  - we can consider structure to be a sum of Fourier modes where particles have been displaced parallel to the wave-vector
  - for a mode with wave-vector along the line of sight there is an additional displacement in redshift-space
  - simplest in the Einstein-de Sitter model ($\Omega_{\rm m} = 1$) where the amplitude of such a mode is doubled
  - and a general mode with wave vector at an angle θ with respect to the line of sight where the amplitude is multiplied by a factor $1 + \mu(\mathbf{k})^2$ where $\mu = \mathbf{k} \cdot \hat{\mathbf{z}} = \cos\theta$
  - the power spectrum in redshift space (subscript s) is then distorted with respect to the (spherically symmetric) real-space power spectrum (subscript r):
  - $P_{\rm s}(\mathbf{k}) = (1 + \mu^2(\mathbf{k}))^2 P_{\rm r}(k)$
  - in more realistic models the effect is less pronounced
  - $\boxed{P_{\rm s}(\mathbf{k}) = (1 + \beta^2\mu^2(\mathbf{k}))^2 P_{\rm r}(k)}$
  - where β is the growth rate of structure which is smaller than in the Einstein-de Sitter model
  - this is a *'quadrupolar'* distortion of the power spectrum and a similar distortion appears in the autocorrelation function (see figure 13)
- so this provides another useful cosmological test
  - sensitive to the density parameter $\Omega_{\rm m}$ and also the acceleration of the background universe
- there is a rather different, but also useful, way to use peculiar velocities in cosmology
- that is to use distance estimates (such as Tully-Fisher and $D_n - \sigma$) together with redshifts to map the deviations from Hubble flow – sometimes called 'cosmic flows'
- as we discussed, that led to the discovery of the 'great attractor'
- but it is limited in range as the errors in the velocities are proportional to distance, so it becomes imprecise beyond the local universe
- the RSD effect, on the other hand, does not rely on any distance estimate, and so can be measured in deep surveys

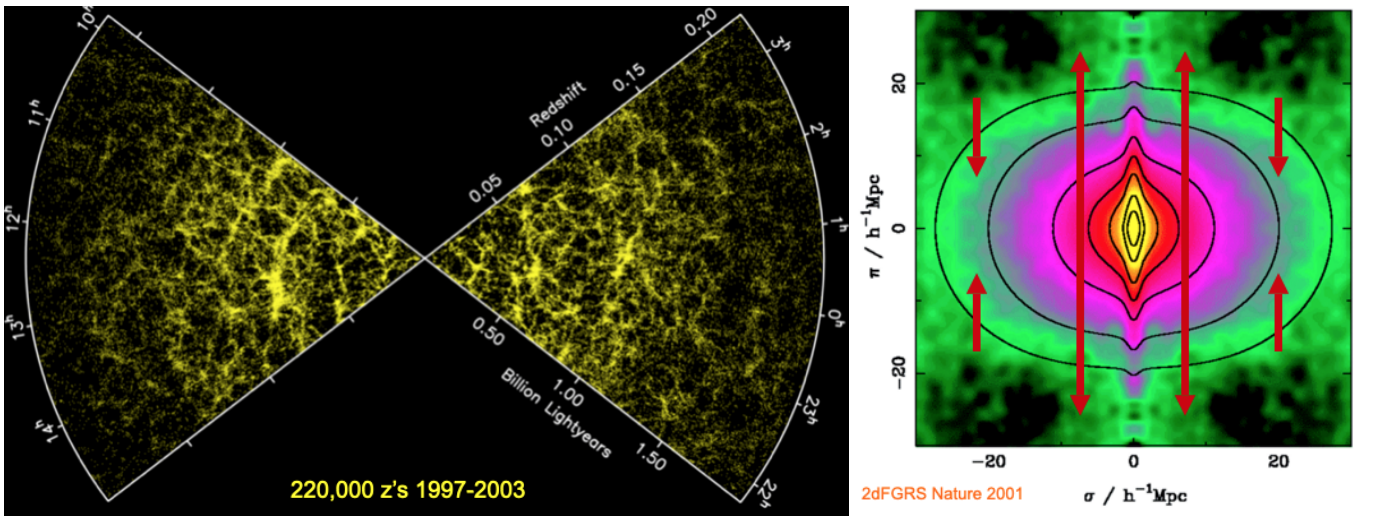

Figure 13: Measurement of the redshift space distortion effect. Left panel is the 'wedge-diagram' showing the 3-dimensional distribution of galaxies measured in the 2df galaxy redshift survey. The 3rd dimension here is redshift, which is not a perfect measurement of distance because of peculiar velocities. These cause dense clusters to appear elongated along the line of sight – the so-called *'fingers of god'* – and they cause super-clusters to appear flattened along the line of sight. The latter is a powerful test of cosmology as it measures the rate at which structures are growing, which in turn is sensitive to the density of the universe and the presence of dark energy. The right panel shows the 2-point function which shows these effects on the 2-point function measured in 2-dimensions (here π is the line-of-sight separation that is affected by velocities and r_p is the separation in the 'plane of the sky'.

2.4 Baryon Acoustic Oscillations

We considered earlier the apparent distances in cosmology; the angular diameter distance $D_{\rm a}(z)$ and the luminosity distance $D_{\rm L}(z)$. These are functions of redshift, but depend on the cosmological parameters (such as $\Omega_{\rm m}$ and Ω_Λ).

It turns out there is a rather accurate 'standard ruler' known as the *'baryon acoustic oscillation'* (or BAO) scale, which is a feature imprinted in the spatial distribution of galaxies and which enables a measure of $D_{\rm a}(z)$ and hence a powerful test of cosmology (see figure 15). This reinforces the evidence for dark energy and helps determine the cosmological parameters.

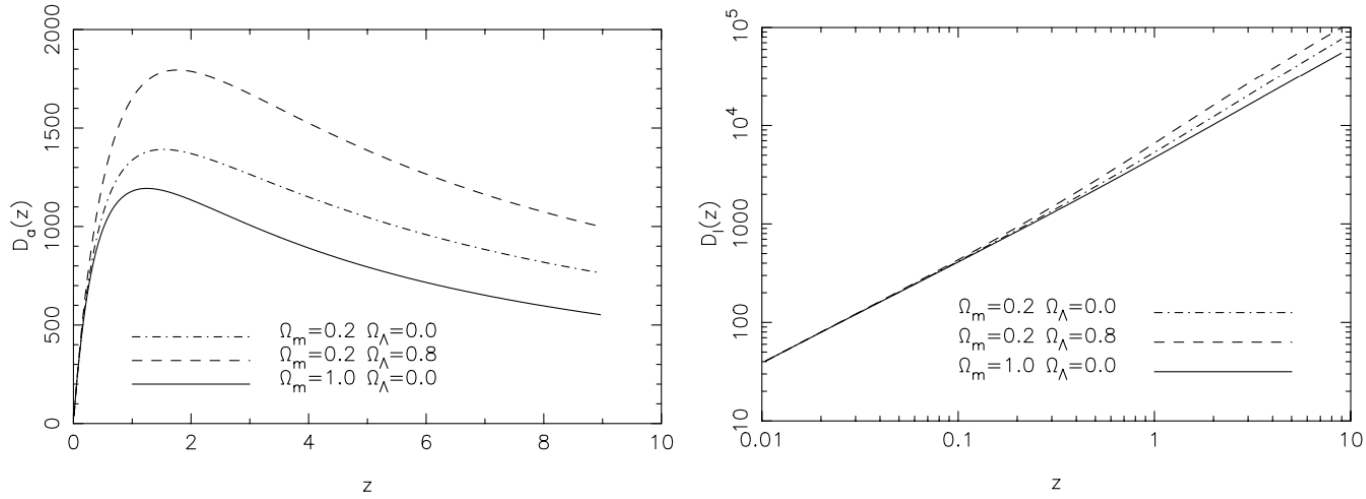

Figure 14: Angular diameter distance (left) and luminosity distance (right) as a function of redshift for various cosmological models.

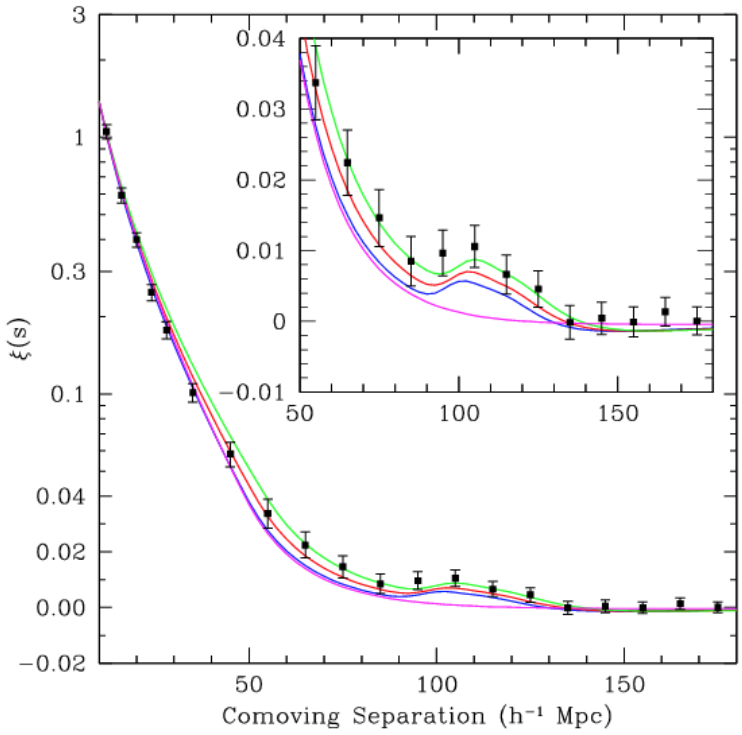

Figure 15: Baryon acoustic oscillations. In the conventional model for the origin of cosmological structure, the 'seeds' of structure were laid down during inflation in the very early universe, with a nearly scale-invariant spectrum. During the period immediately before the plasma reionized these triggered sound waves in the plasma-radiation fluid. This imprinted a feature in the spectrum of density fluctuations that emerged and later developed into fluctuations in the observed large-scale structure traced by galaxies. The feature is rather weak, but it is highly valuable as its scale is predicted from CMB observations (it is essentially the 'sound-horizon'; the product of the sound speed and the age of the universe at the decoupling epoch). Massive redshift surveys were able to measure this feature by means of the 'two-point' function characterising the galaxy distribution at left. This provides a powerful constraint on the cosmological model.

3 Cosmic flows as a probe of large-scale structure

The previous sections described the homogeneous and isotropic expanding FLRW models that are used to describe the properties of the universe on large scales and to interpret cosmological data.

Interest in these models was triggered by the discovery that the universe is expanding; that nearly all other galaxies have positive redshifts and, as Hubble was the first to show, these increase with distance.

However, the Hubble law in not a full description of the 'cosmic flow field'. On small scales, and in our immediate vicinity, the Hubble flow is expected to be quite strongly distorted by the flows associated with the fact that the structures we see – clusters and superclusters – grew from a smooth initial state.

One of the early attempts to measure this was performed by Alan Sandage (who worked with Hubble as a graduate student) and coworkers, who looked for the 'infall' to the local supercluster. They found an effect, but it was very weak; indicating a low mass density.

Around the same time, Vera Rubin and colleagues wer trying to do the same thing, but on a considerably larger scale. What she found was highly surprising as it indicated that the matter density of the universe was surprisingly inhomogeneous on large scales.

Below I will review the ways people have measured distances to galaxies and mapped out the 'cosmic-flows'; departures from a perfectly linear Hubble law.

