

L3 Astro - Section 10 - Clusters of Galaxies

Nick Kaiser

Contents

1	Clusters of galaxies	2
1.1	Overview of cluster-cosmology	2
1.2	George Abell's cluster catalog	2
1.3	Cluster Masses from Galaxy Motions	3
1.4	Clusters observed in X-rays	4
1.4.1	Thermal bremsstrahlung	5
1.4.2	Thermal bremsstrahlung from galaxy clusters	7
1.4.3	Cooling flows in clusters	7
1.5	Gravitational lensing by galaxy clusters	8
1.5.1	Cluster masses from giant arcs	8
1.5.2	The Einstein radius and the critical surface density	8
1.5.3	Caustics and critical curves	10
1.5.4	The optical depth for strong lensing	11
1.5.5	Amplification bias and quasar galaxy associations	11
1.5.6	Micro-lensing by point masses	12
1.5.7	Weak lensing and the bullet cluster	13
1.6	A simple model for the formation of galaxy clusters	14
1.7	Evolution of the cluster mass function	16
1.7.1	Self-similar evolution	16
1.8	The Sunyaev-Zel'dovich effect	18
1.8.1	The thermal SZ effect	18
1.8.2	The kinematic SZ effect	19

List of Figures

1	Two massive Abell clusters	3
2	The Coma cluster	5
3	Bremsstrahlung emission from hot gas	5
4	A 'cooling-flow' cluster	7
5	Giant arcs from gravitational lensing of background galaxies	8
6	Two spectacular 'Einstein rings' produced by galaxy lenses	9
7	The Einstein radius for a point-mass lens	9
8	Model for a lens with a 'fold' caustic	11
9	Caustics and critical curves	12
10	Amplification bias	12
11	Geometry for a point-mass lens	13
12	Microlensing light curve	13
13	Microlensing light curve for a binary lens	14
14	Limits on MACHO dark matter	14
15	Composite image of the 'bullet-cluster'	15
16	The 'top-hat' over-density model for formation of a cluster	15
17	Self-similar model for evolution of the cluster mass function	17
18	The thermal Sunyaev-Zel'dovich effect	18

1 Clusters of galaxies

1.1 Overview of cluster-cosmology

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

- 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

1.2 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 $h \equiv H_0/100\text{km/sec/Mpc}$ 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 $\sigma_v \equiv \langle (v - \bar{v})^2 \rangle \simeq 1000 \text{ km/s}$

- 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 $\sigma_{\perp}^2 = 2\sigma_v^2$, so the root-mean-squared tangential velocity is about $v_{\perp} \simeq 1400 \text{ km/sec}$
 - * for a galaxy that happens to be on a circular orbit (i.e. one with only tangential motion)
 - * $v_{\perp}^2 = G_N M(< r)/r \Rightarrow M(< r) = v_{\perp}^2 r/G_N = 2\sigma_v^2 r/G_N$
 - * if we divide mass interior to r by the volume $V = (4\pi/3)r^3$ we get the density
 - * $\rho(< r) \equiv M(< r)/V \simeq (3/2\pi G_N)\sigma_v^2/r^2$
 - * which we can compare to the critical density of the universe
 - * $\rho_{\text{crit}} = 3H_0^2/8\pi G_N$
 - * so the density, in units of critical, is
 - * $\rho/\rho_{\text{crit}} \simeq 4\sigma_v^2/H_0^2 r^2 \simeq 180(\sigma_v/1000 \text{ km/sec})^2(r/1.5h^{-1}\text{Mpc})^{-2}$
 - * 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_{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 1: Two rich Abell clusters (A1689 and 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).

1.3 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
 - $n^{-1}d(n\sigma_r^2)/dr + 2\beta\sigma_r^2/r = -d\phi/dr$
 - * 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 $n(v_i v_j)$. 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 and 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* $\rho_{\text{crit}} \equiv 3H^2/8\pi G$ you get the *density parameter* $\Omega = \rho/\rho_{\text{crit}} \simeq 0.2$
- 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'

1.4 Clusters observed in X-rays

- Clusters of galaxies emit X-rays via *thermal bremsstrahlung* from gas that is in hydrostatic equilibrium (at $T \sim 10^{7-8}\text{K}$) 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

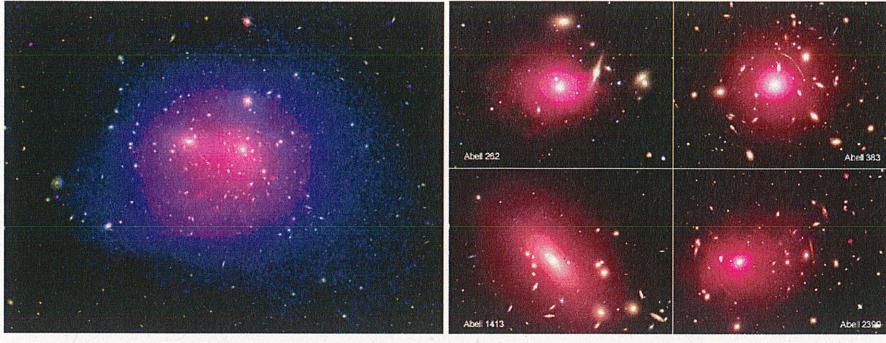


Figure 2: Left: the Coma cluster – a nearby but very rich cluster – in optical with XMM-Newton satellite X-ray (few keV energies) image superposed. Right: 4 Abell clusters observed by the Chandra satellite. These show the X-ray emission to be highly centrally concentrated as compared to the galaxies.

