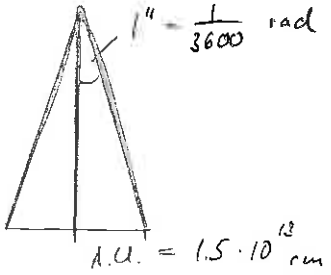
# The Physics of Galaxy Clusters, Christoph Pfrommer

## 0. Preliminaries:

- webpage: www.cita.utoronto.ca/~pfrommer/Lectures
- Literature:

- G.M. Voit, Tracing Cosmic Evolution with Clusters of Galaxies, 2005, RvMP, 77, 207
- J. Peacock, Cosmological Physics, Cambridge University Press
- P.J.E. Peebles, Principles of Physical Cosmology, " "
- T. Padmanabhan, Structure Formation in the Universe, " "
- M. Bartelmann, Lectures on "Cosmology": www.ita.uni-heidelberg.de/research/bartelmann/Lectures/cosmology/

- Units: cgs
  - pc:




$$1\,\mathrm{pc} = \frac{1\,\mathrm{AU}}{\tan\left(1'' \cdot \frac{\pi}{3600 \cdot 180}\right)} \simeq 2 \cdot 10^5\,\mathrm{A.U.} \simeq 3 \cdot 10^{18}\,\mathrm{cm}$$

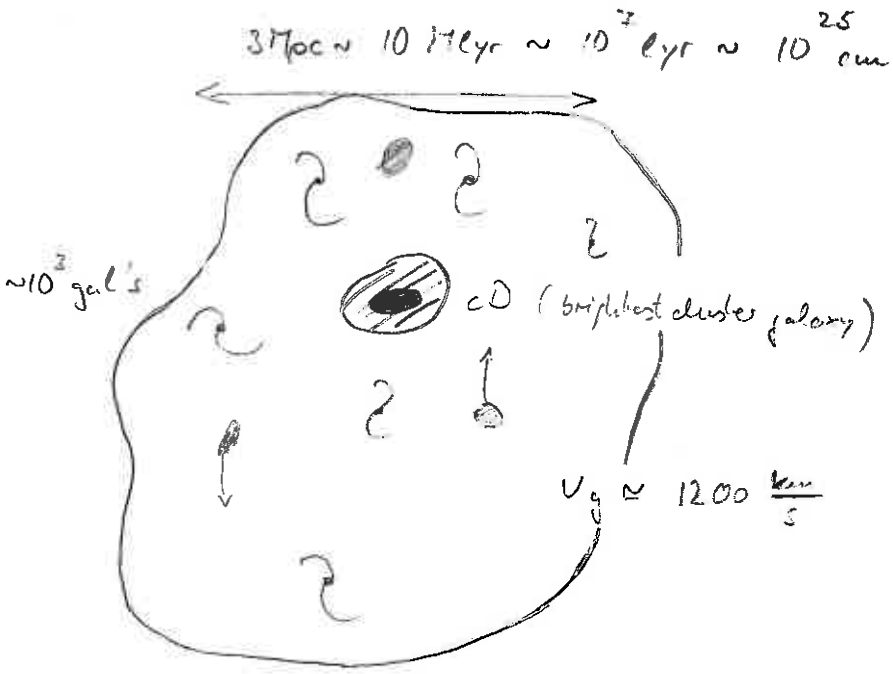
- erg: $\mathrm{erg} = \frac{\mathrm{gm\,cm^2}}{\mathrm{s^2}} = 10^{-7}\,\mathrm{J}$
- $1\,M_\odot \simeq 2 \cdot 10^{33}\,\mathrm{gm}$
- $10^{12}\,\mathrm{eV} = 1\,\mathrm{TeV} \simeq 1.6\,\mathrm{erg}$
- $1\,\mathrm{esu} = \frac{\mathrm{cm}^{3/2}\,\mathrm{gm}^{1/2}}{\mathrm{s}}$

$$\left[\; r_0 = \frac{e^2}{m_e c^2} = \frac{2.3 \cdot 10^{-19}\,\mathrm{cm^3\,gm}}{10^{-27}\,\mathrm{gm}\;10^{21}\,\frac{\mathrm{cm^2}}{\mathrm{s^2}}\;\mathrm{s^2}} = 3 \cdot 10^{-13}\,\mathrm{cm}\;\right]$$

# 1. Overview

## 1.1 What is a cluster of galaxies ?

### 1.1.1. Optical




- Virial Theorem:

$$2T + U = 0 \qquad \text{(in equilibrium)}$$

$$M_g v_g^2 - \frac{G M_c M_g}{R_c} = 0$$

$$M_c = \frac{R_c v_g^2}{G} \simeq \frac{10^{25}\,cm \cdot 1.4 \cdot 10^{14}\,\frac{cm^2}{s^2}}{\frac{2}{3}\,10^{-7}\,erg\,cm\,gm^{-2}}$$

$\sim M_c \simeq 2 \cdot 10^{48}$ gm $\simeq 10^{15} M_\odot$ gravitating mass of a cluster !

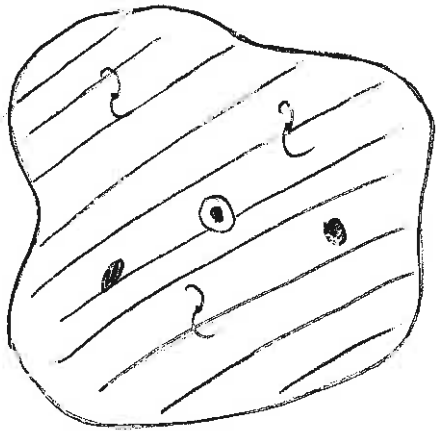
→ typical masses range from $(10^{14} - 10^{15}) M_\odot$

but $M_* \sim \frac{1}{50} M_c$ (Zwicky 1930s)

⇒ inferred existence of DM 80 yrs ago !

## 1.1.2. X-rays

- ) 1970s: X-ray astronomy

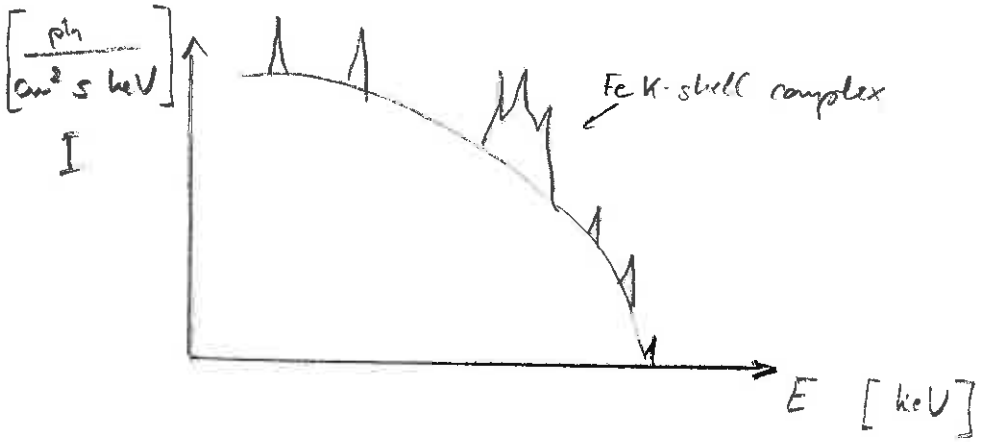


X-ray emission in between galaxies,

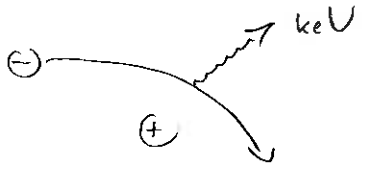
what is this ?

either bremsstrahlung emission from hot thermal electrons or inverse Compton emission from relativistic electrons ?

- ) X-ray spectrum:




→ bremsstrahlung continuum + line emission




BS: $j_X \propto n_e n_i \sqrt{kT}$

$\rightsquigarrow n \sim (10^{-1} - 10^{-3})\ cm^{-3}$

$T \sim (10^7 - 10^8)\ K$ => hot gas !

- talk about temperature of gas in terms of particle energies

$kT \propto (1 - 10)\ keV \propto (10^3 - 10^4)\ eV$

- at these temperatures, most of the elements are fully ionized, except for highly-ionized iron, e.g., hydrogen-like Fe which is an Fe nucleus + 1 bound electron : Fe XXVI

- line emission:

Fe XXV: $h\nu \sim Z(Z-1)\,Ry = 26 \cdot 25 \cdot 13.6\,eV \sim 8.8\,keV$

Fe XXVI: $h\nu \sim Z^2\,Ry \simeq 26^2 \cdot 13.6\,eV \sim 9\,keV$

=> the higher the temperature, the higher the ionization stage!

- hydrostatic masses for a 6 keV cluster:

$E \sim U$

$\frac{3}{2} kT \sim \mu m_p \frac{G M_c}{R_c}$ (assuming monoatomic gas)

- $\mu \simeq 0.6$ mean molecular weight (next sheet)

$$M_c = \frac{3\,kT\,R_c}{2\mu\,m_p\,G} \simeq \frac{3 \cdot 1.6 \cdot 0.6 \cdot 10^{-8}\,erg \;\; 10^{25}\,cm}{2 \cdot 0.6 \cdot 1.7 \cdot 10^{-24}\,gm \;\; \frac{2}{3} \cdot 10^{-7}\,erg\,cm\,gm^{-2}}$$

$$\simeq 2 \cdot 10^{48}\,gm \sim 10^{15}\,M_\odot$$

- X-ray imaging of hot gas: $M_{gas} \sim \frac{1}{7} M_c$

~ we found some of the matter that was "dark" in the optical by looking at a different waveband

~ rest cannot be seen in any other waveband (at least no significant amounts)

=> dark matter, which reflects our ignorance of the composition.

- inventory of cluster mass:

~ 2% stars in galaxies
~ 13% hot gas (1–10 keV)
~ 85% dark matter

} cosmic baryon fraction ~ 16.6%
(slight mismatch points to interesting physics, including non-gravitational energy input from supernovae + black holes)

·) Primordial element composition :

$n_H \equiv \frac{\rho}{m_p} X_H \qquad X_H = 0.76$ mass fraction of hydrogen

$n_{He} \equiv \frac{\rho}{m_p} \left( \frac{1 - X_H}{4} \right)$

- number density of nuclei :

$$n_{nuclei} = n_H + n_{He} = \frac{\rho}{m_p} \cdot \frac{4X + 1 - X}{4} = \frac{\rho}{m_p} \frac{3X + 1}{4} \simeq \frac{\rho}{m_p} \cdot 0.82$$

- number density of electrons : employing charge conservation ∿

$$n_e = n_H + 2 n_{He} = \frac{\rho}{m_p} \frac{X + 1}{2} = \frac{\rho}{m_p} \cdot 0.88 \quad \text{(fully ionized)}$$

$x_e \equiv \frac{n_e}{n_H}$ ∿ $n_e = \frac{\rho}{m_p} X_H x_e$ (partially ionized) $\left[ x_e = \frac{X+1}{2X} = 1.157 \text{ „fully ionized"} \right]$

- mean molecular weight $\mu$ :

$n \equiv \frac{\rho}{\mu m_p}$ , $\quad n = n_e + n_{ion} = n_e + n_H + n_{He}$

$$\mu = \frac{\rho}{m_p n} = \frac{n_H + 4 n_{He}}{n_H + n_{He} + n_e}$$

$$\mu = \begin{cases} \frac{4}{5X_H + 3} = 0.588 & \text{(fully ionized)} \\ \frac{4}{3X_H + 1 + 4X_H x_e} & \text{(partially ionized)} = 1.22 \text{ (not ionized, } x_e = 0) \end{cases}$$

derivation :