3.1 Measuring distances to galaxies

- Astronomers can measure distances to stars in the galaxy by measuring their *'parallaxes'*
  - the reflex of the Earth's motion around the Sun
- but this does not work for galaxies as they are too far away
- extragalactic distances can be measured with *'standard candles'*
  - objects whose luminosity is known, or can be determined from some distance independent observable
  - examples are variable stars such as Cepheids and RR-Lyrae where there is a tight relation between period and luminosity
  - another effective standard candle is the luminosity of the stars in the tip of the red-giant branch (TRGB)

  - and type 1a supernovae

- variable and TRGB stars are only useful in the local universe
- supernovae can be seen to great distances, but are limited in number
- there are, however, other ways to extend the so-called *'distance ladder"* to larger scales using the properties of galaxies themselves, as we now explore

3.1.1 The 'knee' in the luminosity function

- The galaxy luminosity function (number of galaxies per unit volume per unit luminosity) is observed to have a power-law behaviour at low luminosities – with a slope that depends on the morphological type of galaxy – multiplied by. an exponential factor $\exp(-L/L_*)$ that quite sharply truncates the distribution for $L > L_*$, with the characteristic luminosity L_* being described as the location of the 'knee' in the luminosity function. See figure 16.
- If one compares measurements of the luminosity function for two clusters, say, then the knee of the LF will appear at a fainter apparent luminosity – or flux density – for the more distant cluster.
- This gives an estimate of distance, though a rather noisy one
  - the uncertainty in log-distance for a single galaxy being of order unity

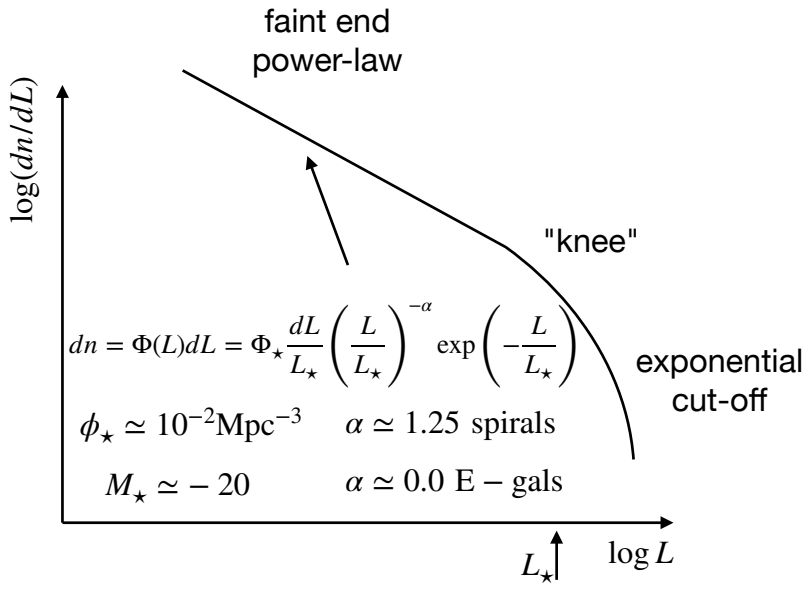

Figure 16: Galaxy (differential) *luminosity function:*. Massive *redshift surveys* (like SDSS) have been performed to determine the 3D distribution of galaxies. The counts of galaxies as a function of luminosity are well described by what's called a *'Schechter function'* after pioneering work by Paul Schechter. It is a power law at low luminosities – with a formal divergence of the number of galaxies – and with a sharp exponential cut-off above what's called the *'characteristic luminosity'* $L_\star$. The characteristic absolute magnitude is $M_\star \simeq -20$, which is about the absolute magnitude of the Milky Way. Elliptical galaxies have a flatter 'faint-end'; with $\alpha \simeq 0.0$.

3.1.2 Applications: the infall to the Virgo supercluster and the Rubin-Ford effect

In the early 70s, Sandage and co-workers tried to use this to measure the expected infall of the MW into the 'Local-Supercluster' (with the Virgo cluster at its centre). They looked at galaxies in a shell around us in redshift space (with recession velocities of about 1000 km/s, like Virgo) and looked for a 'dipole' in the luminosities; the expectation was that galaxies in Virgo would be slightly fainter. They found only a very small effect.

But then Vera Rubin and co-workers applied the same test at a much larger distance (recession velocities of between 3500 and 6500 km/s) and found a surprisingly large effect. This became known as the 'Rubin-Ford effect' and is illustrated in figure 17 It showed that not only are we – or rather the Local Group – moving at a large velocity of about 500 km/s with respect to this shell of galaxies, but that this velocity does *does not* agree with our motion with respect to the cosmic microwave background. So it must be that this very large shell of galaxies has a large net velocity with respect to the 'cosmic reference frame' defined by the CMB. This was wholly unexpected and indicates that there are very large-scale 'peculiar motions' being driven by large-scale inhomogeneity of the matter distribution.

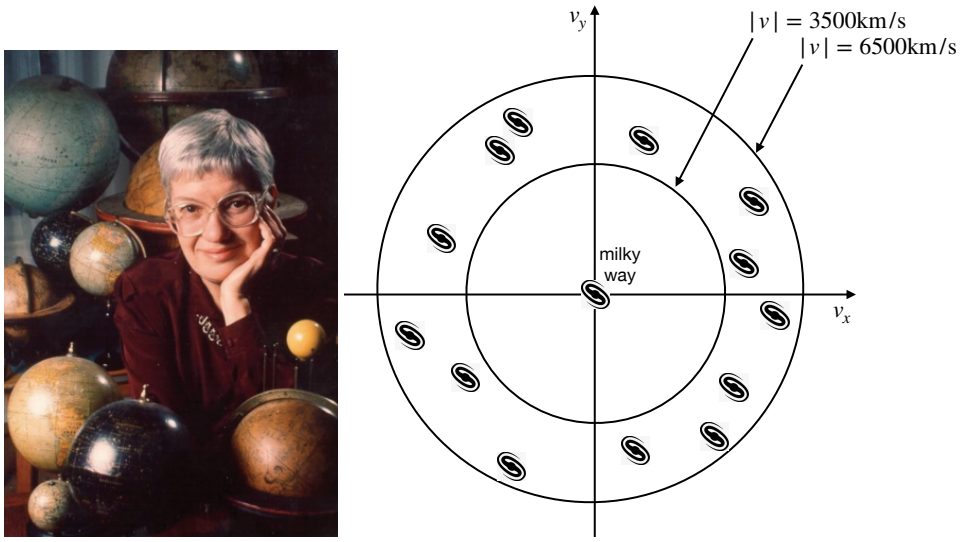

Figure 17: The Rubin-Ford effect. Vera Rubin at left, who did pioneering work establishing that spiral galaxies have flat rotation curves, and after whom the Large Synoptic Survey Telescope has been named, measured luminosities for 96 spiral galaxies with recession velocities indicated and used these to estimate their distance. She found a dipolar asymmetry whose interpretation is that we are moving at about 500 km/s with respect to the frame of reference defined by these galaxies.

3.2 Distance measurement using internal 'scaling relations' of galaxies

3.2.1 The Tully-Fisher relation

- In the mid '70s Brent Tully realized that a much more precise distance estimate could be made for spiral galaxies using the rotation velocity measured from HI spectra (corrected for inclination) as a way to determine the intrinsic luminosity
  - intrinsically more luminous galaxies having greater rotation velocities
  - empircally it was found that $L \propto v^\gamma$ with $\gamma \simeq 3-4$
  - this is known as the *Tully-Fisher relation*
  - this works best if the luminosity is measured in the near infra-red as this is less affected by dust obscuration and because it measures the luminosity of the older stars and is so less affected by 'noise' coming from the recent star formation history
  - the 'IR-TF relation' gives distances with a fractional uncertainty of about 20% (per galaxy; errors go down like $1/\sqrt{N}$ if we average over multiple galaxies)

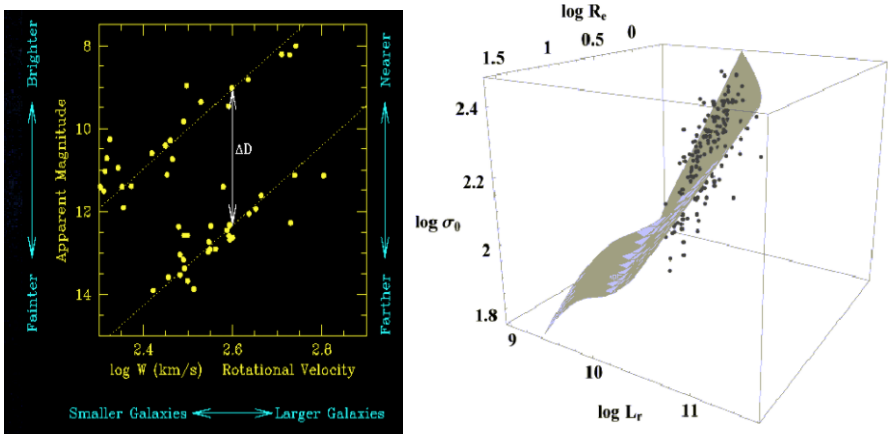

Figure 18: The Tully-Fisher relation (for spirals) and the fundamental plane (for elliptical galaxies) are empirical relations between internal properties of galaxies that enable one to predict the luminosity (or size in the case of ellipticals) from distance independent quantities. As apparent size and flux density are distance dependent (given intrinsic size or luminosity) these allow one to measure distances to galaxies.

3.2.2 The fundamental plane for elliptical galazies

- A similar relation – known as the 'Faber-Jackson relation' – was found to exist for 'early-type' (i.e. elliptical) galaxies
  - for which $L \propto \sigma_v^4$

  - but this gave somewhat less precise distances

- The form of the F-J relation is a reflection of the fact that E-gals have roughly constant surface brightness and mass-to-light ratio
  - the virial theorem says $\sigma_v^2 \sim GM/R$
  - while the surface brightness is $\Sigma \sim L/R^2$
  - so $L \sim \Sigma R^2 \propto \Sigma \times (M/\sigma_v^2)^2 \quad \Rightarrow L \propto \sigma_v^4$

- now elliptical galaxies do not, in fact, have constant surface brightnesses
  - but the surface brightness is measurable
  - as it is independent of distance
  - so we can correct for this

- in the mid-'80s the '7-samurai' (as well as others such as George Djorgovski and Marc Davis) realised that if one combined measurements of the log-velocity dispersion and the log-surface brightness
  - both being independent of distance but both being dependent on the luminosity of the galaxy

- one could obtain distances with precision comparable to that of TF measurements for spirals

- this was exploiting the fact that elliptical galaxies populate a rather well defined plane in the 3-dimensional space of (logarithms of) velocity dispersion, surface brightness and luminosity (or size) known as the 'fundamental plane'.
  - it is also known as the $D_n - \sigma$ relation.

3.3 The Hubble expansion rate and the 'great attractor'

Such distances can be determined out to quite large distances of order 100's of Mpc.

As such they provide a very useful 'rung' at large distances in the 'distance latter' (in which one uses the period-luminosity relation for variable stars to provide a link from distances measured by parallax in the MW to distances to external galaxies). That 'calibrates' the TF and fundamental plane distances (and also extragalactic distances from supernovae).

Comparing the distances and recession velocities tells us the expansion rate H of the universe (from $\mathbf{v} = H\mathbf{r}$).

Another important discovery that came from such measurements – in particular those of the 7-samurai – was that there was an indication of tidal stretching of the region around us out to about 30 Mpc that indicated the presence of a large over-density – the so-called 'great attractor' – at still larger distances. This came as something of a surprise since while it was known that the MW lies in the 'local supercluster', the tidal field required a larger – possibly much larger – mass concentration.

This took a while to sort out, in part because what is now known as the 'great attractor' is partly covered by the 'zone of avoidance' hidden behind the disk of the MW (see figure 19), and in part because the effect is not coming from a single attractor, but is coming from the combination of a number of super-clusters, with a substantial contribution coming, again rather surprisingly, from the 'Shapley supercluster' which is about 200 Mpc away.

4 Gravitational Lensing

4.1 Introduction

4.1.1 History of gravitational lensing

- 1706: Newton, Opticks, Book 3:
  - Query 1. Do not Bodies act upon Light at a distance, and by their action bend its Rays, and is not this action (cæteris paribus) strongest at the least distance?
  - Query 29. Are not the Rays of Light very small Bodies emitted from shining Substances?

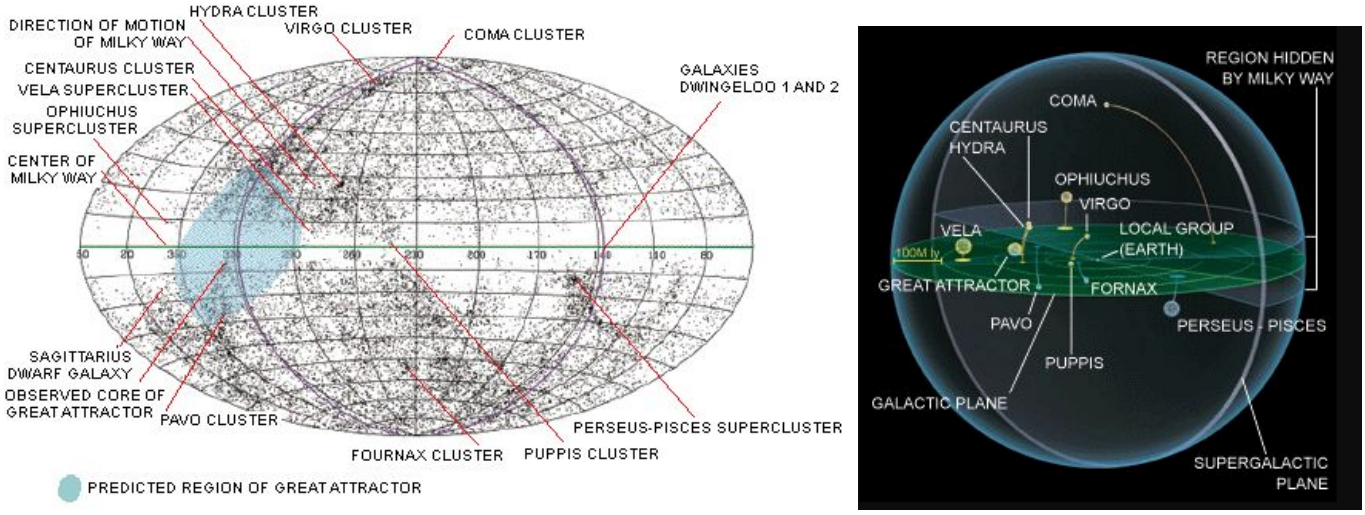

Figure 19: The 'Great Attractor'. The '7-samurai' – a group of astronomers measuring kinematics of stars in elliptical galaxies discovered the $D_n - \sigma$ relation; the observed (and distance independent) velocity dispersion correlated tightly with size of the galaxies (with a correction for surface brightness). The galaxies lie on what is called the 'fundamental plane'. This allowed them to estimate distances to these galaxies and to measure the expansion rate of the universe H in the Hubble relation $v = Hr$. What they found was that the expansion was different in different directions. Fitting for a tensor relation: $\mathbf{v} = \mathbf{H} \cdot \mathbf{r}$ gave a much better fit and showed that there is a tidal stretching of the local universe. Further study pinned down the source of this to an overdensity of galaxies – part of the already known 'Centaurus great wall' supercluster, but hidden by the plane of the Milky Way.