1.4.1 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 3)
 - 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

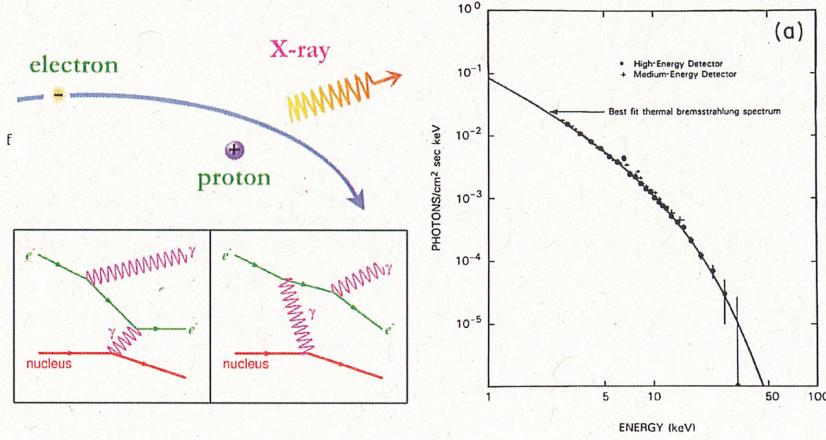


Figure 3: Left: illustration of the process that gives rise to bremsstrahlung emission and its QED Feynman diagrams. Hot gas emits X-rays with $h\nu \sim kT$. Right: X-ray spectrum of the Coma cluster. Shape of spectrum gives the temperature and – via hydrostatic equilibrium – the mass. The brightness gives the density – and hence amount – of the gas.

- the emission depends on how fast the electrons are moving
 - collisions establish Maxwellian distributions with equipartition of energy
 - so pressure $P = nkT = (n_{\text{ions}} + n_e)kT$ comes from both electrons and ions
 - density is strongly dominated the ions: $\rho \sim \rho_{\text{ions}}$
 - so the equation of hydrostatic equilibrium is
 - $\nabla P = g\rho = g\mu m_p n$
 - 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
 - $kT \simeq \mu m_p \phi$
 - 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
 - * $\epsilon_1 = P \times b/v \sim q^6/(\epsilon_0^3 c^3 m_e^2 b^3 v)$
 - which is emitted in waves with frequency $\nu \sim v/b$
 - * 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 $P_1 \sim nb^2v \times \epsilon_1 \sim n_+ q^6/(\epsilon_0^3 c^3 m_e^2 b)$
 - $P_1 \propto 1/b$ 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 wavelength $\lambda_{dB} \sim \hbar/p \rightarrow b_{min} \sim \hbar/mv$
 - this gives the semi-classical power radiated (per electron)
 - * $P_1 \sim n_+ q^6 p / (\epsilon_0^3 c^3 m_e^2 \hbar)$
 - * where $p \sim \sqrt{kT}/m_e$ is the momentum
 - with ‘typical’ photon energy
 - * $E_\gamma = h\nu \sim hv/b_{min} \sim m_e v^2 \sim kT$
 - * 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)
 - * $\epsilon \sim n_e n_+ q^6 p / (\epsilon_0^3 c^3 m_e^2 \hbar)$
 - 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 $\sigma \sim \lambda_{dB}^2 \alpha^3$

1.4.2 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)

1.4.3 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

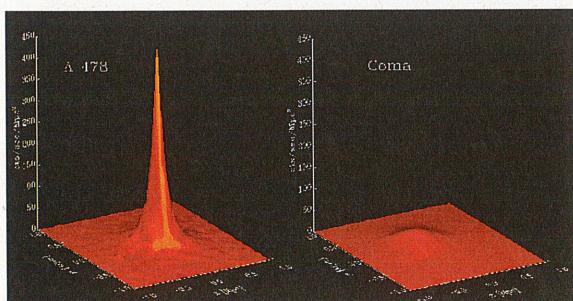


Figure 4: One the left is the X-ray brightness of the cluster Abell 478: a ‘cooling flow’ cluster. In this cluster the relatively intense emission at the centre gives a cooling time that is less than the age of the universe. On the right is the Coma cluster; a more massive cluster, and typical of very massive clusters for which cooling is not significant even in the very centre.

1.5 Gravitational lensing by galaxy clusters

1.5.1 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



Figure 5: Two massive clusters (A370 and RCS2-032727-132623) displaying ‘giant arcs’. These give quite precise measurements of the projected mass along a cylinder.