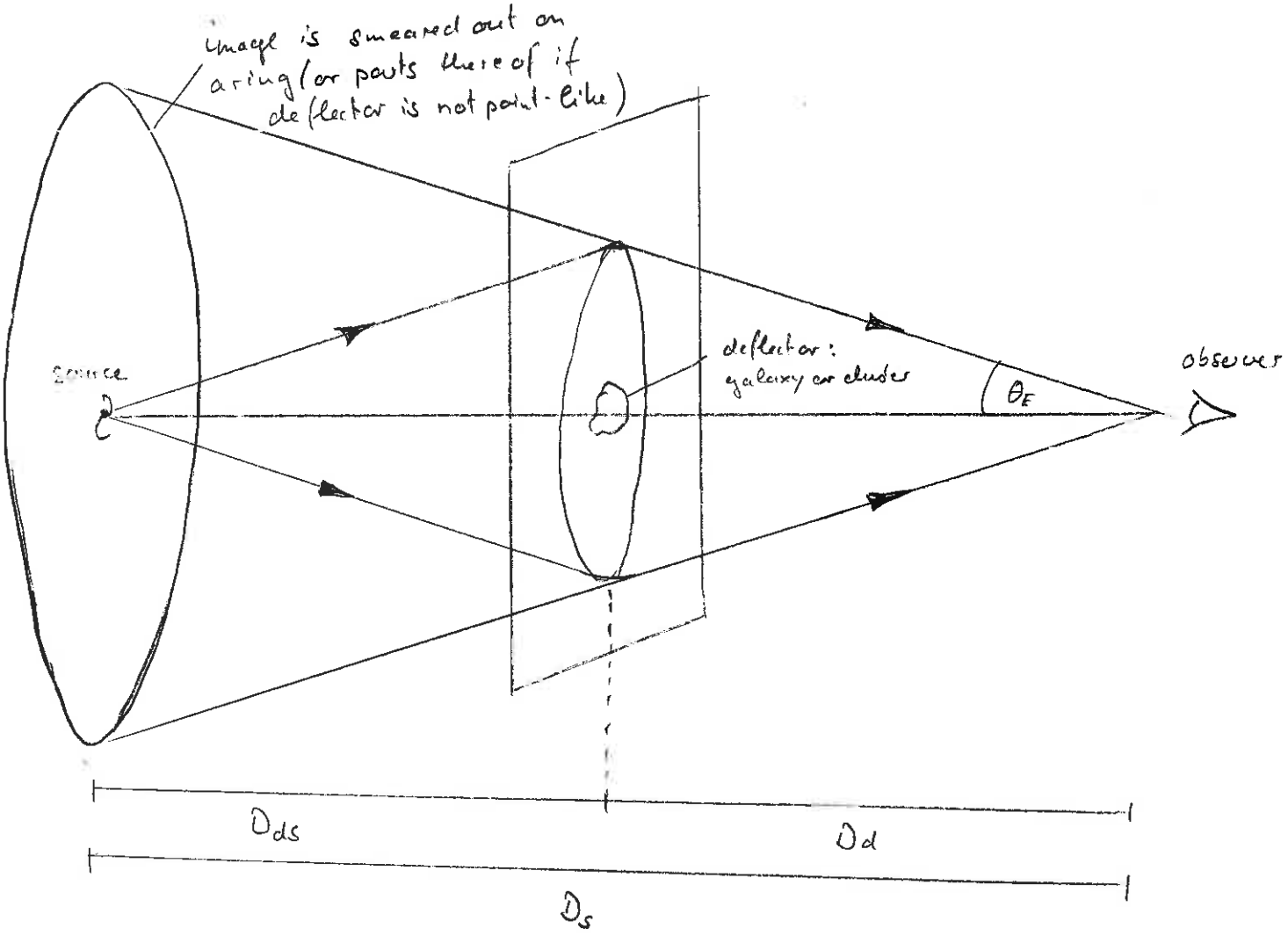
$$\mu = \frac{\left[ X_H + (1 - X_H) \right] \rho}{\left[ X_H + \frac{1 - X_H}{4} + \frac{X_H + 1}{2} \right] \rho} = \frac{4}{4X_H + 1 - X_H + 2X_H + 2} = \frac{4}{5X_H + 3} \quad \text{(fully ionized)}$$

$$\mu = \frac{(X_H + 1 - X_H)\rho}{\left( X_H + \frac{1 - X_H}{4} + x_e X_H \right) \rho} = \frac{4}{4X_H + 1 - X_H + 4x_e X_H} = \frac{4}{3X_H + 1 + 4X_H x_e} \quad \text{(partially ionized)}$$

- number of target nucleons for hadronic reactions (p-p) or bremsstrahlung emission :

$$n_N = n_H + 4 n_{He} = \frac{\rho}{m_p} \left( X_H + 4 \cdot \frac{1 - X_H}{4} \right) = \frac{\rho}{m_p}$$

## 1.1.3 Gravitational lensing




(later, in a lecture on lensing, we will derive

$$\theta_E = \left[ \frac{4GM(\theta_E)}{c^2} \frac{D_{ds}}{D_d D_s} \right]^{1/2}$$ Einstein radius

$$\simeq 3'' \left( \frac{M}{10^{12} M_\odot} \right)^{1/2} \left( \frac{D}{Gpc} \right)^{-1/2},$$ galaxy lens: point-lens approx. Ok

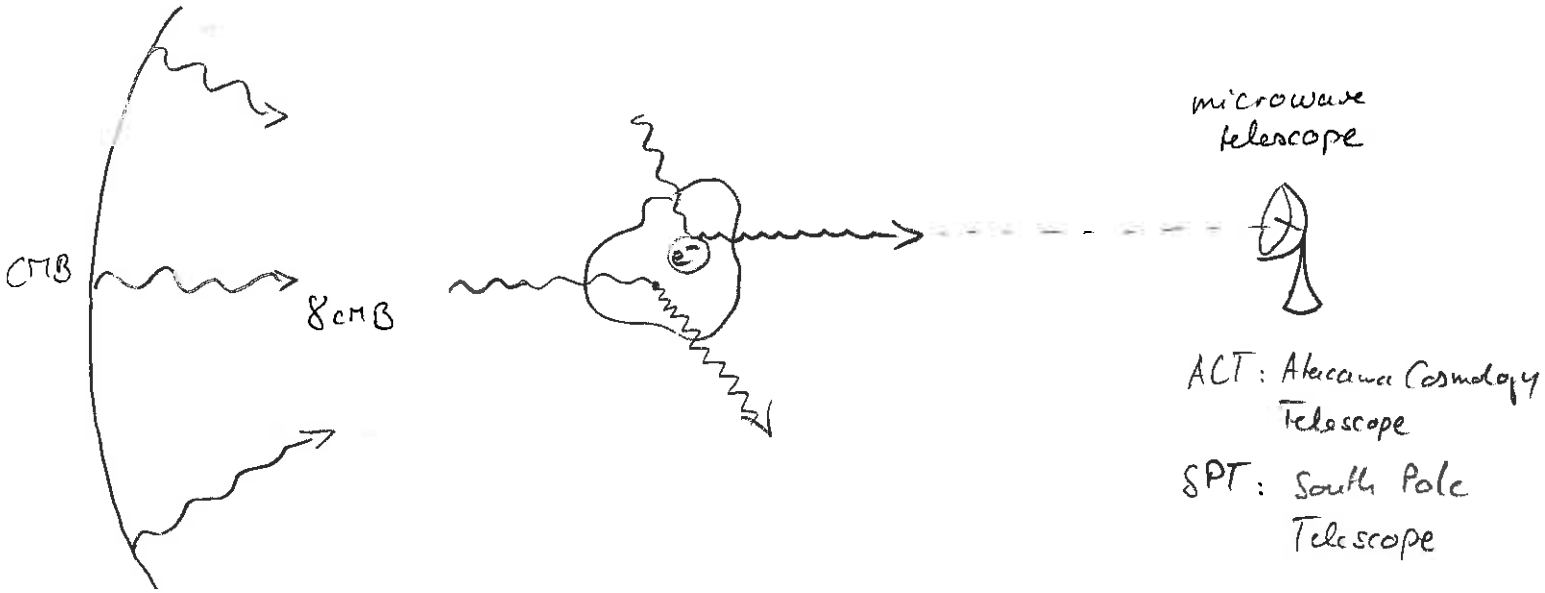
$$\simeq 30'' \left( \frac{M}{10^{14} M_\odot} \right)^{1/2} \left( \frac{D}{Gpc} \right)^{-1/2},$$ $L_d > L_s$: detailed mass-modelling necessary, only fraction of cluster mass within Einstein radius

→ corresponds to typical angular scales of observed giant (tangential) arcs

→ <u>strong lensing</u>: sensitive to <u>projected mass within $\theta_E$</u> → radial arcs

<u>weak lensing</u>: weaker distortion of gal. image in tangential direction
≈ need averaging over neighboring gal.'s to measure signal!

## 1.1.4 Sunyaev – Zel'dovich effect (SZE)

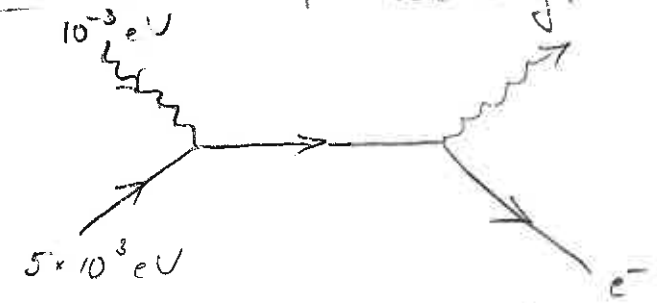



ACT: Atacama Cosmology Telescope

SPT: South Pole Telescope

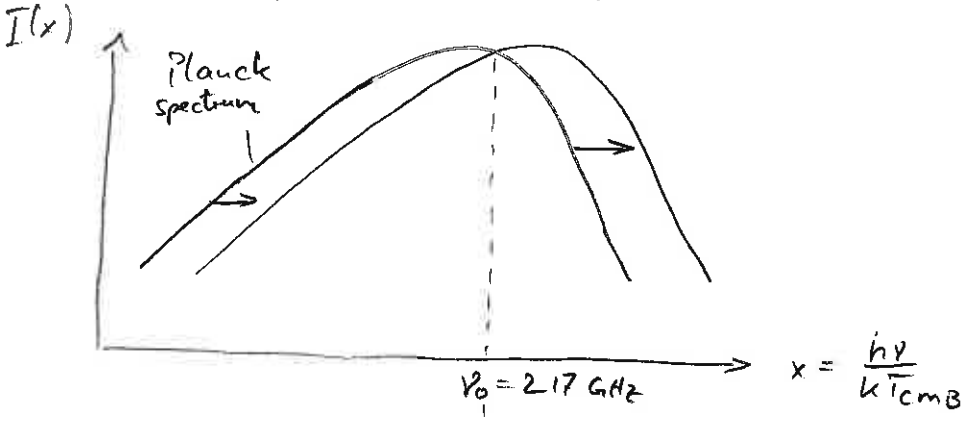
**CMB:** radiation left behind from the early Universe, when hydrogen recombined.

**SZE:** Inverse Compton scattering:




→ mean transfer of energy from "hot" electrons to "cold" photons

→ number of CMB photons remains unchanged (elastic scattering), but CMB spectrum is distorted




reduction of flux → clusters burn "holes" in the CMB for $\nu < \nu_0$

excess flux

→ how many CMB photons get scattered ?

• optical depth $\tau \sim n \sigma_T L$

$$\sigma_T \sim 2\pi r_0^2 \sim 2\pi \left(\frac{e^2}{m_e c^2}\right)^2 \sim 6 \cdot \left[\frac{(4.8 \cdot 10^{-10})^2}{10^{-27}\ 10^{21}}\right] cm^2$$

$$\sim 6 \cdot (3 \cdot 10^{-13})^2 cm^2 \sim 6 \cdot 10^{-25} cm^2$$

$$\tau \sim n \sigma_T L \sim 10^{-4} cm^{-3} \cdot 6 \cdot 10^{-25} cm^2 \cdot 10^{25} cm \sim 6 \cdot 10^{-4} \ll 1$$

• amplitude of the SZ effect : Compton-y parameter

$$y = \left(\begin{array}{c}\text{typical energy gain}\\ \text{by a photon}\end{array}\right) \times \left(\begin{array}{c}\text{probability that a}\\ \text{photon get scattered}\end{array}\right)$$

$$\sim \frac{k T_e}{m_e c^2} \times n_e \sigma_T L \sim 10^{-2} \cdot 10^{-4} cm^{-3} \cdot 6 \cdot 10^{-25} cm^2 \cdot 10^{25}$$

$\Rightarrow y \sim 6 \cdot 10^{-6}$ ~ small change in intensity ~ difficult to detect !

## 1.1.5. Synthesis

Different observables have different strengths and weaknesses

→ complementarity : combine observations to learn more about clusters !