so he would seem to have invented here the concepts used in gravitational lensing – at least in the limit of geometric optics (note that Römer had measured the speed of light ~ 20 years previously)

- 1802: Henry Cavendish: used Newtonian theory to calculate the deflection of light by the Sun. Also done by Baron von Soldner.
  - $\Delta\theta = 2GM/bc^2 \simeq 0.84$"
  - half the correct value
- $\simeq$ 1910: Einstein: Uses the principle of equivalence and the 'rocket' thought experiment
  - gets the same result as Newtonian theory – see figure 20
  - 1912 - 1914: two attempts to measure it – failed (to prove him wrong)
- 1915: Einstein: GR: correctly predicts $\Delta\theta = 4GM/bc^2 \simeq 1.68$"
  - can be obtained from the geodesic equation for a massless particle in the weak-field metric:
  - $ds^2 = -(1+2\Phi)c^2dt^2 + (1-2\Phi)|d\mathbf{x}|^2$
  - extra bending coming from curvature of space
  - as compared to massive particles which see only the warping of time
- 1919: Eddington eclipse expedition: confirms Einstein's prediction
  - Einstein becomes a celebrity
- Einstein considers lensing by other stars but concludes it would be hard to measure
- 1937: Fritz Zwicky: realises that galaxies could be much more effective
  - $\delta\theta \sim \sigma_v^2/c^2 \sim 1$"
  - and clusters of galaxies even more so $\sim 10 - 30$" bending
  - 'giant arcs' and 'Einstein rings' predicted

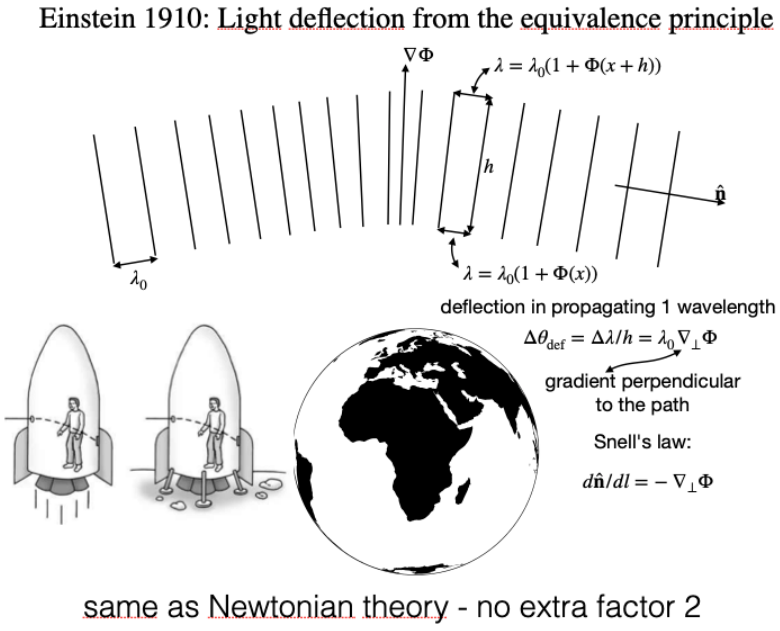

Figure 20: Around 1910, Einstein came up with two arguments from which one can infer the bending of light by a massive object. One was based on his 'tower' thought experiment which convinced him that there should be a gravitational blue-shift for light falling into a potential. The resulting reduction in the wavelength leads to deflection of a beam of radiation passing a massive body. The other argument was based on the equivalence principle, in the form that what we see if we are sitting on Earth – being impeded from falling to its centre by the acceleration provided by the ground under our feet – is the same as if we were in an accelerating rocket in empty space. Both of these lead to predictions identical to that of Newtonian gravity.

4.1.2 The 'lumpy glass' analogy

- weak-field metric
- $ds^2 = -(1+2\Phi)c^2dt^2 + (1-2\Phi)|d\mathbf{x}|^2$
- so photon trajectories ($ds = 0$) have *coordinate speed of light*
  - $|d\mathbf{x}/dt| = \sqrt{\frac{1+2\Phi}{1-2\Phi}}c \simeq (1+2\Phi)c$
  - light moves slower in a potential well
- just light for light in an inhomogeneous medium with refractive index
- $n(\mathbf{x}) = c/|d\mathbf{x}/dt| \simeq 1 - 2\Phi$
- with which we can apply *Snell's law* (or *Snellius-Descartes* – or *Ibn Sahl* 984)
- $d\hat{\mathbf{n}}/d\lambda = -2\boldsymbol{\nabla}_\perp\Phi$
- this analogy is not perfect, however:
  - in lumpy medium, the speed of light (as measured locally) varies but the frequency is fixed
  - in gravity the frequency measured locally changes, but the speed is fixed
    - it is the "coordinate velocity" that varies
  - also, this fails when applied to e.g. measurement of light bending 'in the lab'
    - not that anyone has succeeded in doing this
    - but if they did they would confirm Einstein's 1910 prediction

4.2 Some applications of gravitational lensing in cosmology

4.2.1 The Einstein radius and the critical surface density

- The geometry for calculating the Einstein radius $\theta_{\rm E}$ is shown in figure 21
- we assume that it is very small, so the bending all takes place in a very small range of distances
  - that justifies drawing the light path as two straight lines
- the bending angle $\delta\theta$ for a Newtonian particle can be calculated easily by integrating the transverse acceleration along the line of sight to get $\Delta v_\perp = \int (dz/c)a_\perp$ and then forming

- $\delta\theta = \Delta v_\perp / v = 2GM/bv^2$

- the bending angle for light in GR is twice the Newtonian angle:
  - $\boxed{\delta\theta = 4GM/bc^2}$

- assuming flat spatial geometry (it's not difficult to generalise to hyperbolically curved space) the impact parameter is $b = a_\mathrm{L}\chi_\mathrm{L}\theta_\mathrm{E}$ (where a_L is the scale factor at the time the light passes the lens) which we can use to relate $\delta\theta$ to θ_E:
  - $\delta\theta = 4GM/a_\mathrm{L}\chi_\mathrm{L}\theta_\mathrm{E}c^2$

- while the fact that the physical length of the vertical arrow at the right can be expressed as either $a_\mathrm{S}\chi_\mathrm{LS}\delta\theta$ or as $a_\mathrm{S}\chi_\mathrm{S}\theta_\mathrm{E}$ gives us a second relation between $\delta\theta$ and θ_E:
  - $\delta\theta = \theta_\mathrm{E}\chi_\mathrm{S}/\chi_\mathrm{LS}$

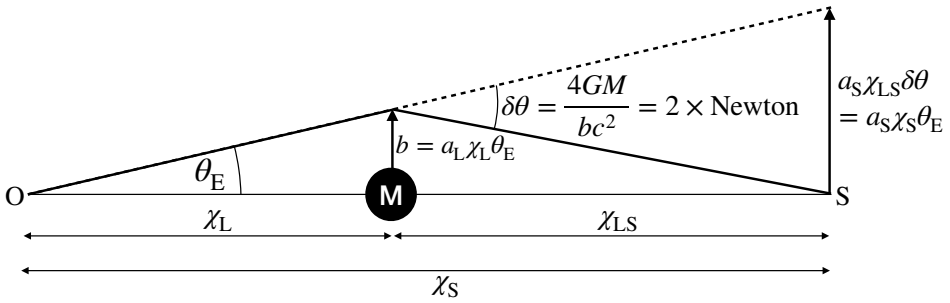

Figure 21: The Einstein radius is the angular radius of the ring that is seen by the observer O for a point-like source S that is exactly on-axis behind a point mass lens.

- equating these gives the *lens equation*
  - $\boxed{\frac{4GM}{a_\mathrm{L}\chi_\mathrm{L}\theta_\mathrm{E}c^2} = \frac{\chi_\mathrm{S}}{\chi_\mathrm{LS}}\theta_\mathrm{E}}$

- whose solution is the *Einstein radius*
  - $\boxed{\theta_\mathrm{E} = \sqrt{\frac{4GM\chi_\mathrm{LS}}{a_\mathrm{L}c^2\chi_\mathrm{L}\chi_\mathrm{S}}}}$

- which, in terms of distances $D(z_\mathrm{L}) \equiv a_0\chi_\mathrm{L}$, $D(z_\mathrm{S}) \equiv a_0\chi_\mathrm{S}$, and using $a_\mathrm{L} = a_0/(1+z_\mathrm{L})$ is
  - $\boxed{\theta_\mathrm{E}(M, z_\mathrm{L}, z_\mathrm{S}) = \sqrt{\frac{4GM(1+z_\mathrm{L})}{c^2}\frac{D(z_\mathrm{S}) - D(z_\mathrm{L})}{D(z_\mathrm{L})D(z_\mathrm{S})}}}$
  - where one should note that the distances are neither angular diameter distances $D_\mathrm{a}(z) \equiv a(z)\chi(z) = (1+z)^{-1}\int cdz/H$, nor luminosity distances $D_L(z) \equiv (1+z)^2 D_a(z) = (1+z)\int cdz/H$, but conformal distances $D(z) \equiv \int cdz/H$, which are the geometric mean of the other distances: $D(z) = \sqrt{D_\mathrm{a}(z)D_L(z)}$

- if we are dealing with a point mass lens, if we know the Einstein radius and the redshifts of the lens and the source then we can invert this to get the mass M

- for extended lenses like clusters then, assuming spherical symmetry, this would give us the projected mass within the cylinder defined by the light rays

- if we divide the mass by the proper area πb^2, using $b = a_\mathrm{L}\chi_\mathrm{L}\theta_\mathrm{E} = D(z_\mathrm{L})\theta_\mathrm{E}/(1+z_\mathrm{L})$ we get the so-called *critical surface density*:
  - $\boxed{\Sigma_\mathrm{crit}(z_\mathrm{L}, z_\mathrm{S}) = \frac{M}{\pi b^2} = \frac{c^2(1+z_\mathrm{L})}{4\pi G}\frac{D(z_\mathrm{S})}{D(z_\mathrm{L})(D(z_\mathrm{S}) - D(z_\mathrm{L}))}}$

- which is the surface density of a mass-sheet at redshift $z_{\rm L}$ that is sufficient to re-focus light from a point source at redshift $z_{\rm S}$

- the critical surface density becomes very large for lenses that are very close to either the observer or the source

- for a source at $z_{\rm S} = 2$, this formula gives a minimum critical surface density of about $0.4{\rm g/cm}^2$ at $z_{\rm L} \simeq 0.5$

  - but with a broad minimum: lenses with $\Sigma > 0.6{\rm g/cm}^2$ being capable of lensing such sources if they have $1 > z_{\rm L} > 0.15$

- only in the cores of clusters does the density reach the critical value

  - so in such 'strong lensing' we are seeing the 'tip of the iceberg' of the lenses

4.2.2 Micro-lensing of stars by compact halo objects

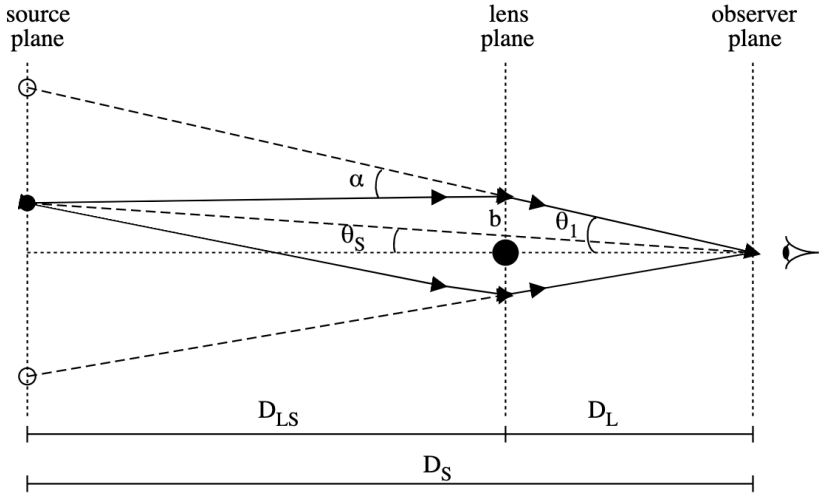

Figure 22: Geometry for lensing of a point-source by a point-mass. There are two images. If the 'impact parameter' is large – that is to say the un-lensed angle $\theta_{\rm S}$ is much larger than the Einstein radius $\theta_{\rm E}$ then the second image – seen close to the lens – will be very faint, but if $\theta_{\rm S} \sim \theta_{\rm E}$ the amplification (the sum of the flux densities for the two stars) will become large; very large for $\theta_{\rm S} \ll \theta_{\rm E}$. Surveys like MACHO were, and continue to be, carried out to monitor the flux-densities of millions of background stars in the bulge and in neighbouring galaxies to search for MACHOS in the MW halo.