1.5.2 The Einstein radius and the critical surface density

- The geometry for calculating the Einstein radius θ_E is shown in figure 7
- 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:

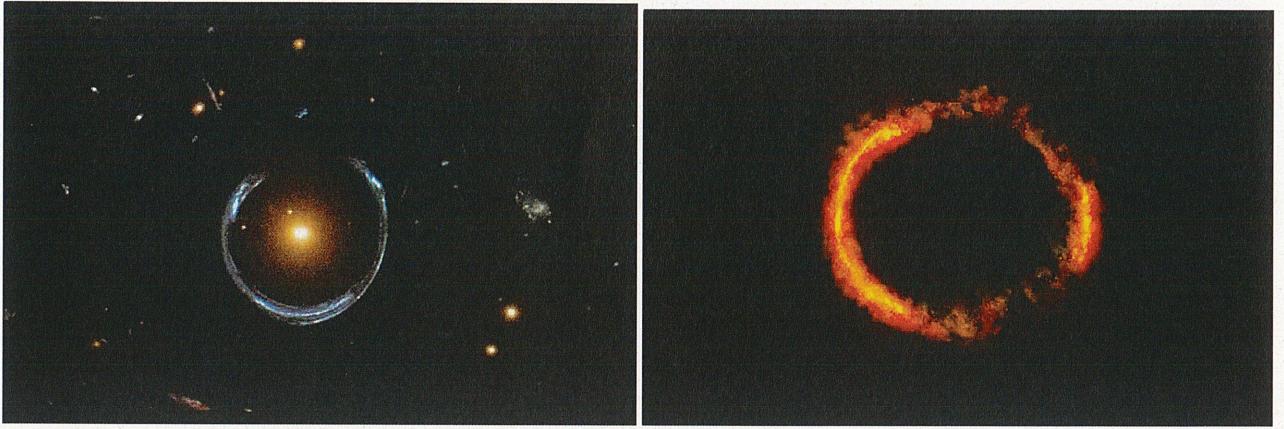


Figure 6: Two spectacular ‘Einstein rings’ – produced by galaxy lenses.

- $\delta\theta = 4GM/bc^2$

- assuming flat spatial geometry (it’s not difficult to generalise to hyperbolically curved space) the impact parameter is $b = a_L \chi_L \theta_E$ (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 :
- $\delta\theta = 4GM/a_L \chi_L \theta_E c^2$
- while the fact that the physical length of the vertical arrow at the right can be expressed as either $a_S \chi_{LS} \delta\theta$ or as $a_S \chi_S \theta_E$ gives us a second relation between $\delta\theta$ and θ_E :
- $\delta\theta = \theta_E \chi_S / \chi_{LS}$

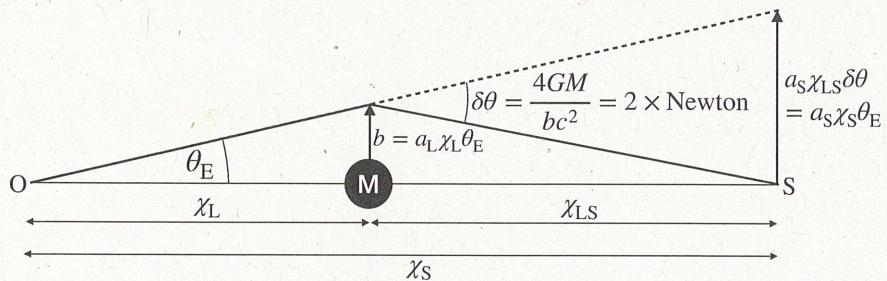


Figure 7: The Einstein radius is the angular radius of the ring that is seen by the observer O for a point-like source S that is exactly on-axis behind a point mass lens.

- equating these gives the *lens equation*

- $$\frac{4GM}{a_L \chi_L \theta_E c^2} = \frac{\chi_S}{\chi_{LS}} \theta_E$$

- whose solution is the *Einstein radius*

- $$\theta_E = \sqrt{\frac{4GM \chi_{LS}}{a_L c^2 \chi_L \chi_S}}$$

- which, in terms of distances $D(z_L) \equiv a_0 \chi_L$, $D(z_S) \equiv a_0 \chi_S$, and using $a_L = a_0/(1+z_L)$ is

- $$\theta_E(M, z_L, z_S) = \sqrt{\frac{4GM(1+z_L)}{c^2} \frac{D(z_S) - D(z_L)}{D(z_L)D(z_S)}}$$

- where one should note that the distances are neither angular diameter distances $D_a(z) \equiv a(z)\chi(z) = (1+z)^{-1} \int cdz/H$, 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: $D(z) = \sqrt{D_a(z)D_L(z)}$

- 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 $b = a_L \chi_L \theta_E = D(z_L) \theta_E / (1+z_L)$ we get the so-called **critical surface density**:

$$-\boxed{\Sigma_{\text{crit}}(z_L, z_S) = \frac{M}{\pi b^2} = \frac{c^2(1+z_L)}{4\pi G} \frac{D(z_S)}{D(z_L)(D(z_S) - D(z_L))}}$$

- which is the surface density of a mass-sheet at redshift z_L that is sufficient to re-focus light from a point source at redshift z_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 $z_S = 2$, this formula gives a **minimum critical surface density of about 0.4 g/cm^2 at $z_L \simeq 0.5$**
 - but with a broad minimum: lenses with $\Sigma > 0.6 \text{ g/cm}^2$ being capable of lensing such sources if they have $1 > z_L > 0.15$
- **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