X-rays:
- good for cluster centers $f_X \sim n^2$ (emphasizes dense cluster gas)
- high-resolution observations possible (Bragg-reflexion)
- bad for clusters at large distances as $F_X \sim \frac{L_X}{D^2}$
- watch out for clumping $C = \frac{\langle n^2 \rangle}{\langle n \rangle^2}$ that biases emission

SZE:
- good for cluster outskirts as $y \sim P_{th} \sim n_e kT$
- good for clusters at large distances : photons are continuously redshifted from the surface-of-last-scattering to us as the Universe expands (independent of when the IC-scattering event happened)
- small signal, difficult to detect ;

optical: – collisionless tracers of gravitational potential and its dynamical state → but needs many gal's to sample the velocity distribution well enough (only possible for med. to large clusters)

– lensing: great for directly probing total mass (dominated by DM) → but problem: projection effects along line-of-sight

## 1.6. Relation to average Universe around us

- critical density of the Universe:

$$\rho_{cr} = \frac{3 H_0^2}{8\pi G} \qquad \left[ \text{local Hubble law } v = H_0 d \right]$$

- $\Omega_m = \frac{\bar{\rho}_m}{\rho_{cr}} \simeq 0.27$, $\Omega_b = \frac{\bar{\rho}_b}{\rho_{cr}} \simeq 0.045$

- $\rho_m \simeq 4 \cdot 10^{10} \frac{M_\odot}{Mpc^3} \simeq 10^9 \frac{M_\odot}{Mly r^3}$ cf. $M_c \sim 10^{15} M_\odot$

⇒ clusters are extremely rare!

⇒ Sum up universe in chunks of $10^{15} M_\odot$, only less than 1% formed an aggregation that makes a big cluster

⇒ collapse of a cluster:

$\bar{\rho}_{cluster} \sim 10^3 \bar{\rho}_m$ ~ radial collapse factor ~ 10

$\bar{n}_b = \frac{\rho_{cr} \Omega_b}{\mu m_p} \sim \frac{10^{-29} \text{ gm cm}^{-3} \; 0.045}{0.6 \cdot 1.7 \cdot 10^{-24} \text{ gm}} \sim 4 \cdot 10^{-7} \text{ cm}^{-3}$

$\xrightarrow{\text{collapse}}$ $4 \cdot 10^{-4} \text{ cm}^{-3}$ (consistent w/ X-ray densities)

# 1.2. Geometry and Dynamics

## 1.2.1. Assumptions

- Basis are the 2 cosmological assumptions:

  (1) When averaged over sufficiently large scales, the observable properties of the Universe are isotropic, i.e., independent of direction.

  (2) Our position in the Universe is not preferred to any other (cosmological principle).

- Remarks:
  - to 1: what are sufficiently large scales? -> nearby galaxy distribution is very anisotropic, only on scales $\gg 100$ Mpc their distribution approaches isotropy. The microwave background is almost perfectly isotropic with a level of anisotropy of $\sim 10^{-5}$.
  - to 2: By assumption 2, the first assumption must be true for every observer in the Universe: if the Universe is isotropic around every point, it's also homogeneous => one often loosely states these assumptions as "the Universe is homogeneous and isotropic"

- We have four interactions in physics:
  - weak + strong are limited to length scales of these elementary interactions
  - electromagnetic is limited to charge shielding scales. In a plasma, this is the skin depth $\lambda_p = \frac{c}{\omega_p} = \sqrt{\frac{m_e c^2}{4\pi e^2 n_e}} \simeq 10^8 \text{ cm} \times \left(\frac{n_e}{2 \cdot 10^{-7} \text{ cm}^{-3}}\right)^{-1/2}$, but magnetic fields can be coherent across large scales and affect the dynamics on these
  - gravity is only relevant force on cosmological scales, it is described by General Relativity. In contrast Newtonian dynamics is constructed for isolated bodies and difficult to describe continuous fields

- General Relativity describes space-time as four-dimensional manifold whose metric tensor $g_{\mu\nu}$ is a dynamical field; its dynamics is governed by Einstein's field equations which couple metric to matter-energy content of space-time. GR is non-linear: space-time structure determines motion of matter and energy, which determine the structure of space-time and so on.

## 1.2.2 Metric

- eigentime element or line element $ds$ (that defines how to measure lengths):

$$ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu \quad , \mu,\nu = 0,1,2,3$$

0 denotes time dimension
$i = 1,2,3$ denote space dimensions

- due to symmetry, the $4\times4$ tensor $g_{\mu\nu}$ has 10 independent components:
1 time-time, 3 space-time, 6 space-space

- define fundamental observers: observers that follow mean motion (of matter and energy) with respect to which all observable properties are isotropic.
→ they imply so-called comoving coordinates (spatial): $dx^i = 0$
which requires their eigentime to be equal the coordinate time $dt$

$$ds^2 = g_{00}\, dt^2 = c^2 dt^2 \Rightarrow g_{00} = c^2$$

- isotropy requires synchronisation of clocks such that $g_{0i} = 0$; otherwise the components $g_{0i}$ could single out a preferred direction in space violating isotropy

$$g_{0i} = 0$$

$$\Rightarrow \quad ds^2 = c^2 dt^2 + g_{ij}\, dx^i dx^j$$

$\sim$ decomposition of space-time into spatial hypersurfaces of constant time!

- Without violating isotropy and homogeneity, spatial hypersurfaces can be scaled by a function $a(t)$ that can only depend on time (otherwise the expansion would be different at different places)

$$ds^2 = c^2 dt^2 - a^2(t)\, dl^2$$

- $dl^2$ is line element of homogeneous and isotropic 3-space

- isotropy requires 3-space to have spherical symmetry → introduce polar coordinates $(w, \varphi, \vartheta)$ where $w$ is radial coordinate, $(\varphi, \vartheta)$ are the usual angles:

$$dl^2 = dw^2 + f_k^2(w)\left(d\vartheta^2 + \sin^2\vartheta\, d\varphi^2\right) = dw^2 + f_k^2(w)\, d\omega^2,$$

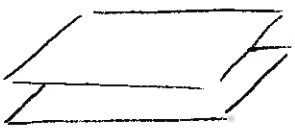
$f_K(w)$ radial function permitted because relation between $w$ and the area of spheres of constant $w$ is still arbitrary, $d\omega$ is the solid angle element

- homogeneity requires $f_k(w)$ to be either trigonometric, hyperbolic, or linear in $w$:

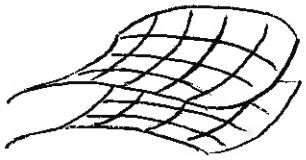
$$f_k(w) = \begin{cases} K^{-1/2} \sin\left(K^{1/2} w\right) & (K>0) \\ w & (K=0) \\ |K|^{-1/2} \sinh\left(|K|^{1/2} w\right) & (K<0) \end{cases}$$

$K$ = const. and parametrizes the curvature of spatial hypersurfaces,
$f_k(w)$ and $|K|^{-1/2}$ have dimension of a length:

flat:



negatively curved:



"Saddle point geometry"

positively curved:

$\Rightarrow$ metric for homogeneous and isotropic universe (Robertson-Walker metric):

$$ds^2 = c^2 dt^2 - a^2(t)\left(dw^2 + f_k^2(w)\, d\omega^2\right)$$

## 1.2.3 Redshift

- depending on the scale function $a(t)$, spatial hypersurfaces can expand or shrink $\Rightarrow$ red- or blueshift of photons expanding through space-time
- consider light that is emitted from comoving source at time $t_e$ and reaches a comoving observer at $w = 0$ at $t_0$. For light, $ds = 0$ (this is a consequence of special relativity and follows from the Minkowsky metric and the constancy of the light speed). Hence we have,

$$c|dt| = a(t)\,dw$$ where the modulus depends on the definition of the direction/sign of the measurement

- the coordinate distance between source and observer is

$$w_{e0} = \int_{t_e}^{t_0} dw = \int_{t_e}^{t_0} \frac{c\,dt}{a(t)} = \text{const.} \quad \text{(by definition)}$$

- Hence the time derivative of $w_{e0}$ w/r to $t_e$ must vanish:

$$\frac{dw_{e0}}{dt_e} = \frac{c}{a(t_0)}\frac{dt_0}{dt_e} - \frac{c}{a(t_e)} = 0 \Rightarrow \frac{dt_0}{dt_e} = \frac{a_0}{a_e}$$

- define $\nu^{-1} = dt$ the frequency of a light wave

$$\frac{\nu_e}{\nu_0} = \frac{\lambda_0}{\lambda_e} = 1 + \frac{\lambda_0 - \lambda_e}{\lambda_e} = 1 + z = \frac{a(t_0)}{a(t_e)}$$

where we introduced the redshift $z = \frac{\lambda_0}{\lambda_e} - 1$, i.e. the relative change in photon wavelength. We see, that wavelengths are continuously stretched as the Universe (and the coordinates on spatial hypersurfaces) expands!

- typically we chose $a(t_0) \equiv 1$, $a(t_e) \equiv a$

$$\Rightarrow \quad a = \frac{1}{1+z}$$

## 1.2.4 Dynamics

- the great degree of symmetry that we imposed upon the metric reduces its dynamics to that of the scale factor $a(t)$
- Einstein's field equations read

$$G_{\alpha\beta} = \frac{8\pi G}{c^2} T_{\alpha\beta} + \Lambda g_{\alpha\beta}$$

  - $\Lambda$ is the cosmological constant introduced by Einstein to allow for static cosmol. models.
  - $G_{\alpha\beta}$ : Einstein tensor, constructed from curvature tensor, depends on metric, its first and second derivatives
  - $T_{\alpha\beta} = \left(\rho + \frac{p}{c^2}\right) u_\alpha u_\beta - p\, g_{\alpha\beta}$ energy-momentum tensor of a perfect fluid characterized by pressure $p$ and (energy) density $\rho$, which can only be functions of time because of homogeneity, $p = p(t)$ and $\rho = \rho(t)$.

$\Rightarrow$ Friedmann's equations (proof in GR books):

$$(\mathrm{I}) \quad \left(\frac{\dot a}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{kc^2}{a^2} + \frac{\Lambda}{3}$$

$$(\mathrm{II}) \quad \frac{\ddot a}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3}$$

- remarks: Newton's analogue: (I) contains first derivatives of distance $\dot a^2 \sim v^2$ $\sim$ integral of motion, i.e., energy equation
  (II) $\ddot a \sim$ force eqns $\sim$ equation of motions
- We want to derive the adiabatic equation by combining those 2 equations:

$$(\mathrm{I}): \quad \frac{\dot a^2 + kc^2}{a^2} - \frac{\Lambda}{3} = \frac{8\pi G}{3}\rho$$

$$(\mathrm{II}): \quad 2\frac{\ddot a}{a} = -\frac{8\pi G}{3}\rho - 2\cdot 4\pi G\frac{p}{c^2} + 2\frac{\Lambda}{3}$$

$$(\mathrm{I})+(\mathrm{II}): \quad \frac{2a\ddot a + \dot a^2 + kc^2}{a^2} - \Lambda = -8\pi G\frac{p}{c^2}$$

now differentiate (I) by $t$ and multiply by $\frac{1}{8\pi G}$; multiply (I)+(II) by $\left(-\frac{3\dot{a}}{8\pi G a}\right)$:

$$\frac{3}{8\pi G}\left(\frac{2\dot{a}\ddot{a}}{a^2} - \frac{2(\dot{a}^2+kc^2)}{a^3}\right) = \dot{\rho}$$

$$-\frac{3}{8\pi G}\left(\frac{2\dot{a}\ddot{a}}{a^2} - \frac{\dot{a}^3}{a^3} - \frac{kc^2\dot{a}}{a^2}\right) + \frac{3\dot{a}\Lambda}{8\pi G a} = \frac{3p}{c^2}\frac{\dot{a}}{a}$$

combine:

$$\dot{\rho} + \frac{3p}{c^2}\frac{\dot{a}}{a} = -\frac{3}{8\pi G}\frac{\dot{a}}{a}\left(\frac{3\dot{a}^2}{a^2} + \frac{3kc^2}{a^2} - \Lambda\right) \overset{(I)}{=} -3\rho\frac{\dot{a}}{a}$$

$$\rightsquigarrow \dot{\rho}a^3 + 3\rho a^2\dot{a} + 3\frac{p}{c^2}a^2\dot{a} = 0$$

- We arrive at the "adiabatic expansion equation" of the Universe:

$$(III)\quad \frac{d}{dt}\left(a^3\rho c^2\right) + \frac{p}{c^2}\frac{d}{dt}\left(a^3\right) = 0$$

which intuitively states energy conservation; $E \equiv a^3\rho c^2$ is energy contained in fixed comoving volume and $a^3 = V$ is the proper volume:

$$dE = TdS - pdV = -pdV \quad \text{for adiabatic processes } \left(dS = \frac{dQ}{T} = 0\right).$$

- remarks:
  - $dS \overset{!}{=} 0$, otherwise there would be a heat flow into a direction, violating isotropy!
  - any two of the equations (I), (II) and (III) are linearly independent
  - one can derive this equation directly from the conservation equation $T^{\alpha\beta}{}_{;\beta} = 0$ which follows from the contracted Bianchi identities $\left(G^{\alpha\beta} - \Lambda g^{\alpha\beta}\right)_{;\beta} \equiv 0$, hence this is the "energy conservation" (in the absence of $\Lambda$).