- An important application of gravitational lensing in cosmology is micro-lensing of stars

  - the geometry for lensing by a point mass is shown in figure 22
  - the total flux density amplification is
    - $\boxed{M = (u^2+2)/u\sqrt{4+u^2}}$
    - where $u \equiv \theta_{\rm S}/\theta_{\rm E}$
  - so as a source moves behind a foreground point-mass lens there will be a bump in the light curve
  - Bohdan Paczyński realised that this provides a way to constrain the contribution of black-holes, or other stellar remnants (known collectively as massive compact halo objects or MACHOS) to the dark matter in the halo of our galaxy as they would occasionally amplify stars in the bulge of the MW and in other nearby galaxies and produce bumps in the light curves with a characteristic signature
    - while the optical depth for such events is small, it is possible to monitor millions of background stars
  - the results of the MACHO and OGLE projects we highly successful and gave tight constraints on the contribution of such objects to the DM
  - another interesting feature of such searches is that the tidal field of planets around the lensing stars can break the singularity of the magnification pattern – by generating astroid caustics – and this provided a way of detecting extra-solar planets

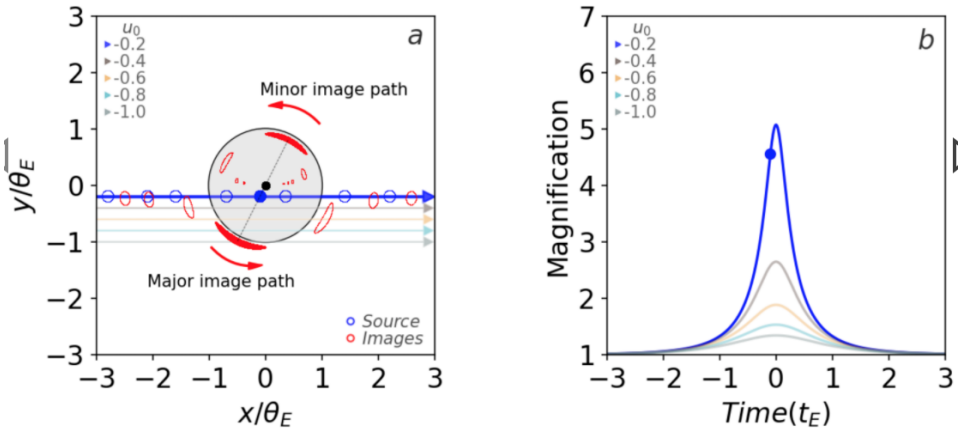

Figure 23: Microlensing light curve. Blue line at left is the path of a source behind a point mass lens. Red 'bananas' are the outline of the images. At right are light curves for various different values of the impact parameter. Images from Yiannis Tsapras.

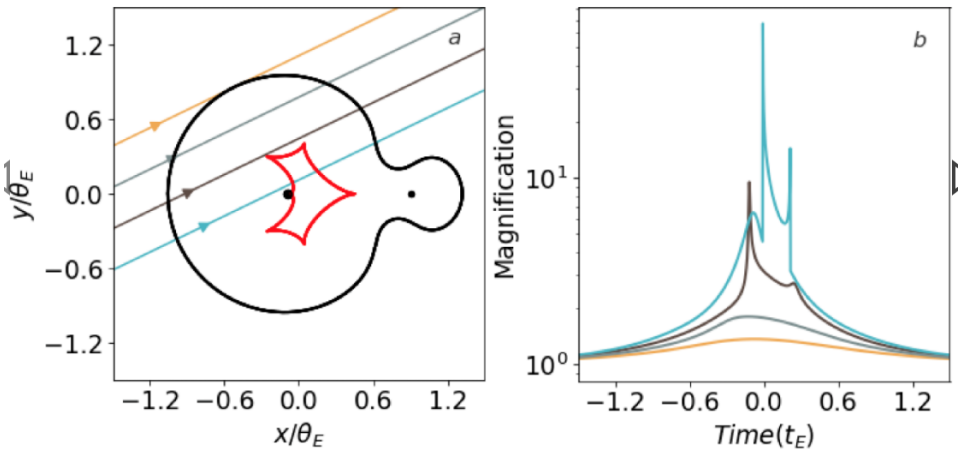

Figure 24: Microlensing light curve for a binary lens. The asymmetry of the lens has caused the line-like singularity lying along the axis behind a point mass lens to 'degenerate' and caustics (fold catastrophes) have formed. The special form of the point mass lens is very sensitive to even small extenal tidal fields and can be used to detect extra-solar planets. Images from Yiannis Tsapras.

4.2.3 Micro-lensing at cosmological distances

- Gunn and Gott pointed out that the 'optical depth' for microlensing of compact sources at cosmological distances would be on the order of $\Omega_{\rm DM}$
  - assuming the DM is composed of sufficiently compact objects
  - and the EInstein radius is greater than the Fresnel scale $r_{\rm F} \sim \sqrt{D\lambda}$.
  - and the source is sufficiently compact also
- Andy Gould pointed out that *gamma-ray bursts* could be lensed by asteroid mass lenses.
- A recent discovery was of a transient event 'Icarus' which is a supergiant star ($L \sim 10^5 L_\odot$) at $z \simeq 1.5$ being lensed (and having its flux density magnified by $\sim 10^4$) by a cluster of galaxies at $z \sim 0.5$.
  - such events are interesting as they can place constraints on the nature of the DM in clusters
  - if, for example, the DM is composed of black-holes
  - or fuzzy dark matter

4.3 Strong lensing by galaxies and galaxy clusters

- Galaxies and the central parts of clusters of galaxies have surface density exceeding the critical value and produce multiple images of distant sources (such as QSOs) or, in propitious events, giant arcs and Einstein rings
- These provide very high quality determinations of the projected mass within the Einstein radius
- As was first pointed out by Sjur Refsdal, the liight in the different images in a multiply imaged QSOs have relative **time delays**
  - simple geometry tells us that the light arrival time difference is on the order of θ^2/H

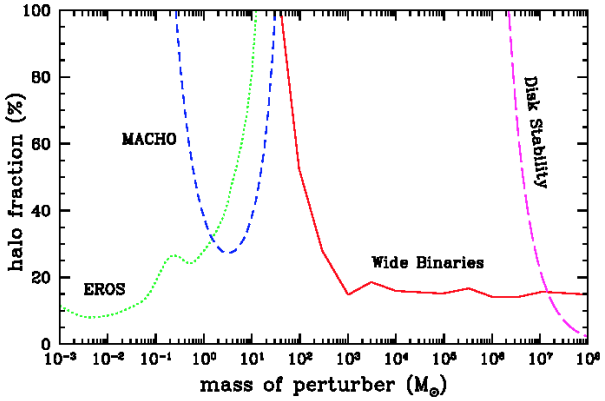

Figure 25: Limits on the fraction of dark matter in Massive Astrophysical Compact Halo Objects (MACHOS) from Yoo et al.. MACHO and EROS are two microlensing surveys. They ruled out the indicated mass ranges. The existence of wide, loosely bound, binary stars in the Milky Way have been argued to give the limits shown in red.

  - so ranging from days to years
  - and QSOs, conveniently, have strong time variation
  - so by cross correlating the 'light curves' one can determine the time delay
  - and thus determine the Hubble constant (without recourse to any complex 'distance ladder')
  - this has come to the fore recently as there appears to be a possible 'tension' between H_0 measured locally and that inferred from the CMB (in the ΛCDM model).
  - but it involved complex modelling

4.4 Gravitational lensing by galaxy clusters

4.4.1 Cluster masses from giant arcs

- Some massive cluster display *'giant arcs'*
- these are highly distorted and elongated images of distant background galaxies
- according to GR light deflection is qualitatively like that for Newtonian test particles but a factor 2 larger
  - galaxy motions are controlled by the *'curvature of time'*
  - which is to say that it is only the perturbation to the time-time component of the metric g_{tt} that affects their motions
  - relativistic particles (like photons), on the other hand, are affected by curvature of space *and* time (equally in GR)
- gravitating systems behave optically like a refractive medium with refractive index $n(\mathbf{r}) = 1 - 2\phi(\mathbf{r})/c^2$
- giant arcs are 'strongly lensed' and probe the inner parts (the 'cores') of clusters
  - generally in agreement with masses from velocity dispersions
- can also determine the mass further out from 'weak lensing'
- potentially useful as a test of alternative theories of gravity

4.4.2 Caustics and critical curves

- clusters are not perfectly circular – far from it in many cases
- but in many cases one has a reasonably large number of strongly lensed background galaxy images
  - note that the critical density depends on source redshift
  - so higher redshift sources have larger Einstein radius and will appear further from the cluster centre

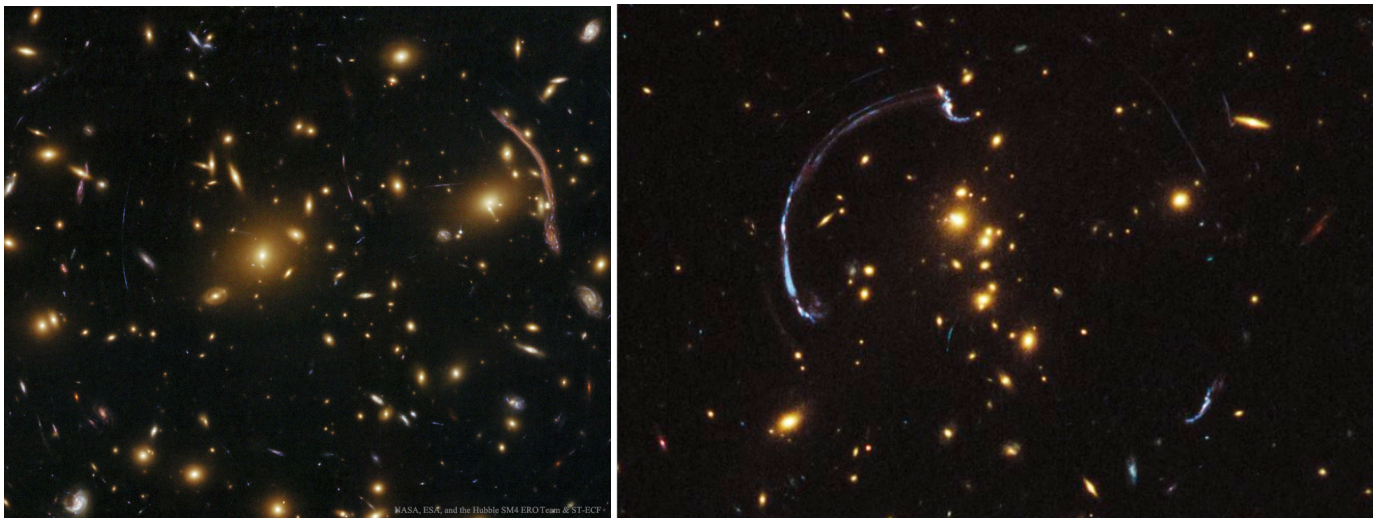

Figure 26: Two massive clusters (A370 an RCS2-032727-132623) displaying 'giant arcs'. These give quite precise measurements of the projected mass along a cylinder.

Figure 27: Two spectacular 'Einstein rings' – produced by galaxy lenses.