1.5.3 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
- 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 our 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

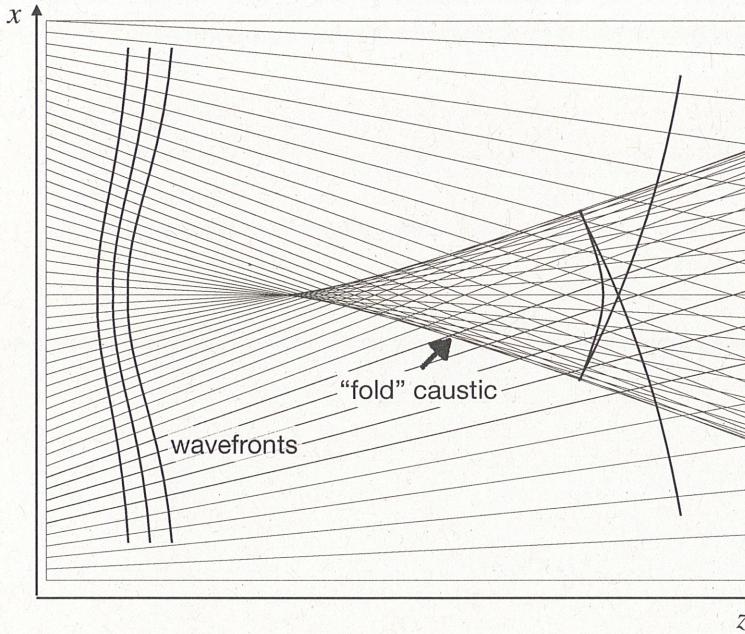


Figure 8: Simple model for a gravitational lens. Straight lines are light rays that have passed through a lens (not shown). This produced a time-delay for the light propagation and hence the wave-fronts are curved. The rays are normal to the wave-fronts and a certain distance down-stream they focus. 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 off 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

$$P(> A) \propto A^{-2}$$
 (as pointed out by Peter Schneider)
 2. there is always an odd number of images

1.5.4 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 $z_s = 2$, 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

1.5.5 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’

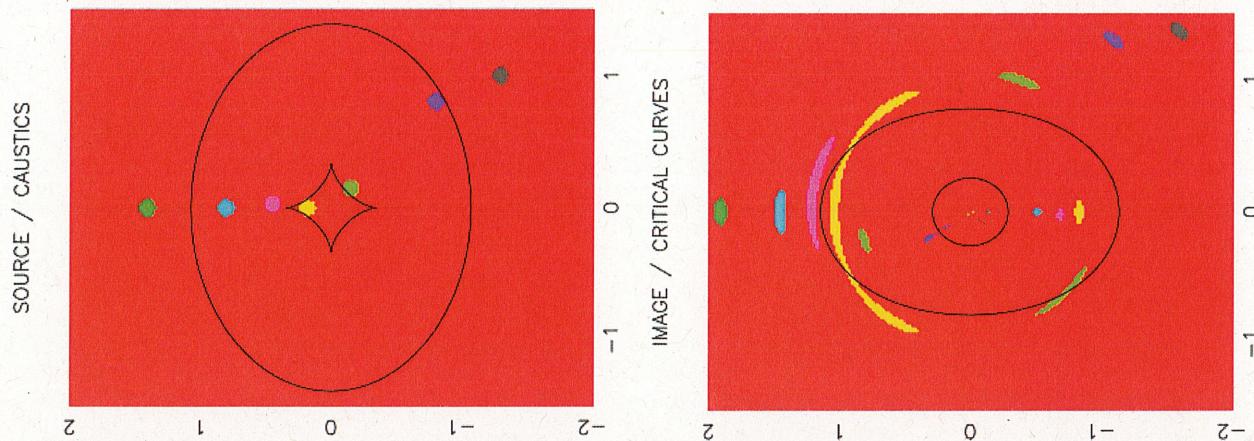


Figure 9: Caustics (left) on the source plane (along with various extended sources) and critical curves (right) on the image (or sky) plane for a slightly flattened, but otherwise spherical lens (made by Konrad Kuijken), are indicated as the black lines. Two of the green sources are outside the outermost 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

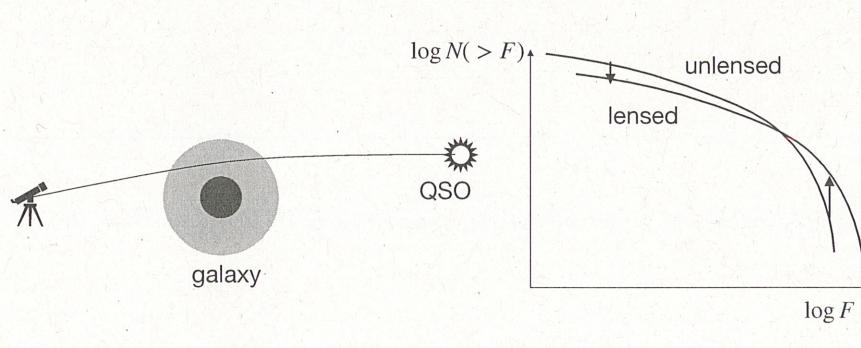


Figure 10: Amplification bias. At the left is shown a distant QSO, the light from which passes through the halo of a foreground galaxy or cluster. It will be magnified and its flux density amplified – so otherwise undetectable sources can be seen close to foreground lenses. But the number density is also diminished. This results in a deficit of faint objects.