# 1.3. Parameters, Age and Distances

## 1.3.1 Forms of Matter

- two forms of matter can be broadly distinguished: relativistic and non-relativistic (radiation and dust, respectively).
- for rel. bosons and fermions, $p = \frac{\varepsilon c^2}{3}$

  non-relativistic matter is approximately pressure-free, $p = 0$, because $p \ll \varepsilon c^2$
- non-rel. matter:

$$(\text{III}): \frac{d}{dt}\left(a^3 \varepsilon c^2\right) = 0 \quad \rightsquigarrow \quad \frac{\dot{\varepsilon}}{\varepsilon} = -3\frac{\dot{a}}{a}$$

$$\Rightarrow \quad \underline{\varepsilon(t) = \varepsilon_0 \cdot a^{-3}}$$

, with $\varepsilon_0$ today, it reflects that density of non-rel. matter is diluted as volume (of expanding space)

- relativistic matter:

$$(\text{III}). \quad \frac{d}{dt}\left(a^3 \varepsilon c^2\right) = \frac{\varepsilon c^2}{3}\frac{d}{dt}\left(a^3\right) = 0$$

$$3a^2\dot{a}\varepsilon + a^3\dot{\varepsilon} - \frac{\varepsilon}{3}3a^2\dot{a} = 0$$

$$\frac{4\dot{a}\varepsilon}{a} + \dot{\varepsilon} = 0 \quad \rightarrow \quad \frac{\dot{\varepsilon}}{\varepsilon} = -4\frac{\dot{a}}{a}$$

$$\Rightarrow \quad \underline{\varepsilon(t) = \varepsilon_0 a^{-4}}$$

the density of rel. matter drops faster than that of non-rel. matter by one more power of $a$ because particles are diluted and lose energy as they are redshifted.

## 1.3.2. Parameters

- introduce Hubble parameter $H(t) \equiv \frac{\dot{a}}{a}$, $H_0 \equiv H(t_0)$

  $H_0 = 100\, h\, \frac{\text{km}}{\text{s Mpc}}$ expansion rate of Universe with $h \simeq 0.70$
- Hubble time:

$$t_H = \frac{1}{H_0} = 3.1 \cdot 10^{17}\, h^{-1}\, \text{s} = 9.8 \cdot 10^8\, h^{-1}\, \text{yr} \simeq 1.4 \cdot 10^{10}\, \text{yr} \simeq 14\ \text{Gyr}$$

- Hubble radius:

$$r_H = \frac{c}{H_0} = 9.3 \cdot 10^{27}\, h^{-1}\, cm = 3 \cdot 10^3\, h^{-1}\, Mpc \simeq 4.3\ Gpc$$

- critical density:

$$\rho_{cr}(t) = \frac{3H^2(t)}{8\pi G}, \quad \rho_{cr,0} = \frac{3H_0^2}{8\pi G} = 1.9 \cdot 10^{-29}\, h^2\, \frac{gm}{cm^3} \simeq 10^{-29}\, \frac{gm}{cm^3}$$

$$\simeq 1.4 \cdot 10^{11}\, \frac{M_\odot}{Mpc^3}$$

writing it in the form

$$\frac{4\pi G}{3}\left(\frac{\rho_{cr}\, a^3}{a}\right) = \frac{\dot{a}^2}{2}$$ illustrates that a sphere filled with matter at critical density produces a gravitational potential that is exactly balanced by the specific kinetic energy.

- dimensionless density parameters:

$$\Omega(t) \equiv \frac{\rho(t)}{\rho_{cr}(t)}, \qquad \Omega_0 \equiv \Omega(t_0) = \frac{\rho(t_0)}{\rho_{cr,0}}$$

$$\Omega_\Lambda(t) = \frac{\Lambda}{3H^2(t)}, \qquad \Omega_{\Lambda,0} = \Omega_\Lambda(t_0) = \frac{\Lambda}{3H_0^2}$$

matter $\Omega_m = \frac{\rho_m}{\rho_{cr}}$ : $\rho_m(t) = \Omega_{m,0}\, \rho_{cr,0}\, a^{-3}$

radiation $\Omega_r = \frac{\rho_r}{\rho_{cr}}$ : $\rho_r(t) = \Omega_{r,0}\, \rho_{cr,0}\, a^{-4}$

- replacing $\rho \to (\rho_r + \rho_m)$ in the Friedmann's equation yields

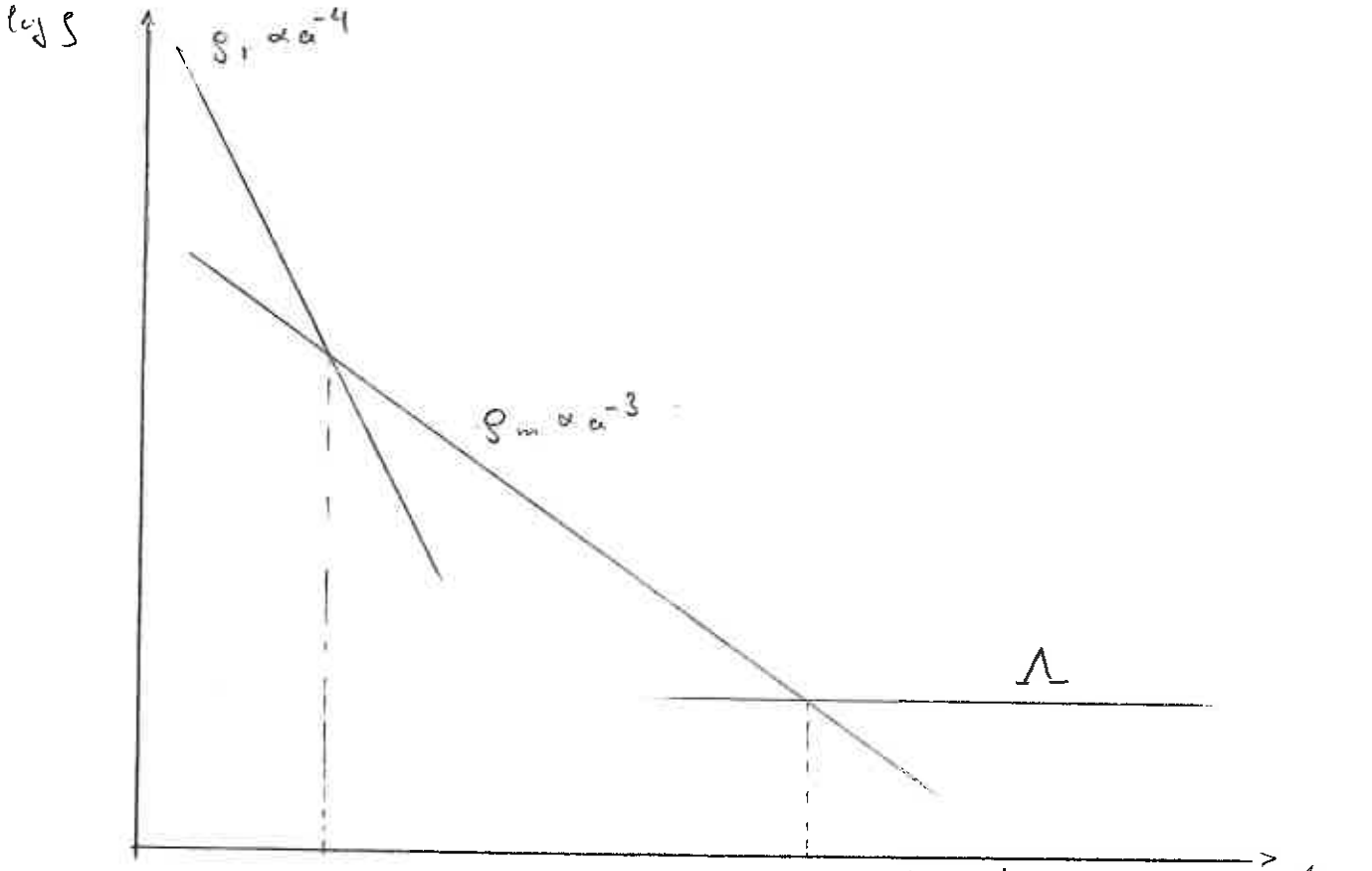
$$H^2(a) = H_0^2\left(\Omega_{r0}\, a^{-4} + \Omega_{m0}\, a^{-3} + \Omega_{\Lambda 0} - \frac{Kc^2}{a^2 H_0^2}\right)$$

- for $a = 1$, left-hand side $H^2(a_0) = H_0^2$ and solving for K-term:

$$-\frac{Kc^2}{H_0^2} = 1 - \Omega_{r0} - \Omega_{m0} - \Omega_{\Lambda 0} = \Omega_K$$ as the curvature parameter

$\Rightarrow$ final form for Friedmann's equation:

$$H^2(a) = H_0^2\left(\Omega_{r0}\, a^{-4} + \Omega_{m0}\, a^{-3} + \Omega_{\Lambda 0} + \Omega_K\, a^{-2}\right) \equiv H_0^2\, E^2(a)$$




$a_{eq} = \frac{\Omega_{r,0}}{\Omega_{m,0}} \approx 2.4 \cdot 10^{-4}$ $\quad a_{eq,\Lambda} = \left(\frac{\Omega_{m0}}{\Omega_{\Lambda 0}}\right)^{1/3} \sim 0.72$

$\rightsquigarrow z_{eq} \approx 4165$ $\quad z_{eq,\Lambda} \approx \frac{1}{a_{eq,\Lambda}} - 1 \approx 0.4$

- $\Omega_{r0} = \Omega_{cmb,0} + 3\Omega_{\nu 0} = 3.2 \cdot 10^{-5} h^{-2}$ $\sim 6.5 \cdot 10^{-5}$ $\sim$ ignore radiation at late times

  $\Omega_{m0} \approx 0.27$

  $\Omega_{\Lambda 0} \approx 0.73$

- $\Omega_m(a) = \frac{8\pi G}{3H^2(a)} \rho_{m0} a^{-3} = \frac{\Omega_{m,0}}{a + \Omega_{m0}(1-a) + \Omega_{\Lambda 0}(a^3 - a)}$

  $\Omega_\Lambda(a) = \frac{\Lambda}{3H^2(a)} = \frac{\Omega_{\Lambda 0} a^3}{a + \Omega_{m0}(1-a) + \Omega_{\Lambda 0}(a^3 - a)}$

- for $a \to 0$ : we find $\Omega_m \to 1$ and $\Omega_\Lambda \to 0$ irrespective of their values

## 1.3.4 Age and Expansion of the Universe

- Since $H = \frac{\dot{a}}{a}$ :

$$\frac{da}{dt} = H_0 a E(a) \Rightarrow H_0 t = \int_0^a \frac{da'}{a' E(a')} \simeq 13.7\ \text{Gyr} \quad \left[\text{solve numerically in general}\right]$$

- Limiting cases:

1) early Universe : radiation dominates, $\rho_r \propto a^{-4}$ : $E(a) = \Omega_{r0}^{1/2} a^{-2}$

$$H_0 t = \frac{a^2}{2\sqrt{\Omega_{r0}}} \iff a = \left[2\sqrt{\Omega_{r0}}\, H_0 t\right]^{1/2}$$

$\rightsquigarrow$ at early times, expansion of Universe scales like $a \propto \sqrt{t}$ as in RD

$$t_{eq} = 5 \cdot 10^{4}\ \text{yr}$$

2) early matter-dominated era : after non-relativistic matter starts dominating, before curvature kicks in, we may approx. $E(a) \simeq \sqrt{\Omega_{m0}}\, a^{-3/2}$

$$H_0 t = \frac{2 a^{3/2}}{3\sqrt{\Omega_{m0}}} \iff a = \left[\frac{3}{2}\sqrt{\Omega_{m0}}\, H_0 t\right]^{2/3}$$

expansion scales faster than in the rad.-dominating era with $a \propto t^{2/3}$

(also called Einstein-de Sitter limit)

3) At late times, if $\Omega_\Lambda \neq 0$ then it starts to dominate and $E(a) = \sqrt{\Omega_\Lambda}$

$$H_0 t = \frac{\ln a}{\sqrt{\Omega_\Lambda}} \Rightarrow a \propto e^{\sqrt{\Omega_\Lambda}\, H_0 t}$$

(where we ignored the lower integration limit as $\Lambda$ becomes only important at late times). The cosmol. constant is driving the Universe exponentially apart (the so-called de-Sitter limit).