- so it is possible to generate quite detailed models of the projected surface mass density for lenses
- for a single point source and a given lens (assumed to have continuous surface density) the rays to the observer plane will suffer a continuous 2D deflection as a function of 2D impact parameter
- that means we can infer, from 'catastrophe theory', that there will be *'caustics'* on the observer plane
  - these are lines on the observer plane where the density of rays becomes formally infinite
    - though the density would remain finite for a finite sized source
  - these are generically one-sided, with the flux density falling off as inverse square-root of distance from the caustic on one side – these are known as 'fold-catastrophes' (there are so-called higher-order catastrophes, with names like 'cusp' and 'swallow-tail', but the most important for are purposes are the folds)
  - an observer outside the outer-most caustic will see a single image of the source
  - but if that observer crosses a caustic he or she will see a pair of infinitely bright images form, on the opposite side of the lens, where there was previously no image
  - these images have opposite 'parity' (the 'odd-parity' image being like that seen in a mirror) and rapidly move apart and dim
- if we hold the observer fixed, then there are analogous caustic surfaces on the source plane where the flux-density magnification of a source would become infinite
  - but the real caustics are on the observer-side of the lens

- there are also so-called *critical curves* on the image plane
  - these are lines where the infinitely bright image pairs can form

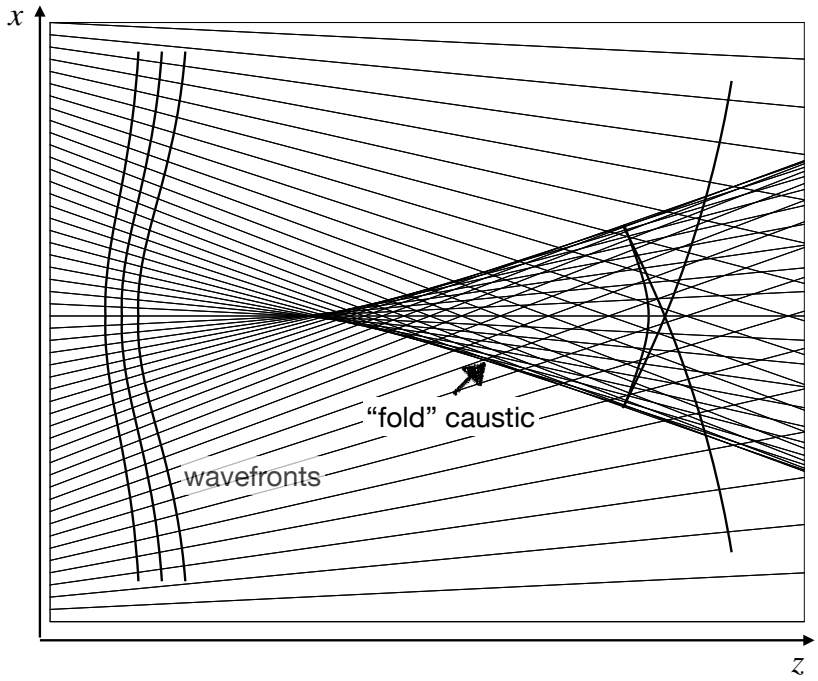

Figure 28: Simple model for a gravitational lens. Straight lines are light rays that have passed through a lens (not shown). This produced a time-delay for the light propagation and hence the wave-fronts are curved. The rays are normal to the wave-fronts and a certain distance down-stream they focus. An ideal lens would focus to a point (and a point-mass lens would generate a line-like singularity) but in general, caustics form. As an observer crossed a caustic two extra images form (so there is always an odd number of images). The extra images are initially infinitely bright, but rapidly move apart. Their flux density drops of as $1/\sqrt{d}$ – it is not difficult to see that the density of rays has this universal scaling law.

- two powerful theorems result from the universal nature of these so-called 'fold catastrophes'
  1. the cumulative probability density for very high flux density amplifications $A \gg 1$ has the universal scaling law
     - $\boxed{P(>A) \propto A^{-2}}$

     (as pointed out by Peter Schneider)
  2. there is always an odd number of images

4.4.3 The optical depth for strong lensing

- from the demographics of clusters (and galaxies) one can calculate an 'optical depth for strong lensing' defined to be the fraction of the source plane for which the amplification exceeds some threshold
  - i.e. it gives the probability that a source have flux-density above the corresponding threshold
- as mentioned, only the very centres of clusters reach the critical density threshold, so we are seeing the tip of the ice-berg
- for a threshold of a factor of a few amplification A, and for sources at $z_{\rm S} = 2$, the optical depth is $\tau \simeq 10^{-3}$
  - and it is roughly equally split between cluster lenses with image splittings of order tens of arcsec
  - and galaxy lenses with splittings of between one to a few arcsec

4.4.4 Amplification bias and quasar galaxy associations

- in surveys of sources selected according to some flux density threshold this gives rise to so-called 'amplification bias'
  - lensing amplification can push otherwise undetectable sources above the flux density threshold
  - one realisation of this is in 'quasar-galaxy associations'

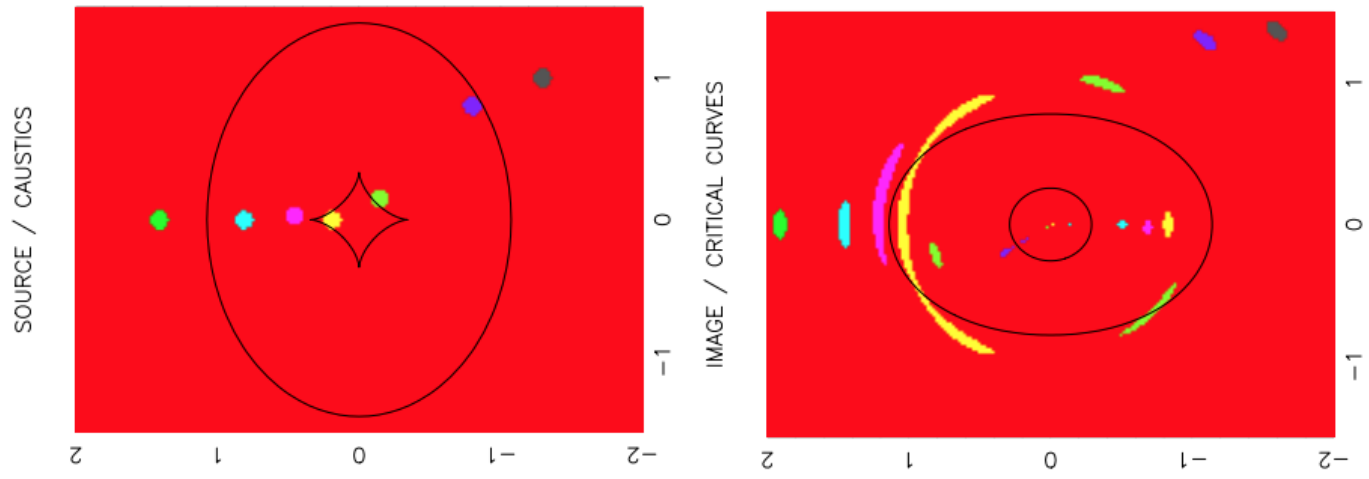

Figure 29: Caustics (left) on the source plane (along with various extended sources) and critical curves (right) on the image (or sky) plane for a slightly flattened, but otherwise spherical lens (made by Konrad Kuijken), are indicated as the black lines. Two of the green sources are outside the outermost caustic and we see only a single image for each of these. The other sources are inside the region where multiple images are seen. For the blue source we see a pair of images either side of the inner critical curve. Had the lens been circular, there would have been a singular amplification pattern with a spike at the origin in the source-plane where a source would produce an Einstein-ring at a radius close to that of the outermost critical curve. The asphericity has caused the singularity to degenerate into a so-called 'astroid' caustic. The bright yellow and green sources straddle the astroid, and produce highly magnified and brightened images.

    - at one time considered to be evidence that quasar redshifts were not cosmological
    - but later realised to be a consequence of lensing; quasars were being detected close on the sky to galaxies because they were being magnified
- it is important to bear in mind that such quasars get their flux densities amplified only by being made larger in size
    - their sizes were not being observed, but had they been, it would have been apparent that the surface brightness was not being affected by these essentially static lenses (Liouville's theorem at work)
- so the number density of lensed sources will be *diluted* at the same time
    - so if we measure the apparent luminosity function for sources behind a lens then for a 'Schechteresque' luminosity distribution one will see an excess of sources at the bright end of the flux-density distribution and a diminution at the faint end

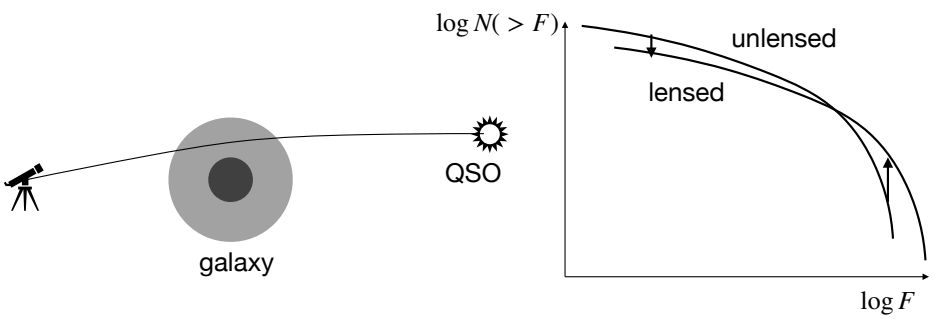

Figure 30: Amplification bias. At the left is shown a distant QSO, the light from which passes through the halo of a foreground galaxy or cluster. It will be magnified and its flux density amplified – so otherwise undetectable sources can be see close to foreground lenses. But the number density is also diminished. This results in a deficit of faint objects.

4.5 Weak Lensing

4.5.1 Introduction

An extremely promising probe of large-scale structure is *weak gravitational lensing*. Light rays propagating to us through the inhomogeneous universe get tugged from side to side by mass concentrations. This gives rise to coherent distortions of the shapes of background galaxies as illustrated in figure 31. Also shown in the figure are the plots from the pioneering paper by Jacob Zel'dovich from 1963 who was the first person – to my knowledge – to point out this effect.

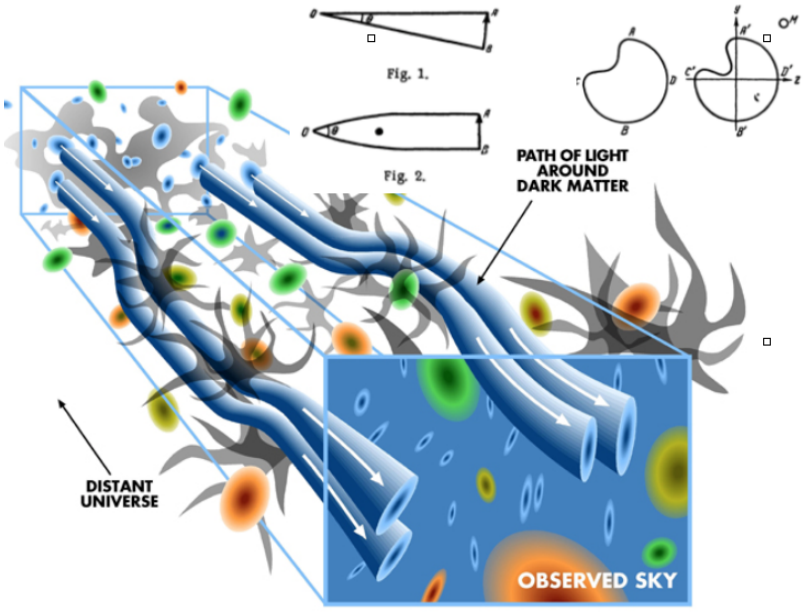

Figure 31: The blue tubes are originally circular bundles of light rays emerging from an observer back off in the distance. These get deflected and tidally distorted as they traverse the universe and end up, on the source-plane, being elliptical. This means that a circular source will appear elliptical, and the shapes of the sources will be polarised. With a single source we cannot infer the lensing induced distortion. But there are huge numbers of background galaxies — the 'cosmic-wallpaper' – and the distortion is coherent over large angular scales. By averaging the elliipticities of the galaxies in a patch we can measure the 'cosmic-shear'. The insets are from Zel'dovich's pioneering 1963 paper.

4.5.2 Order of magnitude of the WL effect

The deflection of a light ray that passes a point mass M at impact parameter b is

$$\theta_{\rm def} = \frac{4GM}{c^2 b}. \tag{5}$$

This is just twice what Newtonian theory would give for the deflection of a test particle moving at $v = c$. In this picture we are imagining the radiation to be test particles being pulled by a gravitational acceleration. There is another useful way to look at this using wave-optics; the inhomogeneity of the mass distribution causes space-time to become curved. As discussed earlier, the space in an over-dense region is positively curved, as illustrated in figure **??**. This means that light rays propagating through the over-density have to propagate a slightly greater distance than they would in the absence of a the density perturbation. Consequently the wave-fronts get retarded slightly in passing through the over-density and this results in focusing of rays. There is still another way to picture the situation: The optical properties of a lumpy universe are, in fact, essentially identical to that of a block of glass of inhomogeneous density where the refractive index is

$$n(\mathbf{r}) = (1 - 2\phi(\mathbf{r})/c^2) \tag{6}$$

with $\phi(\mathbf{r})$ the Newtonian gravitational potential. In an over-dense region, ϕ is negative, so n is slightly greater than unity. In this picture we think of space as being flat, but that the speed of light is slightly retarded in the over-dense region. All three of the above pictures give identical results.