1.5.6 Micro-lensing by point masses

- Another useful application of gravitational lensing in cosmology is micro-lensing of stars

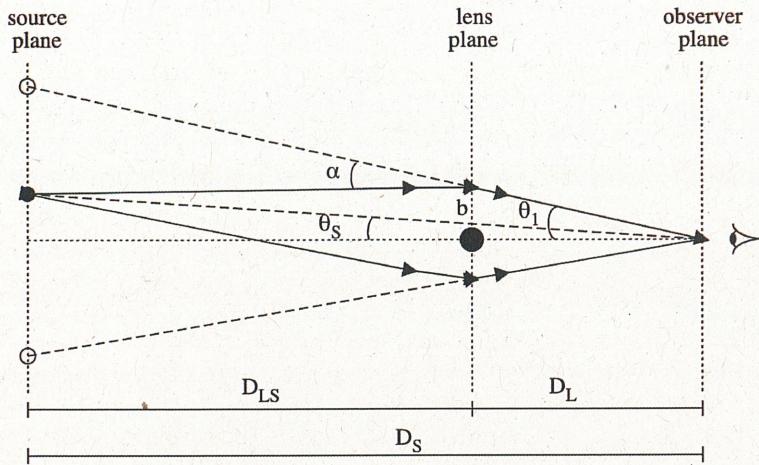


Figure 11: Geometry for lensing of a point-source by a point-mass. There are two images. If the ‘impact parameter’ is large – that is to say the un-lensed angle θ_S is much larger than the Einstein radius θ_E then the second image – seen close to the lens – will be very faint, but if $\theta_S \sim \theta_E$ the amplification (the sum of the flux densities for the two stars) will become large; very large for $\theta_S \ll \theta_E$. 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.

- the geometry for lensing by a point mass is shown in figure 11
- the total flux density amplification is

$$* \quad M = (u^2 + 2)/u\sqrt{4 + u^2}$$

* where $u \equiv \theta_S/\theta_E$

- 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 were 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

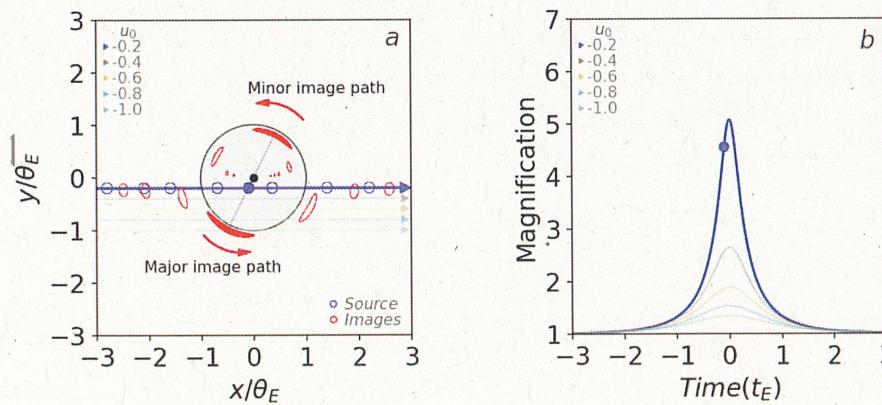


Figure 12: Microlensing light curve. Blue line at left is the path of a source behind a point mass lens. Red ‘bananas’ are the outline of the images. At right are light curves for various different values of the impact parameter. Images from Yiannis Tsapras.

1.5.7 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

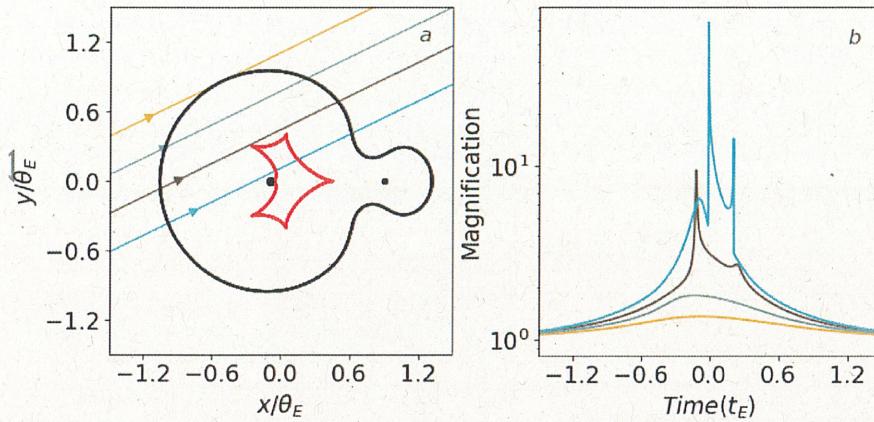


Figure 13: Microlensing light curve for a binary lens. The asymmetry of the lens has caused the line-like singularity lying along the axis behind a point mass lens to ‘degenerate’ and caustics (fold catastrophes) have formed. The special form of the point mass lens is very sensitive to even small external tidal fields and can be used to detect extra-solar planets. Images from Yiannis Tsapras.