Note, this is owing to the constant energy density of $\Lambda$ which implies an increasing energy as spacial hypersurfaces expand!

4) Einstein-de Sitter Universe : $\Omega_{\Lambda 0} = 0$ and $\Omega_{m0} = 1$, then

$$H_0 t = \frac{2}{3} a^{3/2} \iff a = \left(\frac{3}{2} H_0 t\right)^{2/3}$$

## 1.3.5 Distances

- distance measures are no longer unique in general relativity; in Euclidean geometry, a distance between two points is measured at the same time which is impossible to achieve in GR for two reasons:
  1) simultaneous measurements in GR depend on relative motions between two points
  2) connecting two points requires time because of finite speed of light

  $\Rightarrow$ introduce different definitions according to special requirements sought!
- emission and observation time $t_2$ and $t_1$, respectively, are related to scale factors $a_2$ and $a_1 > a_2$ (or $z_2$ and $z_1 < z_2$).

- Proper distance $D_{prop}$: distance measured by time required for light to travel from a source to an observer, $dD_{prop} = -c\,dt = -\frac{c\,da}{\dot{a}}$, minus sign such that $D_{prop}$ increases away from the observer, while $t$ and $a$ increase towards the observer:

$$D_{prop}(z_1, z_2) = c\int_{a(z_2)}^{a(z_1)} \frac{da}{\dot{a}} = \frac{c}{H_0}\int_{a(z_2)}^{a(z_1)} \frac{da}{a\,E(a)}$$

- Comoving distance $D_{com}$: is distance on the spatial hypersurfaces at $t = \text{const}$ between worldlines of source and observer comoving with the mean cosmic flow, identical to coordinate distance between source and observer, $dD_{com} = dw$. Since light moves on $ds = 0$, $a\,dw = -c\,dt = -\frac{c\,da}{\dot{a}}$

$$D_{com}(z_1, z_2) = c\int_{a(z_1)}^{a(z_2)} \frac{da}{a\,\dot{a}} = \frac{c}{H_0}\int_{a(z_1)}^{a(z_2)} \frac{da}{a^2 E(a)} \equiv w(z_1, z_2)$$

- Angular diameter distance $D_{ang}$: is defined in analogy to the relation in Euclidean space between the area $\delta A$ and the solid angle $\delta\omega$ of an object $\delta\omega\, D_{ang}^2 = \delta A$; since the solid angle of spheres of constant radial coordinate $w$ is scaled by $f_k(w)$ in the Robertson-Walker metric, we have

$$\left(\frac{\text{area of object}}{\text{area of full sphere}} = \right) \frac{\delta A}{4\pi \left(\frac{a_2}{a_1}\right)^2 f_k^2\left[w(z_1,z_2)\right]} = \frac{\delta\omega}{4\pi} \quad \left(= \frac{\text{solid angle of object}}{\text{solid angle of sphere}}\right)$$

$$\Rightarrow D_{ang}(z_1,z_2) = \left(\frac{\delta A}{\delta\omega}\right)^{1/2} = \frac{a(z_2)}{a(z_1)} f_k\left[w(z_1,z_2)\right]$$

as the coordinate distance $w(z_1,z_2) = D_{com}(z_1,z_2)$, we have

$$D_{ang}(z_1,z_2) = \frac{a(z_2)}{a(z_1)} f_k\left[D_{com}(z_1,z_2)\right]$$

- Luminosity distance $D_{lum}$ is defined in analogy to the Euclidean relation between intrinsic luminosity of an object and its flux $F = \frac{L}{4\pi D_{lum}^2}$:

$$D_{lum}(z_1,z_2) = \left[\frac{a(z_1)}{a(z_2)}\right]^2 D_{ang}(z_1,z_2)$$

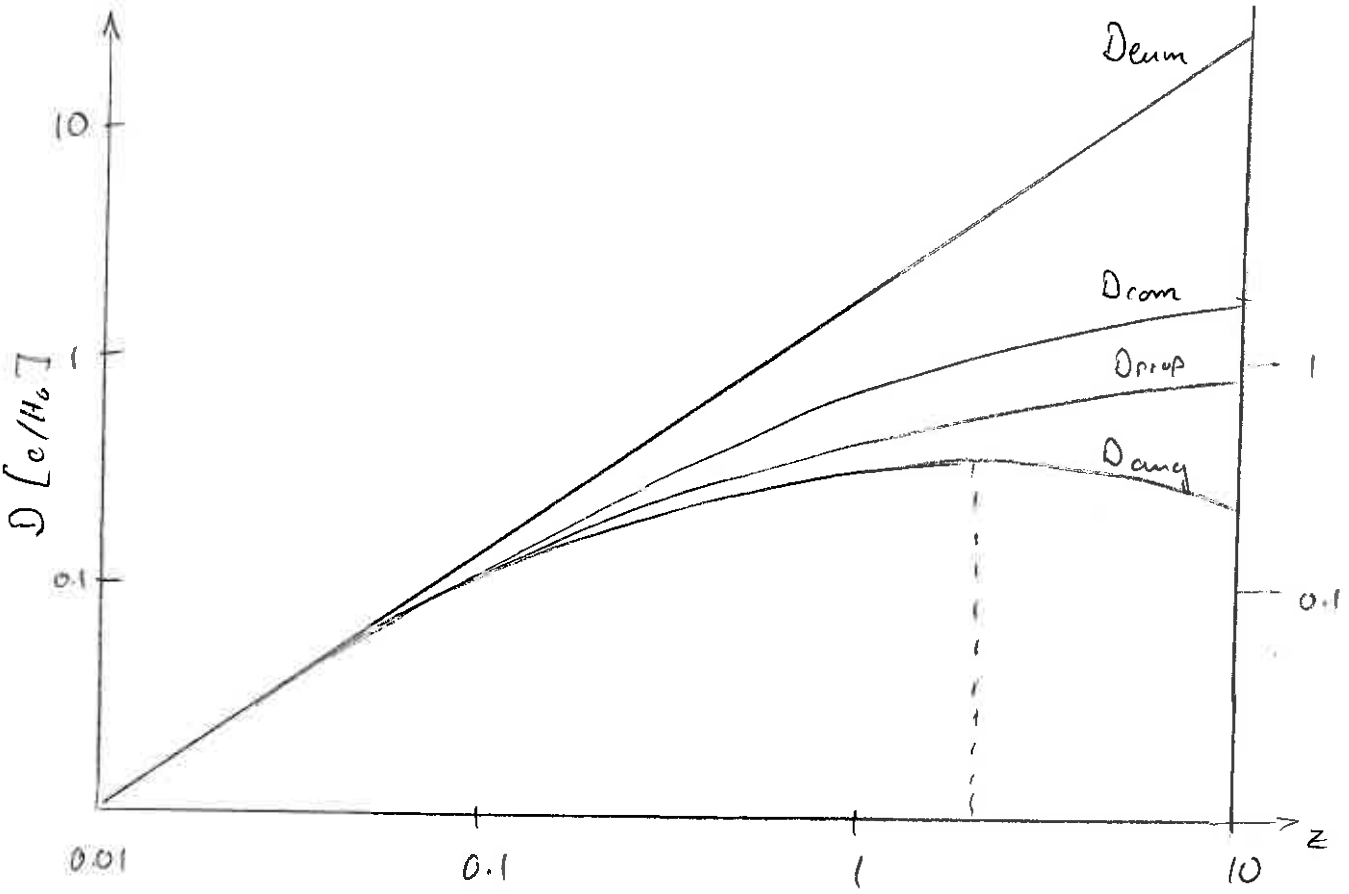
-- photons are redshifted by $\frac{a_1}{a_2}$ between emission and absorption, arrival times are stretched by $\frac{a_1}{a_2}$, and they are spatially diluted by a factor $\left(\frac{a_1}{a_2}\right)^2$; this yields a factor $\left(\frac{a_1}{a_2}\right)^4$ between luminosity and flux and thus a factor $\left(\frac{a_1}{a_2}\right)^2$ in the luminosity distance.

- for $z \ll 1$, $a \approx 1-z$, and $E(a) \approx 1$, we have <u>locally</u>:

$$D \approx \frac{cz}{H_0} + \mathcal{O}(z^2) \quad \text{"Hubble's law"}$$ for all distances above

- $D_{ang}(0,z)$ in Einstein-de Sitter universe ($\Omega_{m0} = 1$, $\Omega_{\Lambda 0} = 0$):

$$D_{ang}(z) = \frac{2c}{H_0}\frac{1}{1+z}\left[1 - \frac{1}{(1+z)^{1/2}}\right]$$ has a maximum for $z = \frac{5}{4}$ and decreases again for increasing $z$! This is a consequence of space-time curvature.




$\Omega_{m0} \simeq 0.27$

$\Omega_{\Lambda 0} \simeq 0.73$

# 2. Evolution of the Dark Component

## 2.1. The Growth of Perturbations

### 2.1.1 The Basic Equations

- There is structure in the universe on all scales: from within galaxies to the large-scale cosmic filamentary web that consists of filaments, clusters at the intersection of filaments, and large voids (that can extend up to $\sim 50\ h^{-1}$ Mpc)
  → basic theory for structure growth in the expanding universe.
- While this theory should be worked out in the framework of general relativity, we restrict ourselves to small scales (relative to the scale of the universe) so that curvature effects and the finite speed of information propagation can be neglected
  → Newtonian dynamics
- we need to consider at least the following components:
  - dark matter (derived from dynamics of stars in galaxies, of galaxies in clusters, grav. lensing ...)
  - baryonic matter
  - radiation
- continuity equation ~ mass conservation:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot (\rho \vec{v}) = 0 \qquad (2.1)$$

$\rho = \rho(t, \vec{r})$ and $\vec{v} = \vec{v}(t, \vec{r})$ are density and velocity field of cosmic fluid at position $\vec{r}$ and time $t$ (in contrast to background quantities, they now depend on position)

- Euler's equation ~ momentum conservation:

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \vec{\nabla})\vec{v} = -\left(\frac{\nabla p}{\rho} + \vec{\nabla}\phi\right) \qquad (2.2)$$

terms on the right-hand side represent the pressure gradient and gravitational forces

- Poisson equation for Newtonian gravitational potential:

$$\nabla^2 \phi = 4\pi G \rho \qquad (2.3)$$

## 2.1.2. Perturbation Equations

- decompose $\rho$ and $\vec{v}$ into their homogeneous background values $\rho_0$ and $\vec{v}_0$ and small perturbations $\delta\rho$ and $\delta\vec{v}$,

$$\rho(t,\vec{r}) = \rho_0(t) + \delta\rho(t,\vec{r}), \qquad \vec{v}(t,\vec{r}) = \vec{v}_0(t) + \delta\vec{v}(t,\vec{r}) \tag{2.4}$$

- let $\vec{r}$ and $\vec{x}$ be physical and comoving coordinates, respectively, and $\vec{r} = a\vec{x}$, we have

$$\vec{v} = \dot{\vec{r}} = \dot{a}\vec{x} + a\dot{\vec{x}} = H\vec{r} + a\dot{\vec{x}} = \vec{v}_0 + \delta\vec{v}, \tag{2.5}$$

i.e. $\vec{v}_0 = H\vec{r}$ is the Hubble velocity and $\delta\vec{v} = a\dot{\vec{x}}$ is the peculiar velocity deviating from the Hubble flow

- inserting (2.4) into (2.1), keeping only first-order terms yields

$$\frac{\partial(\rho_0+\delta\rho)}{\partial t} + \vec{\nabla}\cdot\left(\rho_0\vec{v}_0 + \delta\rho\,\vec{v}_0 + \rho_0\delta\vec{v}\right) = 0 \tag{2.6}$$

- the background quantities need to satisfy mass conservation separately,

$$\frac{\partial\rho_0}{\partial t} + \rho_0\vec{\nabla}\cdot\vec{v}_0 = \frac{\partial\rho_0}{\partial t} + 3H\rho_0 = 0 \tag{2.7}$$

where $\vec{v}_0 = H\vec{r}$ and $\vec{\nabla}_r\cdot\vec{r} = 3$ were used. Hence, we get

$$\underbrace{\frac{\partial\rho_0}{\partial t} + \rho_0\nabla\cdot\vec{v}_0}_{=0 \text{ using (2.7)}} + \frac{\partial\delta\rho}{\partial t} + \delta\rho\,\vec{\nabla}\cdot\vec{v}_0 + \vec{v}_0\cdot\vec{\nabla}\delta\rho + \rho_0\nabla\cdot\delta\vec{v} = \frac{\partial\delta\rho}{\partial t} + \delta\rho\,\vec{\nabla}\cdot\vec{v}_0 + \vec{v}_0\cdot\vec{\nabla}\delta\rho + \rho_0\vec{\nabla}\cdot\delta\vec{v} = 0 \tag{2.8}$$