The gravitational potential of a bound structure is on the order of $\delta\phi \sim \sigma_v^2$, where σ_v is the velocity dispersion or circular velocity. Now the velocity dispersion for a massive cluster of galaxies, for example, is $\sigma_v \simeq 1000$km/s, or about 0.003 times c. One such object can cause a deflection of on the order of $20''$, and multiple objects along the line of sight would give random deflections adding in quadrature to give still larger deflection. Images of distant objects are therefore shifted from their "true" positions (i.e. the positions they would have if we could somehow switch off the gravity perturbation). Unfortunately, this deflection is not easily measurable, since we do not know the true positions. Instead, the effect exploited in weak lensing is

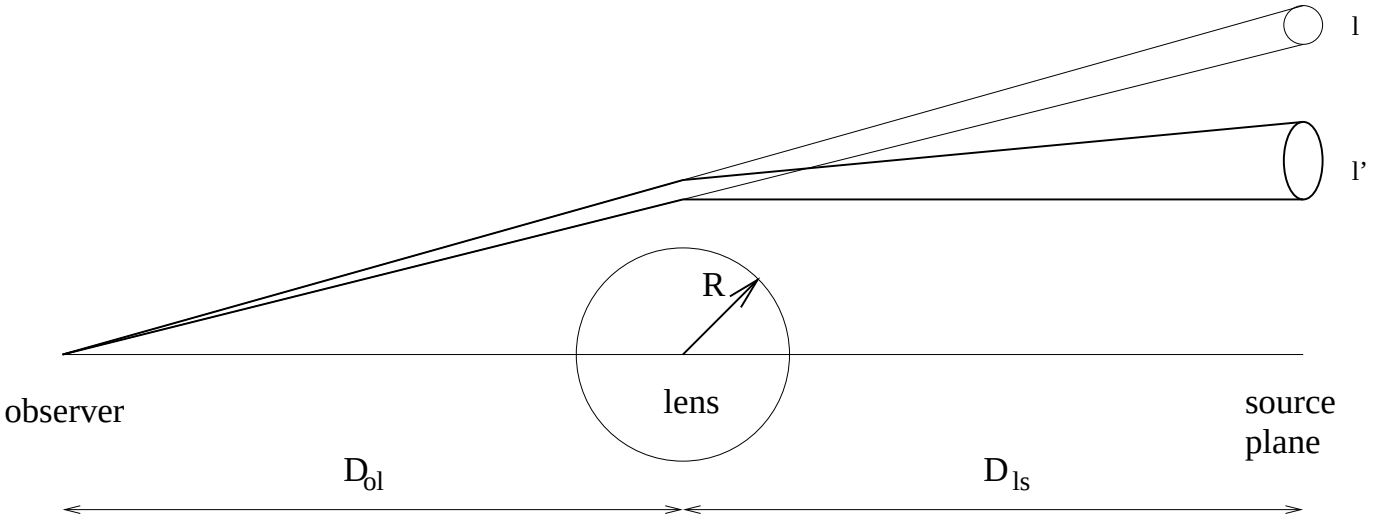

Figure 32: Schematic illustration of weak lensing. In the absence of the deflecting lens, a conical bindle of rays emerging from the observer will intercept the source plane in a circle. In the presence of the lens, the bundle will intercept the source plane in an ellipse. An object of this ellipticity on the source plane will therefore apprear circular. Conversely, a circular object will appear elliptical, but with an ellipticity of the opposite 'sign' (i.e. it will appear stretched along the tangential direction in this example). As shown in the main text, the fractional stretching of the image, also known as the 'image-shear', is $\gamma = l'/l - 1 = (D_{ls}/D_{os})\partial\theta_{\rm def}/\partial\theta$.

the differential deflection — i.e. the fact that the deflection suffered by the light from one side of a distant galaxy is slightly different than the deflection for the other side. This causes a systematic distortion, or 'shearing', of the shapes of the distant galaxies. The effect is similar to the distortion of distant objects seen in a 'mirage', though is a much weaker effect.

We can estimate the size of the effect as follows: Consider a mass concentration of size R and mass M at a distance D_{ol} from the observer as shown schematically in figure 32. Now consider a thin conical bundle of rays with opening angle θ_0 emerging from the observer and propagating back to some distant 'source plane'. In the absence of the lens this bundle would intersect the source plane in a circle of radius $l = D_{\rm os}\theta_0$. Now the lens will introduce some deflection $\theta_{\rm def} \sim GM/c^2 b$, and there will also generally be some change in the deflection across the bundle of $\Delta\theta_{\rm def} \sim GM\delta b/c^2 b^2$. This relative deflection will cause the initially circular bundle of rays to become elliptical. The length of the ellipse in the radial direction will be $l' = l + D_{\rm ls}\Delta\theta_{\rm def}$. The fractional stretching of the ellipse, or 'image shear', is

$$\gamma = \frac{l'}{l} - 1 = \frac{D_{\rm ls}}{D_{\rm os}}\frac{\Delta\theta_{\rm def}}{\theta_0}. \tag{7}$$

Now we can also write the differential deflection as $\Delta\theta_{\rm def} = \theta_0 d\theta_{\rm def}/d\theta$, where $d\theta_{\rm def}/d\theta$ is the *distortion tensor*; it gives the rate of change of deflection angle with position on the sky. The image shear is then

$$\gamma = \frac{D_{\rm ls}}{D_{\rm os}}\frac{d\theta_{\rm def}}{d\theta}. \tag{8}$$

Now consider an over-dense lens of size R and density contrast $\delta\rho/\rho$. The angular size of the lens is $\theta \sim R/D_{\rm ol}$. The deflection angle is

$$\theta_{\rm def} \sim \frac{G\delta M}{Rc^2} \sim \frac{G\delta\rho R^2}{c^2}. \tag{9}$$

The deflection angle will vary smoothly with impact parameter, so the distortion is

$$\frac{d\theta_{\rm def}}{d\theta} \sim \frac{\theta_{\rm def}}{\theta} \sim \frac{G\delta\rho R^2/c^2}{R/D_{\rm ol}} \sim \frac{H^2 D_{\rm ol} R}{c^2}\frac{\delta\rho}{\rho} \tag{10}$$

where we have used $H^2 \sim G\rho$. The image shear is therefore

$$\gamma \sim \frac{H^2}{c^2}\frac{D_{\rm ol}D_{\rm ls}}{D_{\rm os}} R\frac{\delta\rho}{\rho}. \tag{11}$$

Note that the strength of the effect scales as the product of the density and the size of the object; i.e. it is proportional to the surface density of the lens.

This is the effect due to a single object. This may be relevant for highly non-linear objects such as clusters of galaxies, where the space filling factor is small and the probability that a line of sight intercepts such an object is small. For large-scale structures with $\delta\rho/\rho \lesssim 1$ the filling factor is of order unity, and the net effect will be the superposition of a large number of structures along the line of sight. In fact, the shear is the integral of the tidal field along the line of sight. Each structure will give a contribution to the shear with a random direction and strength. This means that the net shear variance $\langle\gamma^2\rangle$ will be the sum of the individual shear variances, or, equivalently, that the net effect will be larger than that from a single object by roughly $\sqrt{N}$ where $N \sim D_{\rm os}/R$ is the number of structures. The factor $D_{\rm ol}D_{\rm ls}$ means that we get a small effect from structures which are very close to either the source or the observer, so, to get a crude estimate of the effect we can assume that all of the distances are of the same order of magnitude. This gives the prediction for the root mean square shear:

$$\langle\gamma^2\rangle^{1/2} \sim \frac{H^2 D_{\rm os}^{3/2} R^{1/2}}{c^2} \left\langle \left(\frac{\delta\rho}{\rho}\right)^2 \right\rangle^{1/2} . \tag{12}$$

This tells us that the strength of the effect increases as the $3/2$ power of the distance to the source; therefore to detect the effect one wants to use sources at cosmological distance (this also gives a large number of background galaxies, with also helps). For sources at $z \sim 1$, the distance is $D \sim c/H$, so the strength of the effect is then

$$\langle\gamma^2\rangle^{1/2} \sim \sqrt{\frac{R}{D}} \left\langle \left(\frac{\delta\rho}{\rho}\right)^2 \right\rangle^{1/2} . \tag{13}$$

This would suggest that, for super-cluster scale structures with $R \sim 10$Mpc and $\delta\rho/\rho \sim 1$, the root mean square shear would be on the order of 5%. More careful estimates — putting in factors like 4π and geometric factors properly — gives a predicition for shear of order 1% on scales of one degree.

This is a very small effect. A 1% shear means that a circular object will appear elliptical, mith major axis about 1% larger than the minor axis. However, galaxies are already elliptical, with root mean square ellipticity of about 30%, so the effect cannot be detected for a single object. What makes the effect measureable is that the shear is *spatially coherent*. The effect from super-cluster and larger scale structure will be coherent over large angular scales. Now the number of background galaxies — the 'cosmic wallpaper' — becomes very large; one can readily detect $\gtrsim 10^5$ galaxies per square degree. By looking for a statistical tendency for the galaxy position angles to be anisotropically distributed one can measure shears on the order of $\sim 0.3/\sqrt{N\,\mathrm{gal}}$, which is comfortably smaller than the prediction on degree scales. This is just the statistical error; to keep the systematic errors below this requires very careful analysis of the images. Current measurements find consistent results on scales of $\lesssim 10'$, and here is little data on larger scales as yet, but several large-scale surveys are being carried out, so the outlook is promising.

4.5.3 State of the art of cosmic shear measurements

Gravitational lensing has been characterised as 'a theorists heaven and an observers hell', and this applies also to weak-lensing. The promise of this technique has been apparent for a long time, but the observations and the required analysis are difficult. The current state-of-the-art is illustrated in figure 33.

The pace of observational development will pick up in the near future as there are two major missions about to enter service. Both of these were designed with weak-lensing as one of the 'science-drivers'; the LSST (or Rubin Observatory) on Chile and the ESA Euclid space-mission. These are shown in figure 34 which also shows an artists impression of the 'square-kilometer-array' radio telescope which will be built on a somewhat longer timescale.

5 Cosmology with clusters of galaxies

5.1 Overview of cluster-cosmology

Clusters of galaxies are the most massive (up to $\sim$ few$10^{15} M_\odot$) bound and virialized objects. They are have many applications in cosmology:

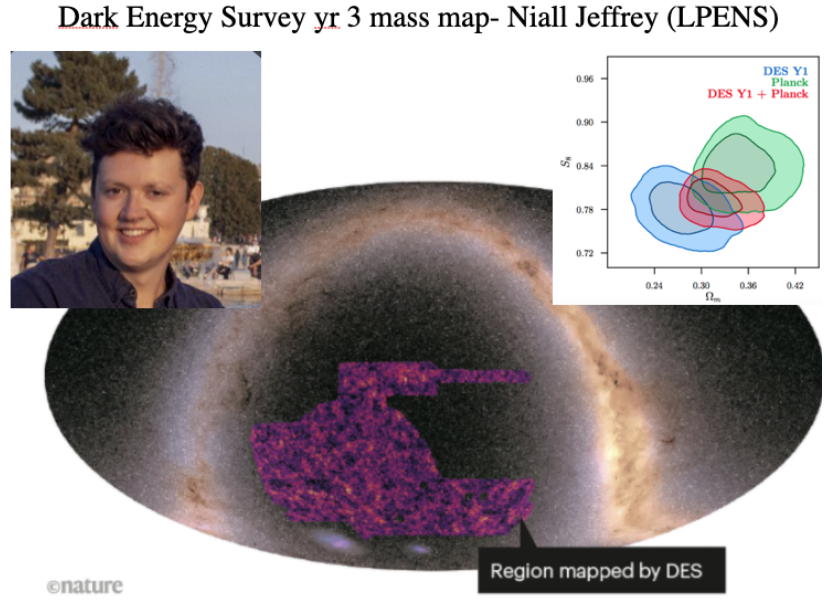

Figure 33: Weak lensing mass map from the 'dark-energy sky surve' project DES. The DES collaboration refitted a 4m class telescope in Chile with a wide-field corrector and a very large CCD detector. This was designed to carry out a massive survey over a period of about 5 years to image a large area of the sky. The 3 year results are shown here in the form of a map of the projected mass density obtained by 'inverting' the shear measurements of many hundreds of millions of background galaxies. The analysis was led by Niall Jeffrey. The inset on the upper right shows the constraints on the cosmological parameters: the amount of DM $\Omega_{\rm m}$ and the clustering strength.

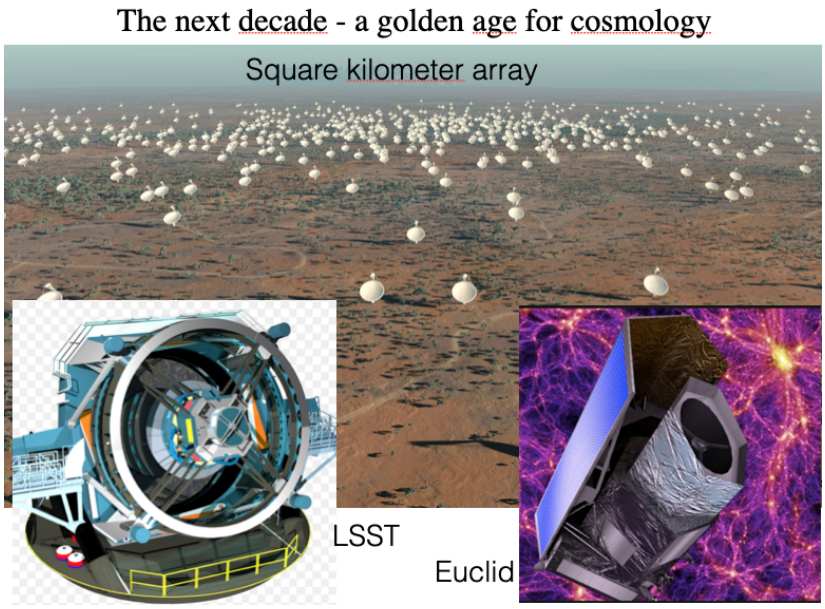

Figure 34: At the bottom left is shown the Rubin-Observatory (a.k.a. the large synoptic survey telescope (LSST)) which is under construction in Chile. It was fathered by Tony Tyson, who pioneered weak lensing measurements, and who realised the potential of this technique and the need for a dedicated survey instrument. It is a 10m class telescope with a huge field of view and will survey half the sky. The Euclid satellite will perform imaging and spectroscopy over about a third of the sky. Its primary mission is to do weak lensing and BAO oscillations. On a longer timescale the SKA will extend these measurements into the radio domain.