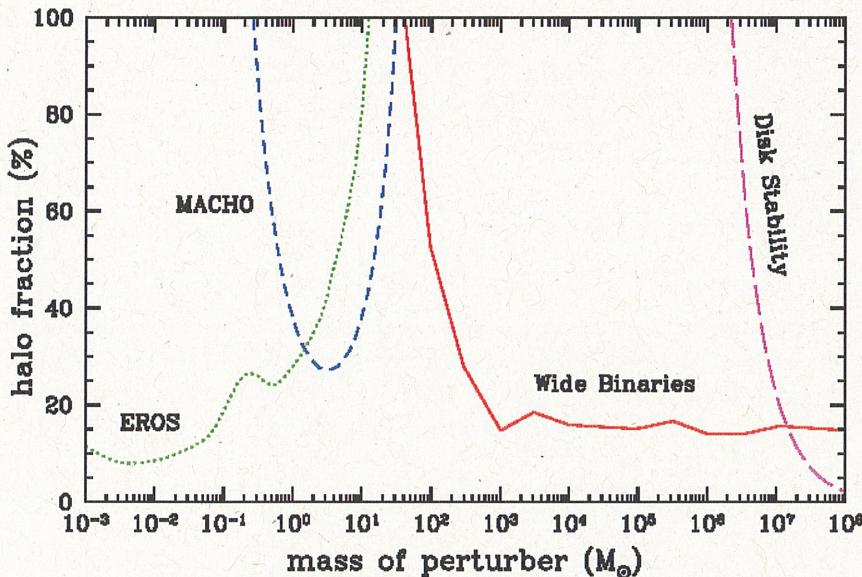


Figure 14: Limits on the fraction of dark matter in Massive Astrophysical Compact Halo Objects (MACHOS) from Yoo et al.. MACHO and EROS are two microlensing surveys. They ruled out the indicated mass ranges. The existence of wide, loosely bound, binary stars in the Milky Way have been argued to give the limits shown in red.

- 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
 - then 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 15)

1.6 A simple model for the formation of galaxy clusters

- We can estimate what we would expect for the density, and *density contrast*, of a recently virialised cluster using the following simple model:
 - we assume there is a ‘background’ cosmology which, for simplicity, is Einstein-de Sitter, so $r \propto t^{2/3}$
 - in that background we carve out a sphere of mass M
 - * which, in the background, would have been marginally bound to itself

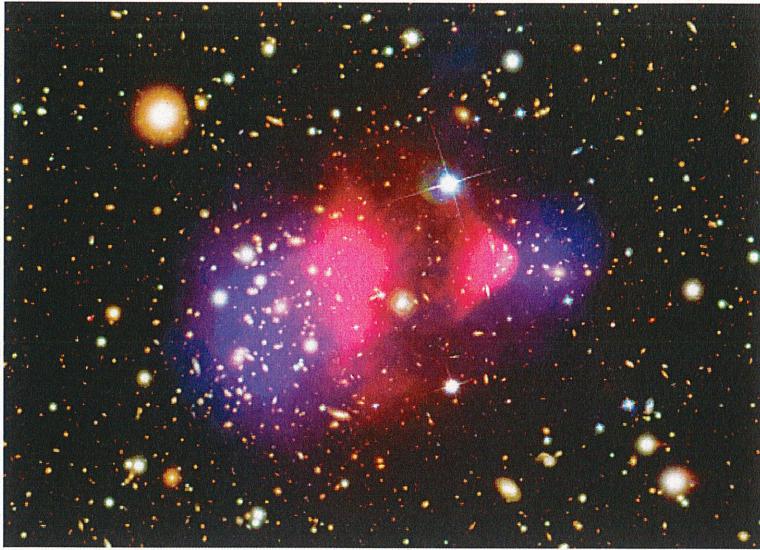


Figure 15: Composite image of the ‘bullet-cluster’. This is a pair of clusters that have recently suffered a collision. Superposed on the optical image is shown the total mass density from weak lensing (blue) and the X-ray emission (red). The main lesson from this is that, unlike the X-ray emitting gas, the dark matter must be collisionless (or nearly so) as it has evidently ‘passed through’ itself during the collision.

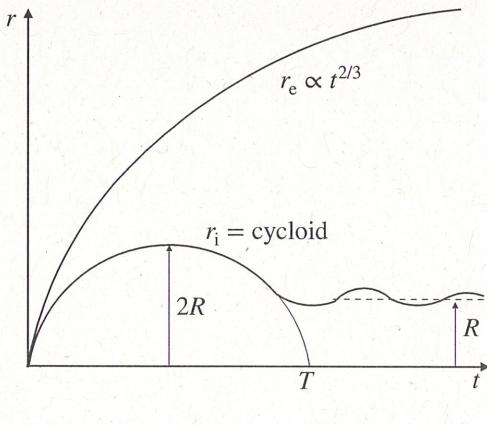


Figure 16: A simple model for the formation of a cluster is that it was initially a uniform density sphere ‘carved out’ of a uniform density ‘background’ universe but with a lower total energy, so it was gravitationally bound to itself and therefore doomed to expand only to some maximum radius and then collapse. The virial theorem tells us that, in order to generate enough kinetic energy to satisfy $2K+U=0$, it must collapse by about a factor 2 from its maximum size. This gives the Kepleresque relation between radius, mass and time of collapse: $R = \sqrt[3]{GM/4\pi^2 T^{2/3}}$. If the background universe were of marginally bound Einstein-de Sitter form, it would have $r_e(T) = \sqrt[3]{9GM/2} T^{2/3}$. It follows that the recently virialised object should have a density contrast $\rho/\bar{\rho} = (r_e/R)^3 = 18\pi^2 \simeq 200$.