- defining the density contrast $\delta := \frac{\delta\rho}{\rho_0}$, we have (2.9)

$$\frac{\partial\delta\rho}{\partial t} = \dot{\delta\rho}_0 + \dot{\delta}\rho_0 = -\delta\rho_0\,\vec{\nabla}\cdot\vec{v}_0 + \dot{\delta}\rho_0, \tag{2.10}$$

using the unperturbed continuity equation (2.7), we can rewrite (2.8) by using (2.10):

$$-\delta\rho_0\vec{\nabla}\cdot\vec{v}_0 + \dot{\delta}\rho_0 + \delta\rho_0\vec{\nabla}\cdot\vec{v}_0 + \vec{v}_0\cdot\nabla\delta\rho + \rho_0\vec{\nabla}\cdot\delta\vec{v} = 0 \quad /:\rho_0$$

$$\dot{\delta} + \vec{v}_0\cdot\vec{\nabla}\delta + \vec{\nabla}\cdot\delta\vec{v} = 0 \tag{2.11}$$

- Similarly, we split Euler's equation (2.2) in an unperturbed and perturbed part, introducing pressure and potential perturbations $\delta p$ and $\delta\phi$, respectively, and keep only terms to linear order:

$$\frac{\partial\delta\vec{v}}{\partial t} + (\delta\vec{v}\cdot\vec{\nabla})\vec{v}_0 + (\vec{v}_0\cdot\vec{\nabla})\delta\vec{v} = -\frac{\vec{\nabla}\delta p}{\rho_0} - \vec{\nabla}\delta\phi \tag{2.12}$$

- the term $(\delta\vec{v}\cdot\vec{\nabla})\vec{v}_0$ reads (written componentwise):

$$\left[(\delta\vec{v}\cdot\vec{\nabla})\vec{v}_0\right]_i = \left(\delta v_j \frac{\partial}{\partial r_j}\right) H r_i = H\,\delta_{ij}\,\delta v_j = H\,\delta v_i = \left[H\delta\vec{v}\right]_i \qquad (2.13)$$

hence we have for (2.12):

$$\frac{\partial\delta\vec{v}}{\partial t} + H\delta\vec{v} + H\vec{r}\cdot\vec{\nabla}_r\,\delta\vec{v} = -\frac{\vec{\nabla}\delta p}{\rho_0} - \vec{\nabla}\delta\phi$$

$$\left[\frac{\partial}{\partial t} + H\vec{r}\cdot\vec{\nabla}_r\right]\delta\vec{v}_{=a\dot{u}} + H\delta\vec{v} = \partial_t(a\vec{u}) + Ha\vec{u} = \dot{a}\vec{u} + a\dot{\vec{u}} + Ha\vec{u} = a\dot{\vec{u}} + 2aH\vec{u} \quad \Big| :a \quad \text{(↳ (2.17))}$$

- a perturbed/unperturbed split of Poisson's equation yields

$$\nabla^2\delta\phi = 4\pi G\rho_0\delta \qquad (2.14)$$

- Now we convert to comoving positions, $\vec{x} := \frac{\vec{r}}{a}$ and comoving peculiar velocities, $\vec{u} := \frac{\delta\vec{v}}{a}$, and introduce the gradient with respect to comoving coordinates,

$$\vec{\nabla}_r = \frac{1}{a}\vec{\nabla}_x \quad \text{such that} \quad \vec{r}\cdot\vec{\nabla}_r = \vec{x}\cdot\vec{\nabla}_x \qquad (2.15)$$

- also, we have to transform the time derivatives; the total differential of $f(\vec{r},t)$ is

$$df = \frac{\partial f}{\partial t}dt + \vec{\nabla}_r f\cdot d\vec{r} = \frac{\partial f}{\partial t}dt + \vec{\nabla}_r f\cdot a\left(H\vec{x}\,dt + d\vec{x}\right)$$
$$= \left(\frac{\partial f}{\partial t} + H\vec{x}\cdot\vec{\nabla}_x f\right)dt + \vec{\nabla}_x f\cdot d\vec{x} \qquad \left[\text{using (2.5): } \frac{d\vec{r}}{dt} = a\frac{\dot{a}}{a}\vec{x} + a\frac{d\vec{x}}{dt}\right]$$

hence, we obtain a replacement rule for the partial time derivative in physical coordinates:

$$\left(\frac{\partial}{\partial t} + H\vec{x}\cdot\vec{\nabla}_x\right) = \left(\frac{\partial}{\partial t} + \vec{v}_0\cdot\vec{\nabla}_r\right) \longrightarrow \frac{\partial}{\partial t} \qquad (2.16)$$

- We now use the abbreviation $\vec{\nabla}$ for $\vec{\nabla}_x$ and have the 3 basic perturbation equations:

$$\dot{\delta} + \vec{\nabla}\cdot\vec{u} = 0$$
$$\dot{\vec{u}} + 2H\vec{u} = -\frac{\vec{\nabla}\delta p}{a^2\rho_0} - \frac{\vec{\nabla}\delta\phi}{a^2} \qquad (2.17)$$
$$\nabla^2\delta\phi = 4\pi G\rho_0 a^2\delta$$

for four variables $\delta$, $\vec{u}$, $\delta\phi$, and $\delta p$. We need an equation of state to close the system:

assume adiabatic perturbations

$$\delta p = \delta p(\delta, S) \overset{\downarrow}{=} \delta p(\delta) = c_s^2\,\delta\rho = c_s^2\rho_0\delta \qquad (2.18)$$

with sound speed $c_s$.

## 2.1.3. Density Perturbations

- to derive a single equation for $\delta$, we take the divergence of the Euler equation,

$$\vec{\nabla}\cdot\dot{\vec{u}} + 2H\,\vec{\nabla}\cdot\vec{u} = \frac{d}{dt}\left(\vec{\nabla}\cdot\vec{u}\right) + 2H\,\vec{\nabla}\cdot\vec{u} = -\frac{1}{a^2\rho_0}\nabla^2\delta p - \frac{1}{a^2}\nabla^2\delta\phi$$

and the total time derivative of the continuity equation,

$$\ddot{\delta} + \frac{d}{dt}\,\vec{\nabla}\cdot\vec{u} = 0$$

Combining those and the Poisson equation & EOS yields (since $\vec{\nabla}\cdot\dot{\vec{u}} = \frac{d}{dt}\,\vec{\nabla}\cdot\vec{u}$)

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G\rho_0\delta + \frac{c_s^2\nabla^2\delta}{a^2} \qquad (2.19)$$

- discussion:
  - second term on lhs is the "Hubble drag" term which suppresses perturbation growth due to the expansion of the universe
  - first term on lhs is the gravitational term, which causes perturbations to grow via gravitational instability
  - last term on lhs is a pressure term due to spatial variations in the density

- strategy: we can decompose $\delta$ into plane waves,

$$\delta(\vec{x},t) = \int \frac{d^3k}{(2\pi)^3}\,\hat{\delta}(\vec{k},t)\,e^{-i\vec{k}\cdot\vec{x}} \qquad (2.20)$$

where $\hat{\delta}(\vec{k},t)$ are time-dependent Fourier amplitudes; inserting into (2.19) yields

$$\ddot{\hat{\delta}} + 2H\dot{\hat{\delta}} = \hat{\delta}\left(4\pi G\rho_0 - \frac{c_s^2k^2}{a^2}\right) \qquad (2.21)$$

→ note that this equation was derived from the non-relativistic Euler equation, hence it is only valid in the <u>matter-dominated era</u>!

→ what about radiation-dominated era?

- using special-relativistic fluid mechanics, ignoring pressure gradients, the perturbation equations for an ideal relativistic fluid (e.g. photons) can be derived in a similar way with the following replacements:

continuity equation: $\vec{\nabla}\cdot(\rho\vec{v}) \longrightarrow \vec{\nabla}\cdot\left[\left(\rho+\frac{p}{c^2}\right)\vec{v}\right] = \frac{4}{3}\vec{\nabla}(\rho\vec{v})$

Poisson equation: $\rho \longrightarrow \rho + 3\frac{p}{c^2} = 2\rho$

with the pressure of a relativistic gas is $p = \frac{\rho c^2}{3}$

$\Rightarrow$ In the radiation-dominated era, we get:

$$\ddot{\hat{\delta}} + 2H\dot{\hat{\delta}} = 2\cdot\frac{4}{3}\cdot 4\pi G\rho_0\hat{\delta} = \frac{32\pi}{3}G\rho_0\hat{\delta} \qquad (2.22)$$

- Solutions:

1) assume a static background, $H=0$, (2.21) becomes the oscillator equation

$$\ddot{\hat{\delta}} + \omega_0^2\hat{\delta} = 0, \qquad \omega_0 := \sqrt{\frac{c_s^2k^2}{a^2} - 4\pi G\rho_0} \qquad (2.23)$$

$\omega_0$ is real for $k \geq k_J := \frac{2a\sqrt{\pi G\rho_0}}{c_s}$, the so-called Jeans wavenumber (2.24)

which defines the Jeans length $\lambda_J := \frac{2\pi}{k_J} = \frac{c_s}{a}\sqrt{\frac{\pi}{G\rho_0}}$ (comoving) (2.25)

$\Rightarrow$ perturbations smaller than $\lambda_J$ oscillate, others grow or decay

Note that one can derive the Jeans length by balancing the sound crossing time $\tau_s = \frac{a\lambda_J}{c_s} = \frac{2\pi a}{k_J c_s}$ with the gravitational free-fall time; $\tau_{ff} = \sqrt{\frac{\pi}{G\rho_0}}$; equating yields $k_J = \frac{a}{c_s}\sqrt{4\pi G\rho_0}$

2) What about scales $\lambda \gg \lambda_J$ (or equivalently in pressureless fluids)?

If $\Omega = 1$, we have $\rho_{cr} = \rho_0 = \frac{3H^2}{8\pi G} \Rightarrow 4\pi G\rho_0 = \frac{3}{2}H^2$

$$\ddot{\hat{\delta}} + 2H\dot{\hat{\delta}} = \frac{3}{2}H^2\hat{\delta} \text{ (matter-dom.)}, \quad \ddot{\hat{\delta}} + 2H\dot{\hat{\delta}} = 4H^2\hat{\delta} \text{ (rad.-dom.)} \qquad (2.26)$$

We have solutions for $H$ in these two eras (see Sect. 1.3.4, cases 4 and 1),

$$\frac{\dot{a}}{a} = H(t) = \frac{2}{3t} \text{ (matter-dom.)}, \quad \frac{\dot{a}}{a} = H(t) = \frac{1}{2t} \qquad (2.27)$$

- ansatz $\hat{\delta}(k,t) \propto t^n$ yields

matter-dom. era: $\ddot{\hat{\delta}} + \frac{4}{3}\frac{\dot{\hat{\delta}}}{t} = \frac{3}{2}\frac{4}{9t^2}\hat{\delta}$

$$n(n-1) + \frac{4}{3}n - \frac{2}{3} = 0$$

$$n^2 + \frac{1}{3}n - \frac{2}{3} = 0 \quad \rightarrow \quad n = -1, \frac{2}{3} \qquad (2.28a)$$

rad.-dom. era: $\ddot{\hat{\delta}} + \frac{\dot{\hat{\delta}}}{t} = \frac{4}{4t^2}\hat{\delta}$

$$n(n-1) + n - 1 = 0$$

$$n^2 - 1 = 0 \quad \rightarrow \quad n = \pm 1 \qquad (2.28b)$$

Hence, we get the growing solutions (in the Einstein-de Sitter limit, $z \gtrsim 2$ in our universe):

$$\hat{\delta} \propto \begin{cases} a^2 & \text{radiation-dominated era} \\ a & \text{matter-dominated era} \end{cases} \qquad (2.29)$$

and the decaying solutions (which are irrelevant for cosmological structure growth):

$$\hat{\delta} \propto \begin{cases} a^{-2} & \text{rad.-dominated era} \\ a^{-3/2} & \text{matter-dominated era} \end{cases}$$