- their *masses*, determined from velocity dispersions, together with their stellar luminosities, provide strong evidence for *dark matter*
  - this was known since Zwicky's work in the 30's, but under-appreciated
- their *gravitational lensing effect:*
  - tests the *general relativity* prediction for light bending
  - provides us with powerful *natural telescopes* to image faint distant galaxies that are otherwise too faint to detect
  - constrains the *nature* of the dark matter
    - the dynamics of the *bullet cluster* shows that the dark matter must be collisionless
    - micro-lensing of high-z stars constrains how much of the DM can be in the form of black holes or compact 'mini-haloes' or in 'granules' in the case of fuzzy DM
  - can constrain the age of the universe through the *time delay effect*
- their *X-ray emission* from thermal bremsstrahlung provides another mass estimate
  - this constrains hypothetical *fifth forces* acting in the 'dark-sector' and

  - allows a detailed accounting of the *mass budget* – i.e. the relative abundance of dark matter and 'baryons', as well as a break-down of how much of the latter is in stars and how much remains gaseous

- the *evolution of the cluster mass function* is sensitive to the expansion history of the universe
  - this provided one of the earlier indications of the need for *dark energy*
- they affect the cosmic microwave background through the *Sunyaev-Zel'dovich (SZ) effects*
  - the *thermal SZ effect*
    * measures the integral of the ionized gas pressure along the line of sight
    * provides an efficient way to detect clusters at high redshift
  - while the *kinetic SZ effect* provides
    * a probe of *'bulk-flows'* that can potentially be applied at high redshift
    * a near-unique test of *homogeneity* (rather than just *isotropy*) of the universe

5.2 George Abell's cluster catalog

- George Abell discovered thousands of galaxy clusters by visual inspection of photographic plates from the Palomar (Schmidt telescope) survey
- He identified them by eye, obtained redshifts, and then classified them according to his 'richness' – the number of galaxies within 2 magnitudes of the 3rd brightest galaxy lying within projected distance $1.5h^{-1}$ Mpc.
  - where $h \equiv H_0/100$km/sec/Mpc is a historical relic from the time when the actual value of H_0 was uncertain by a factor ~ 2
- Rich clusters have velocity dispersion $\sigma_v \equiv \langle (v - \overline{v})^2 \rangle \simeq 1000$ km/s
  - that is the line-of-sight velocity – i.e. only one component out of three
  - but it is enough to get a reasonable estimate of the mass of such a cluster as follows:
    * if we assume that the velocities of galaxies are locally isotropic the mean square tangential velocity (having 2-components) is $\sigma_\perp^2 = 2\sigma_v$, so the root-mean-squared tangential velocity is about $v_\perp \simeq 1400$km/sec
    * for a galaxy that happens to be on a circular orbit (i.e. one with only tangential motion)
    * $v_\perp^2 = G_{\rm N} M(<r)/r \quad \Rightarrow \boxed{M(<r) = v_\perp^2 r/G_{\rm N} = 2\sigma_v^2 r/G_{\rm N}}$
    * if we divide mass interior to r by the volume $V = (4\pi/3)r^3$ we get the density
    * $\rho(<r) \equiv M(<r)/V \simeq (3/2\pi G_{\rm N})\sigma_v^2/r^2$
    * which we can compare to the critical density of the universe
    * $\rho_{\rm crit} = 3H_0^2/8\pi G_{\rm N}$
    * so the density, in units of critical, is
    * $\rho/\rho_{\rm crit} \simeq 4\sigma_v^2/H_0^2 r^2 \simeq 180(\sigma_v/1000{\rm km/sec})^2 (r/1.5h^{-1}{\rm Mpc})^{-2}$
    * where we see that the uncertainty in the Hubble constant has dropped out
- Thus the more massive objects have a mass density – mean density interior to $r_{\rm Abell}$ – of about 200 times the critical density
- Interestingly – and not coincidentally – this is roughly equal to the density expected for recently *virialized* objects
  - objects which turned around when the universe was about 1/2 its present age (see below for details)

Figure 35: Two rich Abell clusters (A1689 an A2274). Both of these are dominated by elliptical galaxies having old – 'red and dead' – stellar populations (don't pay any attention to the apparent difference in colour here; that is purely cosmetic).

5.3 Cluster Masses from Galaxy Motions

- cluster masses can be obtained from the virial theorem or from Jeans's equation
  - the latter has the advantage that it doesn't assume that 'light-traces-mass'
    - it requires measurement of $n(r)$ and velocity dispersion (tensor)
    - which require 'de-projection' which usually assumes spherical symmetry
    - a somewhat dubious assumption as clusters tend to be highly irregular
  - a significant uncertainty comes from the unknown *orbital velocity dispersion anisotropy*
  - the inferred mass is larger (smaller) if orbits are preferentially tangential (radial)
    - for a spherical, equilibrated, cluster
    - $\boxed{n^{-1}d(n\sigma_r^2)/dr + 2\beta\sigma_r^2/r = -d\phi/dr}$
    - where $\beta \equiv 1 - (\sigma_\theta^2 + \sigma_\phi^2)/2\sigma_r^2$ is the *velocity dispersion anisotropy parameter*
      - $\beta = 0$ for isotropic dispersion with $\sigma_r^2 = \sigma_\theta^2 = \sigma_\phi^2$
      - $\beta = 1$ for highly *radial* anisotropy ($\sigma_\theta^2,\ \sigma_\phi^2 \ll \sigma_r^2$).
      - $\beta < 0$ for tangential anisotropy
    - as an example, if
      - $n \propto r^{-2}$ (quite reasonable) and
      - σ^2 independent of r (also reasonable) and
      - $\beta = 1$ (extreme)
    - then left hand side vanishes $\Rightarrow d\phi/dr = 0 \Rightarrow M = 0$
      - recall the physical meaning of Jeans's equation
      - LHS is the *divergence* of the momentum flux density $n\langle v_i v_j\rangle$. Consider a conical 'plug'. With radial orbits, no momentum is flowing through the walls of the plug. And if $n \propto 1/r^2$ the flux at the top of the plug is the same as at the bottom. Hence, *kinematically,* there is no 'build up' of momentum in the volume. And hence the required 'gravitational return current' is zero.
  - N-body simulations suggest clusters have mild radial anisotropy an that this is a $\sim 20\%$ 'nuisance factor'

- Fritz Zwicky was the first to do this (in the '30s)
  - he found mass-to-light ratios $M/L \simeq$ few hundred times solar!

- much bigger than that for normal stellar population or luminous regions of galaxies
- *there is much more mass* ($\sim 30\times$) *than we see in stars*
- this became the so-called *'missing mass problem'* (though it was really the *light* that was missing

- If you take that M/L for clusters and multiply by the measured mean luminosity density $\mathcal{L}$ of the universe from redshift surveys, you get the mean *mass* density of the universe ρ.
  - this assumes M/L of clusters is representative of the universe – I.e. $M/L = \rho/\mathcal{L}$
- comparing to the *critical density* $\rho_{\rm crit} \equiv 3H^2/8\pi G$ you get the *density parameter* $\Omega = \rho/\rho_{\rm crit} \simeq 0.2$
- a more elegant way to do this is to use the *cosmic virial theorem* which relates ζ to the *pairwise velocity dispersion*
  - gives essentially the same result
- conclusions:
  - there is *a lot* of dark matter
  - but apparently not enough to 'close the universe'

5.4 Clusters observed in X-rays

- Clusters of galaxies emit X-rays via *thermal bremsstrahlung* from gas that is in hydrostatic equilibrium (at $T \sim 10^{7-8}$K) in the potential well of the DM
  - in the 'hierarchical' model of structure formation, clusters form by merger of smaller cluster, groups etc.
  - In the process, the gas gets *shock heated* in order to come to hydrostatic equilibrium

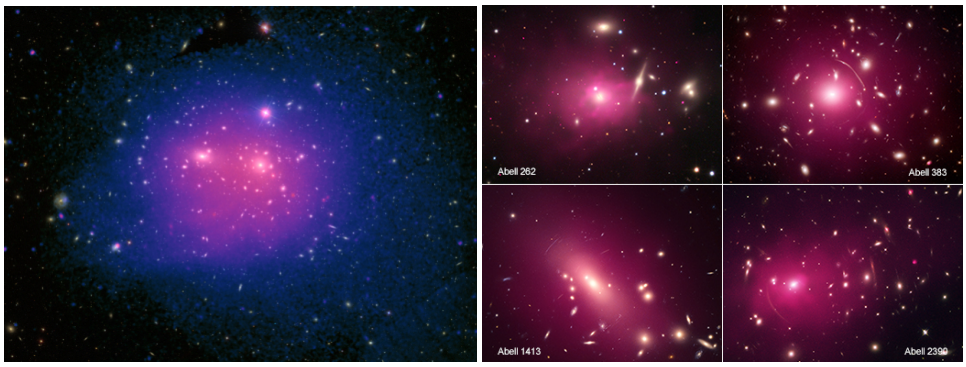

Figure 36: Left: the Coma cluster – a nearby but very rich cluster – in optical with XMM-Newton satellite X-ray (few keV energies) image superposed. Right: 4 Abell clusters observed by the Chandra satellite. These show the X-ray emission to be highly centrally concentrated as compared to the galaxies.

5.4.1 Thermal bremsstrahlung

- the key features of *thermal bremsstrahlung* –'braking radiation' – emission can be understood *semi-classically:*
  - electrons are accelerated as they pass by ions and radiate (see figure 37)
    - called *free-free emission*
    - negligible emission from the more massive ions
    - and negligible emission from electron-electron collisions
      - as there is no time varying dipole moment

- the emission depends on how fast the electrons are moving
  - collisions establish Maxwellian distributions with equipartition of energy

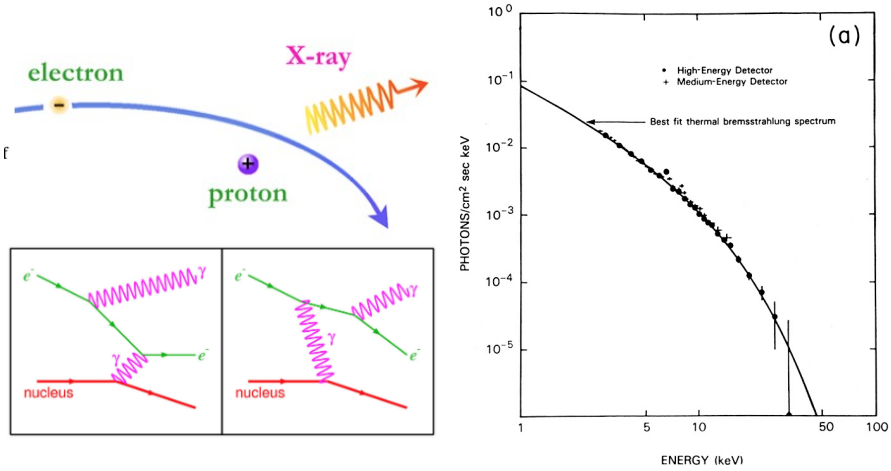

Figure 37: Left: illustration of the process that gives rise to bremsstrahlung emission and its QED Feynman diagrams. Hot gas emits X-rays with $h\nu \sim kT$. Right: X-ray spectrum of the Coma cluster. Shape of spectrum gives the temperature and – via hydrostatic equilibrium – the mass. The brightness gives the density – and hence amount – of the gas.