- and we replace it by a sphere of the same mass with negative binding energy that will expand to some maximum radius and then recollapse
- at the time t_{\max} of maximum expansion the kinetic energy K was zero and the potential energy was $U(t_{\max}) \sim GM/r_{\max}$ with some coefficient determined by the shape
- the virial theorem tells us that, after it has collapsed and virialised, it will have $2K+U=0$
- with $E = K + U = U(t_{\max})$, and $K = -U/2$, this implies that the final binding energy must be

$$- U(t_{\text{vir}}) = 2U(t_{\max})$$

- so, since $U \propto 1/r$, it must have collapsed by a factor 2 in order to generate the kinetic energy required to stabilise itself against further collapse
- the equations of motion are

$$\text{exterior : } v^2 = 2GM/r$$

$$\text{interior : } v^2 = 2GM/r - \text{constant}$$

- the solution for the exterior is
 - $r_e = \alpha t^{2/3}$
 - with α a constant, which implies, for the velocity
 - $v = dr/dt = (2/3)\alpha t^{-1/3} \Rightarrow v^2 = (4/9)\alpha^2 t^{-2/3} = (4/9)\alpha^3/r$

- and which, with the equation of motion, implies $\alpha^3 = 9GM/2$ and so

$$r_e(t) = \sqrt[3]{9GM/2} t^{2/3}$$

- the solution for the interior is the cycloid

$$\begin{aligned} r &= R(1 - \cos \eta) \\ t &= (T/2\pi)(\eta - \sin \eta) \end{aligned}$$

- which we can verify by computing $v = dr/dt = (dr/d\eta)/(dt/d\eta) = (2\pi R/T) \sin \eta / (1 - \cos \eta)$

- which, with a little algebra, implies $v^2 = (2\pi R/T)^2 (2/(1 - \cos \eta) - 1)$, or

$$v^2 = 8\pi^2 R^3/T^2 r - 4\pi R^2/T^2$$

- and which, with the equation of motion, implies

$$R = \sqrt[3]{GM/4\pi^2} T^{2/3}$$

- comparing these we can estimate the density of a recently virialised object with respect to that of a critical density of the same age as

$$\rho/\bar{\rho} = (r_e(t=T)/R)^3 = 18\pi^2 \simeq 200$$

- while a crude model – it neglects completely the effect of dark energy, for instance – this is borne out by numerical simulations

- these show that, if we consider a sphere around a simulated cluster, or ‘halo’, within which the density contrast is about 200 then this delineates quite well the exterior ‘infall region’ from the virialised interior where we have multiple streams of matter

- it also turns out, for rich clusters, with velocity dispersions of about 1000km/sec, to be about the same as the ‘Abell radius’ ($1.5\text{Mpc}/h$) that George Abell arrived at empirically

1.7 Evolution of the cluster mass function

- the cluster mass- or X-ray luminosity-function has a form rather similar to the galaxy luminosity function with a ‘knee’, above which the number of clusters drops exponentially
- the evolution of the cluster mass function $n(M)dM$ depends on
 1. the initial seeds for structure
 2. how the structure grows (which depends on expansion history)

1.7.1 Self-similar evolution

- during the matter dominated era
 - which, it used to be thought continued up to the present day,

the ‘background’ cosmology is ‘scale invariant’: density, scale factor etc. just vary as power laws with time.
- and the primordial fluctuations are also approximately scale invariant
 - e.g. Gaussian random field with power-law power spectrum $\Delta_\rho^2(k) \equiv k^3 P_\rho(k) \propto k^{n+3}$
 - where k is a ‘co-moving’ wavenumber $\mathbf{k} = a(t)\mathbf{k}_{\text{phys}}$
 - and $n \simeq -1$ on relevant scales (spectral index varies slowly with wavelength)

the fact that (as we will see later) the density fluctuations grow with time as $\Delta\rho/\rho \propto a(t) \propto 1/(1+z)$ means that the wave-number (inverse comoving scale) of non-linearity grows like