3) <u>During the matter-dominated era, for $\Omega_{m,0} \neq 1$ and $\Omega_{\Lambda,0} \neq 0$:</u>

$$\ddot{\delta} + 2H\dot{\delta} = 4\pi G \rho_0 \delta \qquad \text{(we omit the 'hats' for brevity)}$$

* change of variables: $\frac{d}{dt} = \dot{a}\frac{d}{da} = aH\frac{d}{da}$, $\dot{\delta} = aH\delta'$, $(\cdot)' \equiv \frac{d}{da}$ (2.30)

$$\ddot{\delta} = aH\frac{d}{da}(aH\delta') = aH\left[H\delta' + aH'\delta' + aH\delta''\right]$$

$$= (aH)^2\delta'' + aH^2\left(1 + a\frac{H'}{H}\right)\delta' = (aH)^2\left[\delta'' + \left(\frac{1}{a} + \frac{H'}{H}\right)\delta'\right]$$

$$\ddot{\delta} + 2H\dot{\delta} = (aH)^2\left[\delta'' + \left(\frac{1}{a} + \frac{H'}{H}\right)\delta'\right] + 2aH^2\delta'$$

$$= (aH)^2\left[\delta'' + \left(\frac{1}{a} + \frac{H'}{H} + \frac{2}{a}\right)\delta'\right]$$

$$= (aH)^2\left[\delta'' + \left(\frac{3}{a} + \frac{H'}{H}\right)\delta'\right] = 4\pi G\rho_0\delta \quad , \quad \rho_0 = \frac{\rho_{cr,0}}{a^3}\Omega_{m,0}$$

$$\delta'' + \left(\frac{3}{a} + \frac{H'}{H}\right)\delta' = \frac{4\pi G}{a^2H^2}\frac{3H_0^2}{8\pi G a^3}\Omega_{m,0}\delta$$

$$\delta'' + \left(\frac{3}{a} + \frac{H'}{H}\right)\delta' = \frac{3}{2}\frac{\Omega_{m,0}H_0^2}{H^2a^5}\delta \qquad (2.31)$$

* next we rewrite the l.h.s as a total derivative:

$$\frac{1}{a^3 H}\left(a^3 H \delta'\right)' = \frac{1}{a^3 H}\left[\left(3a^2 H + a^3 H'\right)\delta' + a^3 H \delta''\right] = \delta'' + \left(\frac{3}{a} + \frac{H'}{H}\right)\delta' = \frac{3}{2}\frac{\Omega_{m0} H_0^2}{H^2 a^3}\delta \qquad (2.32)$$

* 1st solution: $\delta_1 = H$ insert into (2.32) (2.33)

$$\left(a^3 H H'\right)' = \frac{3}{2}\frac{\Omega_{m0} H_0^2}{a^2} \rightsquigarrow a^3 H H' = -\frac{3\,\Omega_{m0} H_0^2}{2a} + A_1$$

$$H H' = \frac{1}{2}\left(H^2\right)' = -\frac{3\,\Omega_{m0} H_0^2}{2a^4} + \frac{A_1}{a^3}$$

$$H^2 = \frac{3\,\Omega_{m0} H_0^2}{a^3\,3} - \frac{2A_1}{2a^2} + A_2 = H_0^2\left(\frac{\Omega_{m0}}{a^3} + \frac{\Omega_k}{a^2} + \Omega_\Lambda\right) \checkmark$$

with $A_1 = -H_0^2 \Omega_K$, $A_2 = H_0^2 \Omega_\Lambda$

* 2nd solution: $\delta_2 = \delta_1 f$ with $\delta_1 = H$ (2.34)

(2.32): $\left(a^3 H \delta'\right)' = \frac{C\delta}{H a^2}$ ; $C = \frac{3}{2}\Omega_{m0} H_0^2$

$$\left[a^3 H\left(\delta_1' f + \delta_1 f'\right)\right]' = \frac{C\delta_1 f}{H a^2}$$

$$\left(a^3 H\right)'\left(\delta_1' f + \delta_1 f'\right) + a^3 H\left(\delta_1'' f + 2\delta_1' f' + \delta_1 f''\right) = \frac{C\delta_1 f}{H a^2} \qquad (2.35)$$

$$f'' \cdot a^3 H \delta_1 + f' \cdot\left[\left(a^3 H\right)'\delta_1 + 2a^3 H \delta_1'\right] + f\cdot\left[\left(a^3 H\right)'\delta_1' + a^3 H \delta_1''\right] = \frac{C\delta_1 f}{H a^2}$$

$$f\left[\left(a^3 H\right)'\delta_1' + a^3 H \delta_1''\right] = f\left(a^3 H \delta_1'\right)' = \frac{C\delta_1 f}{H a^2} \quad : \delta_1 = H \text{ solves this, cf. (2.33)}$$

Hence we are left with first two terms in (2.35):

$$\frac{f''}{f'} = -\frac{\left(a^3 H\right)'\delta_1 + 2a^3 H\delta_1'}{a^3 H \delta_1} = -\frac{3a^2 H^2 + a^3 H' H + 2a^3 H H'}{a^3 H^2}$$

$$= -\frac{3a^2 H\left(H + aH'\right)}{a^3 H^2} = -\frac{3(aH)'}{aH} = -3\left[\ln(aH)\right]'$$

$$\Rightarrow \ln f' = -3\ln(aH) \rightarrow f' = \frac{1}{(aH)^3}$$

$$\Rightarrow \delta_2 = H\int_0^a \frac{d\tilde{a}}{(\tilde{a}H)^3} \qquad (2.36)$$

The solution $\delta_2 = H(a) \int_0^a \frac{d\tilde{a}}{(\tilde{a}H)^3}$ is determined up to an overall constant factor.

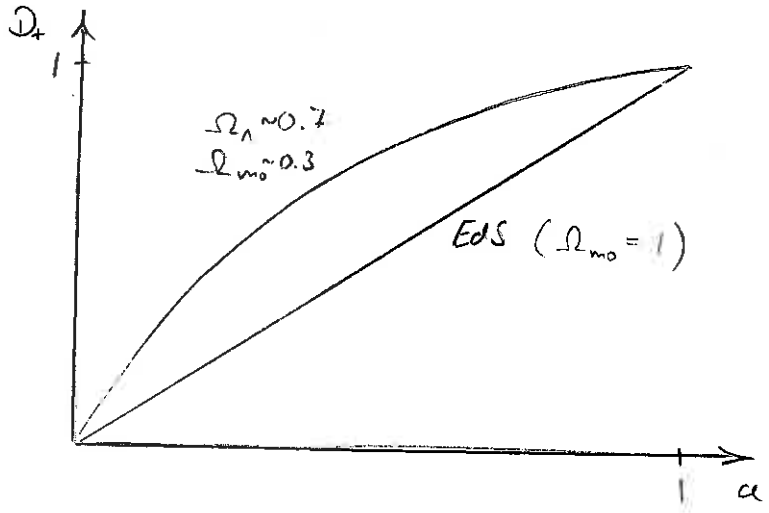
Usually one makes the choice

$$\delta(a) = \delta_0 D_+(a) \tag{2.37}$$

$$D_+(a) = \frac{5}{2} \Omega_{m0} E(a) \int_0^a \frac{d\tilde{a}}{\tilde{a}^3 E^3(\tilde{a})} \tag{2.38}$$

an excellent approximation is given by ( Carroll et al. 1992 )

$$D_+(a) = \frac{5a}{2} \Omega_m(a) \left[ \Omega_m^{4/3}(a) - \Omega_\Lambda + \left(1 + \frac{\Omega_m(a)}{2}\right)\left(1 + \frac{\Omega_\Lambda}{70}\right) \right]^{-1} \tag{2.39}$$




- Effect of $\Lambda$-term : - earlier times ~ faster growth
  - later times ~ "slowing down" of structure formation and eventually halting and reversing it !

## 2.2. Statistics and Non-linear Evolution

### 2.2.1 Power spectra

- let's decompose the density contrast into plane waves; introduce Fourier transform $\hat{\delta}$ of density contrast $\delta$ as

$$\delta(\vec{x}) = \int \frac{d^3k}{(2\pi)^3} \hat{\delta}(\vec{k}) e^{-i\vec{k}\cdot\vec{x}}, \qquad \hat{\delta}(\vec{k}) = \int d^3x\, \delta(\vec{x}) e^{i\vec{k}\cdot\vec{x}} \tag{2.40}$$

- $\delta$ is a random field that must be isotropic and homogeneous (fundamental cosmological assumptions); the statistical properties of $\delta$ (mean, variance, etc.) must not change under rotations and translations
- mean of $\delta$ vanishes by definition

$$\langle \delta \rangle = \left\langle \frac{\rho - \bar{\rho}}{\bar{\rho}} \right\rangle = \frac{\langle \rho \rangle}{\bar{\rho}} - 1 = 0 \tag{2.41}$$

variance of $\delta$ in Fourier space defines the power spectrum $P(k)$

$$\langle \hat{\delta}(\vec{k}) \hat{\delta}^*(\vec{k}') \rangle =: (2\pi)^3 P(k)\, \delta_D(\vec{k} - \vec{k}') \tag{2.42}$$

where $\delta_D$ is Dirac's delta distribution, which ensures that modes of different wave vector $\vec{k}$ are uncorrelated in Fourier space as required by homogeneity; the power spectrum can not depend on a preferred direction in Fourier space because of isotropy

- the correlation function of $\delta$ in real space is defined as

$$\xi(y) := \langle \delta(\vec{x})\, \delta(\vec{x} + \vec{y}) \rangle \tag{2.43}$$

where the brackets resemble ensemble averages at all positions $\vec{x}$; $\xi$ measures the coherence of the density contrast between <u>all</u> points separated by a distance $|\vec{y}|$; as for the power spectrum, $\xi$ cannot depend on direction of $\vec{y}$ because of isotropy; applying the ergodic hypothesis enables to interpret the ensemble averages as averages over all positions $\vec{x}$ and directions $\vec{y}$

- inserting Fourier integrals (2.40) into (2.43), we have

$$\xi(y) = \left\langle \int \frac{d^3k}{(2\pi)^3} \int \frac{d^3k'}{(2\pi)^3} \hat{\delta}(\vec{k}) \hat{\delta}(\vec{k}') e^{-i\vec{k}\cdot\vec{x}} e^{-i\vec{k}'\cdot(\vec{x}+\vec{y})} \right\rangle$$

$$= \int \frac{d^3k}{(2\pi)^3} \int \frac{d^3k'}{(2\pi)^3} \langle \hat{\delta}(\vec{k}) \hat{\delta}^*(\vec{k}') \rangle e^{-i\vec{k}\cdot\vec{x}} e^{+i\vec{k}'\cdot(\vec{x}+\vec{y})}$$

↑ phase reversal by complex conjugate

$$= \int \frac{d^3k}{(2\pi)^3} P(k) e^{i\vec{k}\cdot\vec{y}} \qquad (2.44)$$

⇒ the correlation function is the Fourier transform of the power spectrum (and vice versa)! They carry an equivalent amount of information.