  - so pressure $P = nkT = (n_{\rm ions} + n_e)kT$ comes from both electrons and ions
  - density is strongly dominated the ions: $\rho \sim \rho_{\rm ions}$
  - so the equation of hydrostatic equilibrium is
    * $\boxed{\nabla P = g\rho = g\mu m_p n}$
    * where $\mu \equiv \sum_i m_i n_i / m_p \sum_i n_i$ is the mean mass per particle in units of the proton mass, and which implies
    * $\boxed{kT \simeq \mu m_p \phi}$
  - as a consequence,
    * the velocity of ions is similar to the velocity of galaxies, or $\sim 0.003c$
    * while the velocity of the electrons is larger by factor $\sqrt{m_p/m_e} = \sqrt{2000} \simeq 50$
    * so electrons are non-relativistic (though not *very* non-relativistic)

- consider the energy emitted from a single collision with impact parameter b:
  - acceleration is $a \simeq q^2/\epsilon_0 m_e b^2$ so $\ddot{d} = q\ddot{r} = qa \simeq q^3/(\epsilon_0 m_e b^2)$
  - Larmor: radiated power (classical) $P \simeq \ddot{d}^2/\epsilon_0 c^3 \simeq q^6/(\epsilon_0^3 c^3 m_e^2 b^4)$
  - one collision lasts a time $t \sim b/v$
  - so the (classical) energy radiated per collision is
    * $\boxed{\epsilon_1 = P \times b/v \sim q^6/(\epsilon_0^3 c^3 m_e^2 b^3 v)}$
  - which is emitted in waves with frequency $\boxed{\nu \sim v/b}$
    * Fourier transform of a 'pulse' of width $\tau \sim 1$ has power dominated by frequencies $\nu \sim 1/\tau$

- power radiated by one electron
  - for one electron, the rate of collisions with impact parameter b is $\sim n_+ b^2 v$
    * where n_+ is the density of protons and we will assume, for simplicity, a plasma composed purely of hydrogen
  - so the power radiated by one electron is $P_1 \sim nb^2 v \times \epsilon_1 \sim n_+ q^6/(\epsilon_0^3 c^3 m_e^2 b)$
  - $P_1 \propto 1/b$ so most of the power comes from lowest b (highest ν)
  - Q: what is the 'cut-off' for the impact parameter?
  - A: depends on T, but for relevant temperatures is it quantum mechanics: electrons have a de Broglie wavelegth $\lambda_{\rm dB} \sim \hbar/p \rightarrow b_{\rm min} \sim \hbar/mv$
  - this gives the *semi-classical* power radiated (per electron)
    * $\boxed{P_1 \sim n_+ q^6 p/(\epsilon_0^3 c^3 m_e^2 \hbar)}$
    * where $p \sim \sqrt{kT}/m_e$ is the momentum

- with 'typical' photon energy
  - $\boxed{E_\gamma = h\nu \sim hv/b_{\rm min} \sim m_e v^2 \sim kT}$
  - caution:
    - this is typical in the sense that the photons that carry most of the energy have this sort of energy
    - there are actually a logarithmically divergent *number* of low energy photons emitted in softer collisions,

- emissivity for a plasma
  - multiplying P_1 by the number density of electrons gives the *free-free emissivity* (energy/time/volume)
    - $\boxed{\epsilon \sim n_e n_+ q^6 p/(\epsilon_0^3 c^3 m_e^2 \hbar)}$

  1. this analysis is 'semi-classical' in the following sense:
     - we computed the mean power radiated using a picture of classical particles radiating as predicted by Larmor
     - but with a cut-off that comes from quantum mechanics
     - this is confirmed by quantum electro-dynamics (QED)
  2. while the characteristic photon energy is on the order of $kT \simeq m_e v^2/2$ the mean energy released per 'encounter' is $\epsilon_1 \sim (q^2/\epsilon_0 c\hbar)^3 m v^2/2$
     - smaller by a rather large factor $\sim \alpha^3$
     - consistent with QED
       - diagrams have 3 vertices, gives amplitude proportional to q^3 (or $\alpha^{3/2}$)
       - so probability is proportional to q^6 (or α^3)
       - cross-section is $\sigma \sim \lambda_{\rm dB}^2 \alpha^3$

5.4.2 Thermal bremsstrahlung from galaxy clusters

- Key properties of thermal bremsstrahlung:
  - bolometric emissivity scales as $n^2 T^{1/2}$
    - so with gas density profile $n \propto r^{-2}$ most emission (volume r^3 times emissivity $\propto r^{-4}$ comes from where the plasma is densest – the emission being dominated by the 'core' of the cluster
    - this explains why clusters look very compact in X-rays
  - the observed brightness gives the integral of $n^2 \Rightarrow$ along the line of sight
  - the energy cut-off (with the detailed spectrum computable from QED) at $E_\gamma \sim kT$
    - allows one to estimate the temperature and hence (via hydrostatic equilibrium) the cluster mass
    - with no worries about orbital anisotropy
    - some worries about 'non-thermal' sources of pressure (magnetic fields, turbulent motions ...)

- Key results:
  - mass agrees well with dynamical mass estimates
    - constrains theories with '5th forces' augmenting gravity in the 'dark-sector' if these forces affect the galaxy motions
  - amount of mass in gas is
    - $\sim 5\times$ more than in stars, but
    - $\sim 5\times$ less than the DM
  - assuming clusters contain a representative sample of stuff
    - density of ordinary (i.e. non-dark) matter is $\sim$ few percent of critical density
    - which agrees with the density inferred from big-bang nucleosynthesis (BBN)

5.4.3 Cooling flows in clusters

- for the bulk of the plasma in a cluster the 'cooling rate' Γ due to thermal bremsstrahlung
  - defined to be the emissivity divided by the thermal energy density

  is small compared to the expansion rate for the universe
- so the cooling time is longer than the age of the universe
- but for some clusters, the density of gas in the core is sufficiently high that they lose a significant fraction of energy in the age of the universe
- we can still model these using hydrostatic equilibrium
  - as the cooling rate is smaller than the inverse dynamical time

  but as they lose energy, the gas in the centre will become more concentrated, which is an unstable process which must be limited by either the gas turning into stars or there being some kind of feedback

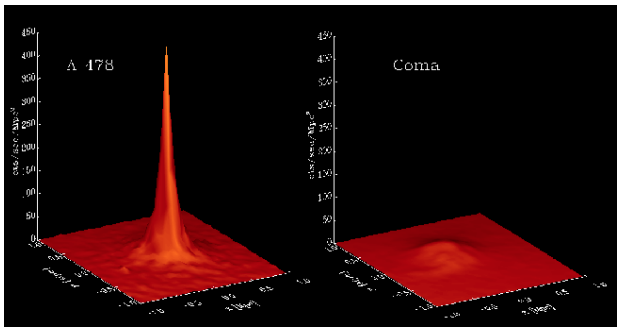

Figure 38: One the left is the X-ray brightness of the cluster Abell 478: a 'cooling flow' cluster. In this cluster the relatively intense emission at the centre gives a cooling time that is less than the age of the universe. On the right is the Coma cluster; a more massive cluster, and typical of very massive clusters for which cooling is not significant even in the very centre.

5.5 The Sunyaev-Zel'dovich effect

- if we take the observed X-ray luminosity $L_{\rm X}$ and temperature $T_{\rm X}$ (inferred from spectrum fitting) for a rich cluster and use $L_{\rm X} \sim \epsilon r^3$ with thermal bremstrahlung emissivity $\epsilon \propto n_e^2\sqrt{T_{\rm X}}$ then we find the density of electrons
  - $n_{\rm e} \sim 3\times10^{-3}h^{-1/2}{\rm cm}^{-3}(r/200{\rm kpc}/h)^{-3/2}(T_{\rm X}/4{\rm keV})^{-1/4}(L_{\rm X}/10^{44}h^{-2}{\rm erg/sec})^{1/2}$
  - where we have used fiducial values for luminosity, temperature and core radius characteristic of a typical rich cluster
- if we multiply this by the Thomson cross section $\sigma_{\rm T} \simeq 6\times10^{-29}{\rm m}^2$ and the radius we get an estimate for the optical depth for electron scattering
  - $\boxed{\tau_{\rm es} \sim 3\times10^{-3}}$
- this means that a fraction of a percent of the photons coming to us from the CMB will have been scattered 'out of the beam' coming through a cluster
- and a similar number of photons that we would otherwise not have seen will have been scattering into the beam
- Rashid Sunyaev an Jacob B. Zel'dovich realised that, owing to the motion of the electrons doing the scattering, this would have very interesting observable effects

5.5.1 The thermal SZ effect

- The electrons in the hot plasma in clusters scatter photons from the *cosmic microwave background* (CMB) radiation
- electrons are non-relativistic: so this is *Thompson scattering*
  - scattering is 'elastic' in the electron rest-frame
  - in cluster frame ('lab-frame') there is change in photon energy
  - net effect is to shift the spectrum slightly to higher energies
  - results in a decrement (enhancement) of flux density F_ν at low (high) frequencies
  - illustrated schematically in figure 39

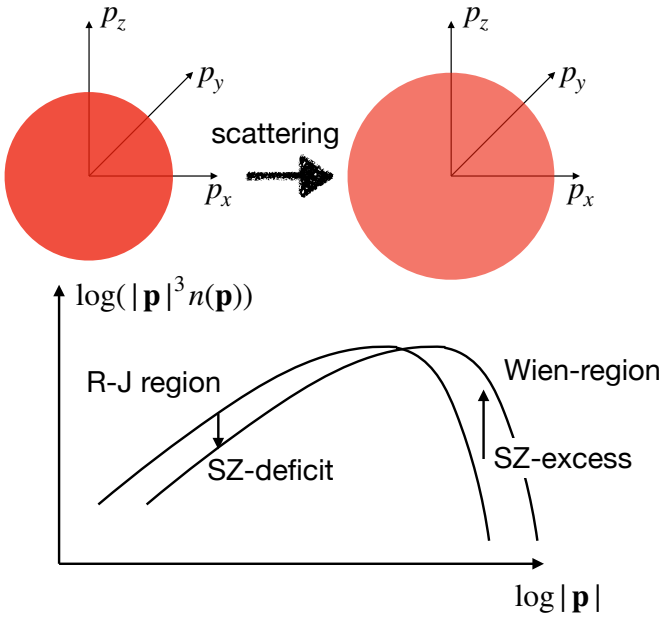

Figure 39: The thermal Sunyaev-Zel'dovich effect. Consider a box containing thermal radiation and hot plasma. Initially the photons have a Planck spectrum with occupation numbers $n(\mathbf{p}) = 1/(e^{h\nu/kT}-1)$. A fraction of these photons get scattered by the rapidly moving electrons – giving a random kick in 3-dimensional momentum space. The number of photons is not changed, but, the momentum being a vector' it's mean modulus is increased. That means that the photons occupy a larger spherical volume in $\mathbf{p}$-space. The result, when we observe the microwave background in a direction towards a cluster we see an excess at high frequencies and a deficit in the Rayleigh-Jeans region. There is a certain weighted integral of the spectrum that 'nulls-out' the thermal SZ effect. What is left is the 'kinematic' effect caused by any net motion of the clusters.

- described quantitatively by the collisional Boltzmann equation
  - a.k.a. 'Focker-Planck' or 'Kompaneets' equation.
- the observed 'decrement' measures the integral of $n_e\langle v_e^2\rangle$
  - i.e. the line integral of the pressure
  - linear in n_e so more sensitive to outer parts of cluster than X-ray emission
- unlike the X-ray, or other types of emission which suffer $(1+z)^{-4}$ surface brightness dimming
  - which makes cluster detection very difficult at even quite modest redshift
- the brightness fluctuation induced by the SZ effect is relatively redshift independent
  - making SZ surveys a good way to detect clusters at high redshift

5.5.2 The kinematic SZ effect

- The flip in sign of the thermal SZ effect means there is an integral of the intensity over frequency that 'nulls-out' the thermal SZ signal
- what is left is a *'kinematic'* effect arising from any line-of-sight peculiar peculiar motion
- this gives another probe of growth rate of large-scale structure
- becomes the most promising probe outside of the local universe

- alternatives like TF, $D_n - \sigma$ have constant *fractional* error in distance
- so error in peculiar velocity increases with distance
- not so for kinematic SZ

- note that this is one of the few ways to test the assumption of *homogeneity* in FRW cosmology
  - we know the universe is *isotropic* around us to high precision
  - but do we know that it is *homogeneous*?
  - a global inhomogeneity would require us to live in, or very close to, a special place in the Universe
  - but otherwise compatible with what's seen for the most part
  - and there are some claims that galaxy counts support the idea that we live near the centre of a large 'local underdensity'
  - such models are tightly constrained by lack of 'monopole' in the kinematic SZ measurements

5.5.3 Weak lensing and the bullet cluster

- the background galaxies that we see outside the giant arcs are also affected by gravitational lensing, but to a lesser degree: their shapes are distorted

- this is essentially a statistical effect
  - since galaxies are intrinsically elliptical it is impossible to say how much shape distortion an individual galaxy has suffered
  - but if we measure the shapes of hundreds of galaxies on a patch of sky
  - and assume that, on average, they are intrinsically isotropic
  - them we can determine the 'shear'
    - a 2-component quantity giving the 'polarisation' of the galaxy shapes
  - and from this one can determine the 'convergence' and hence the mass surface density

- a spectacular application of weak lensing is in the 'bullet cluster' (figure 40)

Figure 40: Composite image of the 'bullet-cluster'. This is a pair of clusters that have recently suffered a collision. Superposed on the optical image is shown the total mass density from weak lensing (blue) and the X-ray emission (red). The main lesson from this is that, unlike the X-ray emitting gas, the dark matter must be collisionless (or nearly so) as it has evidently 'passed through' itself during the collision.