$$k_* \propto (1+z)^{2/(n+3)}$$

- and hence the characteristic mass (knee of the mass function) varies as

$$M_* \propto k_*^{-3} \propto (1+z)^{-6/(n+3)}$$

- which is quite a strong rate of evolution

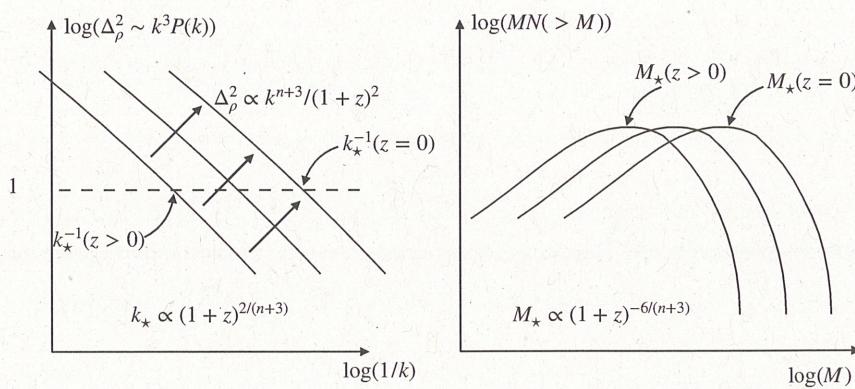


Figure 17: Self-similar model for evolution of the cluster mass function. Left panel shows the evolution of the power spectrum of density perturbations for (nearly) scale invariant initial conditions. In an Einstein-de Sitter model this evolves with time as $(1+z)^2$ so the ‘characteristic scale of non-linearity’ k_*^{-1} increases with time as a power-law also. Right hand panel shows mass function.

- the radius scales like the cube root of the mass divided by the density
- the latter scales like $\rho \propto (1+z)^3$ so the characteristic radius scales as $R_* \propto a/k_* \propto (1+z)^{-1}(1+z)^{-2/n+3}$
- or
- $R_* \propto (1+z)^{-(n+5)/(n+3)}$
- and the characteristic temperature (from hydrostatic equilibrium) goes like $T_* \propto M_*/R_*$ or
- $T_* \propto (1+z)^{(n-1)/(n+3)}$
 - the idea here is that the clustering is growing ‘hierarchically’ with small halos merging into larger ones as the universe ages
 - this is a highly complex process – at the time these models were developed it was not possible to model this using hydrodynamical simulations. Even today this is challenging.
 - the beauty of the model is that the scale invariance of the initial fluctuations and the background within which these are evolving means that one can predict the population at one time from observations at another simply by scaling the physical quantities appropriately
- this model worked quite well, but not perfectly
 - understandable since lower-mass clusters were quite likely to have been affected by early energy ejection
 - gives the gas an entropy larger than would have arisen from shocking in the self-similar evolution
but even allowing for this (e.g. focussing on high-mass end of distribution function) there was a problem
- models predicted too rapid evolution
- this was an early indication of the need for dark-energy or cosmological constant (Λ CDM) (e.g. Pat Henry)
 - along with problems with the age of the stars vs. age of the universe

1.8 The Sunyaev-Zel'dovich effect

- if we take the observed X-ray luminosity L_X and temperature T_X (inferred from spectrum fitting) for a rich cluster and use $L_X \sim \epsilon r^3$ with thermal bremsstrahlung emissivity $\epsilon \propto n_e^2 \sqrt{T_X}$ then we find the density of electrons
 - $n_e \sim 3 \times 10^{-3} h^{-1/2} \text{ cm}^{-3} (r/200 \text{ kpc}/h)^{-3/2} (T_X/4 \text{ keV})^{-1/4} (L_X/10^{44} h^{-2} \text{ erg/sec})^{1/2}$
 - 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 $\sigma_T \simeq 6 \times 10^{-29} \text{ m}^2$ and the radius we get an estimate for the optical depth for electron scattering
 - $\tau_{\text{es}} \sim 3 \times 10^{-3}$
- 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 and Jacob B. Zel'dovich realised that, owing to the motion of the electrons doing the scattering, this would have very interesting observable effects

1.8.1 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 18

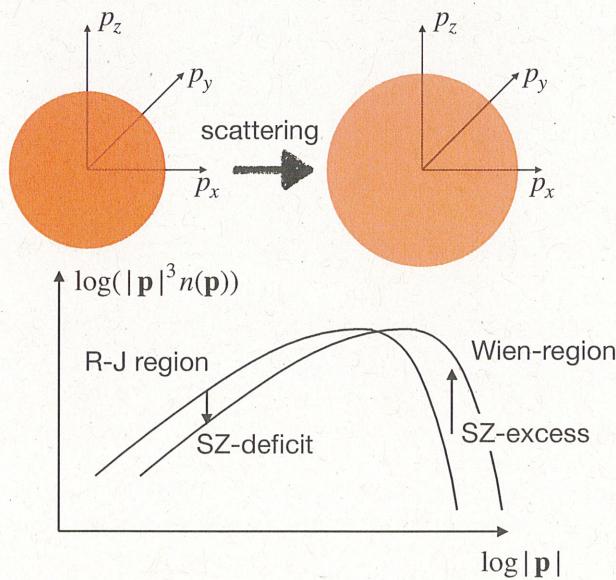


Figure 18: The thermal Sunyaev-Zel'dovich effect. Consider a box containing thermal radiation and hot plasma. Initially the photons have a Planck spectrum with occupation numbers $n(\mathbf{p}) = 1/(e^{\hbar\nu/kT} - 1)$. A fraction of these photons get scattered by the rapidly moving electrons – giving 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

1.8.2 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