- $$\mathrm{Re} \int_0^\pi \sin\vartheta d\vartheta e^{i\vec{k}\cdot\vec{y}} = \mathrm{Re} \int_0^\pi \sin\vartheta d\vartheta e^{iky\cos\vartheta} = 2\int_0^1 d\cos\vartheta \cos(ky\cos\vartheta)$$

$$= 2\,\frac{\sin ky}{ky}$$

$$\Rightarrow \xi(y) = \int \frac{d^3k}{(2\pi)^3} P(k) e^{i\vec{k}\cdot\vec{y}} = 2\pi \int \frac{k^2 dk}{(2\pi)^3} P(k) \int_0^\pi \sin\vartheta d\vartheta e^{iky\cos\vartheta}$$

$$\underline{\xi(y) = 4\pi \int \frac{k^2 dk}{(2\pi)^3} P(k) \frac{\sin ky}{ky}} \qquad (2.45)$$

- Clearly, the variance of $\delta$ is $\xi(y=0)$:

$$\sigma^2 = 4\pi \int \frac{k^2 dk}{(2\pi)^3} P(k) \qquad (2.46)$$

- The variance in real space depends on the scale we are interested in; we introduce a window function $W_R(|\vec{x}|)$ that approaches unity (or a finite constant) for scales $x \ll R$ and drops to zero outside $R$; we introduce

$$\bar{\delta}(\vec{x}) := \int d^3y \, \delta(\vec{x}) \, W_R(|\vec{x}-\vec{y}|) \qquad (2.47)$$

the density contrast averaged on scale $R$ with the window function $W_R$

- According to the Fourier convolution theorem, $\widehat{f \otimes g} = \hat{f}\,\hat{g}$, i.e. the Fourier transform of a convolution is the product of the Fourier transforms of the two functions; hence

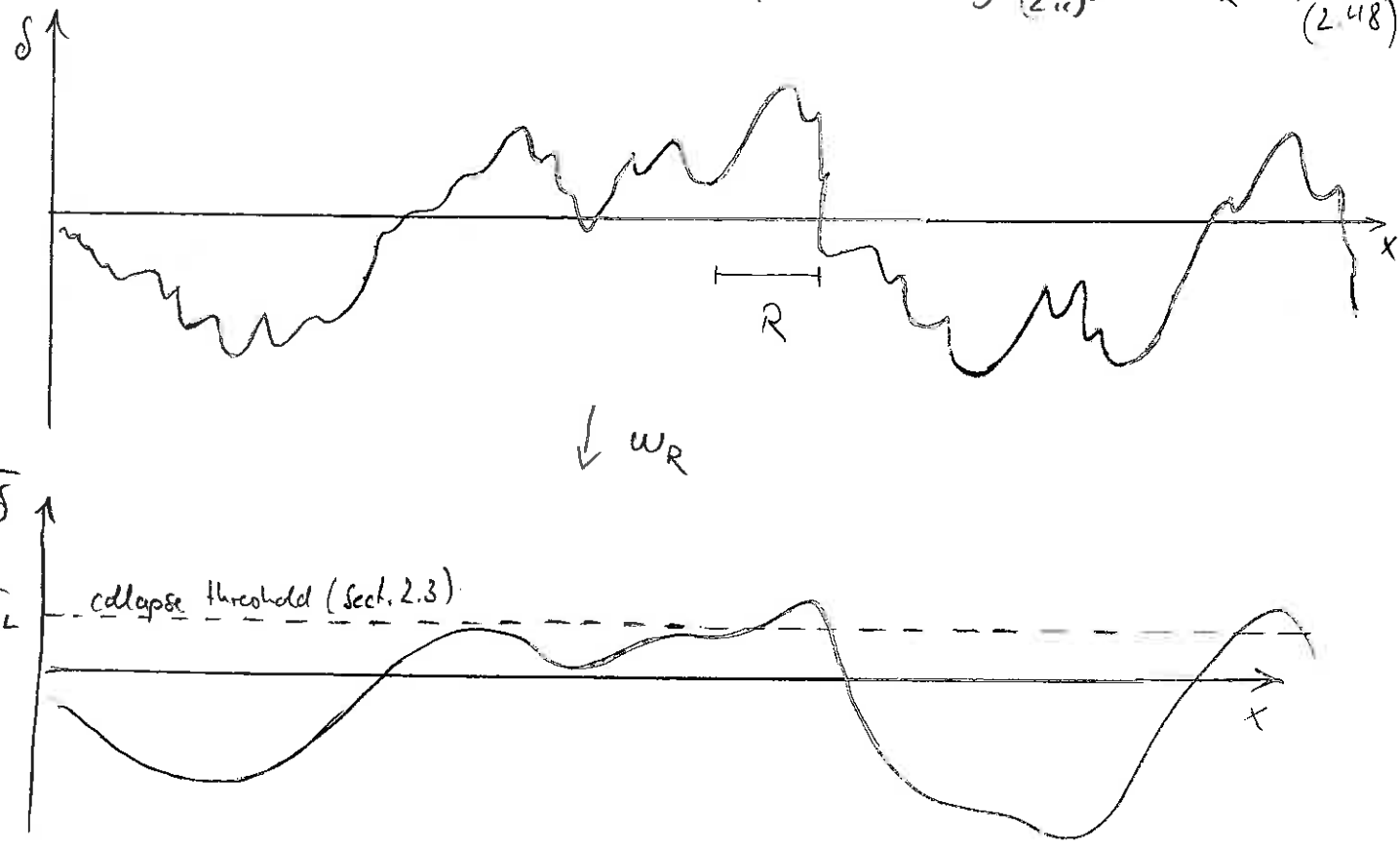
$$\hat{\bar{\delta}} = \widehat{\delta \otimes W_R} = \hat{\delta}\,\hat{W}_R$$

Hence the power spectrum of the density contrast, filtered on the scale $R$ is

$$\bar{P}(k) = P(k)\,\hat{W}_R^2(k)$$

The variance of the filtered density contrast is

$$\sigma_R^2 = \langle \delta^2(\vec{x}) \rangle = \left\langle \left| \frac{\delta M}{M} \right|^2 \right\rangle = 4\pi \int \frac{k^2 dk}{(2\pi)^3} P(k)\,\hat{W}_R^2(k) \qquad (2.48)$$




→ density and mass per volume fluctuate in space, where there is more mass, gravity will collect this and collapse; varying $R$ singles out specific objects, such as clusters;

→ the variance on a scale $8\,h^{-1}$ Mpc, $\sigma_8$, is often used for characterising the amplitude of the power spectrum.

## 2.2.2. Evolution of the power spectrum

- in the radiation-dominated era, $\delta \propto a^2$, and the growth slows down in the matter-dominated era with $\delta \propto a$.
- the Hubble radius grows as the universe expands. Since we are interested in early times, i.e. what happens around the matter-radiation equality $a_{eq} = \frac{\Omega_{r0}}{\Omega_{m0}}$, we can use the Einstein-de Sitter limit of the Hubble function

$$\frac{H}{H_0} = \sqrt{\frac{\Omega_{m0}}{a^3} + \frac{\Omega_{r0}}{a^4}} = \frac{\sqrt{\Omega_{m0}}}{a^{3/2}} \left(1 + \frac{a_{eq}}{a}\right)^{1/2} \qquad (2.49)$$

$$\text{hence, } r_H = \frac{c}{H(a)} = \frac{c}{H_0} \frac{a^{3/2}}{\sqrt{\Omega_{m0}}} \left(1 + \frac{a_{eq}}{a}\right)^{-1/2} \propto \begin{cases} a^2 & \text{rad.-dominated era} \\ a^{3/2} & \text{matter-dominated era} \end{cases} \qquad (2.50)$$

$$r_H(a_{eq}) = \frac{c}{H_0} \frac{a_{eq}^{3/2}}{\sqrt{2\,\Omega_{m0}}} \qquad (2.51)$$

- a density perturbation "enters the horizon" when its wave length $\lambda$ equals the Hubble radius, $r_H(a) = \lambda$; if $\lambda < r_H(a_{eq})$, the perturbation enters the horizon while radiation is still dominating the expansion. Until $a_{eq}$, the expansion time scale, $t_{exp} = H^{-1}$, is determined by the radiation pressure $\rho_r$, which is shorter than the gravitational collapse time scale of the dark matter, $t_{collapse}$:

$$t_{exp} \sim (G\rho_r)^{-1/2} < (G\rho_m)^{-1/2} \sim t_{collapse} \qquad (2.52)$$

In other words, the fast radiation-driven expansion prevents (dark-matter) density perturbations from collapsing. Since light can only cross regions smaller than the horizon size, the suppression of growth is thus only restricted to scales smaller than the horizon, $r_H$; and large-scale perturbations remain unaffected that have a wave length $\lambda > r_H(a_{eq})$. After matter-radiation equality, $a_{eq}$, matter starts dominating and the growth continues (albeit at a slower rate according to $\delta \propto a$).

- a mode with comoving wave number $k$ enters the horizon at $a_{eq}$

$$\lambda = \lambda_0 = a_{eq}\frac{2\pi}{k_0} = r_{H,eq} = \frac{c}{H_0}\frac{a_{eq}^{3/2}}{\sqrt{2\Omega_{m0}}} \qquad (2.52)$$

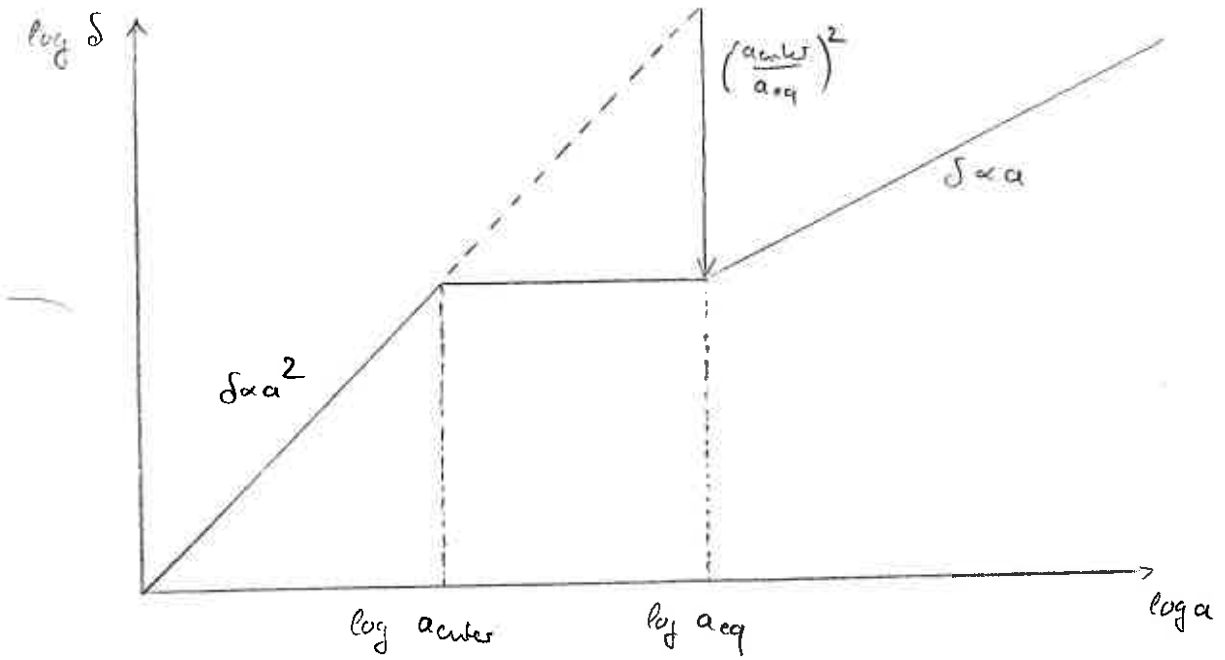
thus $$k_0 = 2\pi\frac{H_0}{c}\sqrt{\frac{2\Omega_{m0}}{a_{eq}}} = 2\pi\frac{H_0}{c}\Omega_{m0}\sqrt{\frac{2}{\Omega_{r0}}} \qquad (2.53)$$

modes larger than $\lambda_0$, i.e. with $k < k_0$, continue growing; modes with $k > k_0$ stop growing when they enter the horizon at $a_{eq}$ and only continue for $a > a_{eq}$, when radiation pressure ceases to be important.

- according to (2.50), $r_H \propto a^2$ in the rad.-dominated era, and $r \propto a^{3/2}$ afterwards, hence

$$\lambda_{phys} = a_{enter}\frac{2\pi}{k} \propto \begin{cases} a_{enter}^2 & (a_{enter} < a_{eq}) \\ a_{enter}^{3/2} & (a_{enter} > a_{eq}) \end{cases}$$

$$\Rightarrow \quad a_{enter} \propto \begin{cases} k^{-1} & (a_{enter} < a_{eq}) \\ k^{-2} & (a_{enter} > a_{eq}) \end{cases} \qquad (2.54)$$




- while the growth of small modes is suppressed, modes with $\lambda > \lambda_0$ continue growing as $\delta \propto a^2$ during radiation domination, hence the relative suppression of small modes is

$$f_{sup} = \left(\frac{a_{enter}}{a_{eq}}\right)^2 = \left(\frac{k_0}{k}\right)^2 \qquad (2.55)$$

- suppose the initial power spectrum at very early times is $P_i(k)$; when modes enter the horizon before $a_{eq}$, the spectrum is $P_{enter}(k) = a_{enter}^4 P_i(k)$ and $P_{enter}(k) = a_{enter}^2 P_i(k)$ if they enter afterwards:

$$P_{enter}(k) \propto \langle \hat{\delta}^2 \rangle \propto \begin{cases} a_{enter}^4 P_i(k) & (a_{enter} < a_{eq} \text{ or equival. } k > k_0) \\ a_{enter}^2 P_i(k) & (a_{enter} > a_{eq} \text{ or equival. } k < k_0) \end{cases} \qquad (2.56)$$

in both cases $P_{enter}(k) = k^{-4} P_i(k)$ because of (2.54)

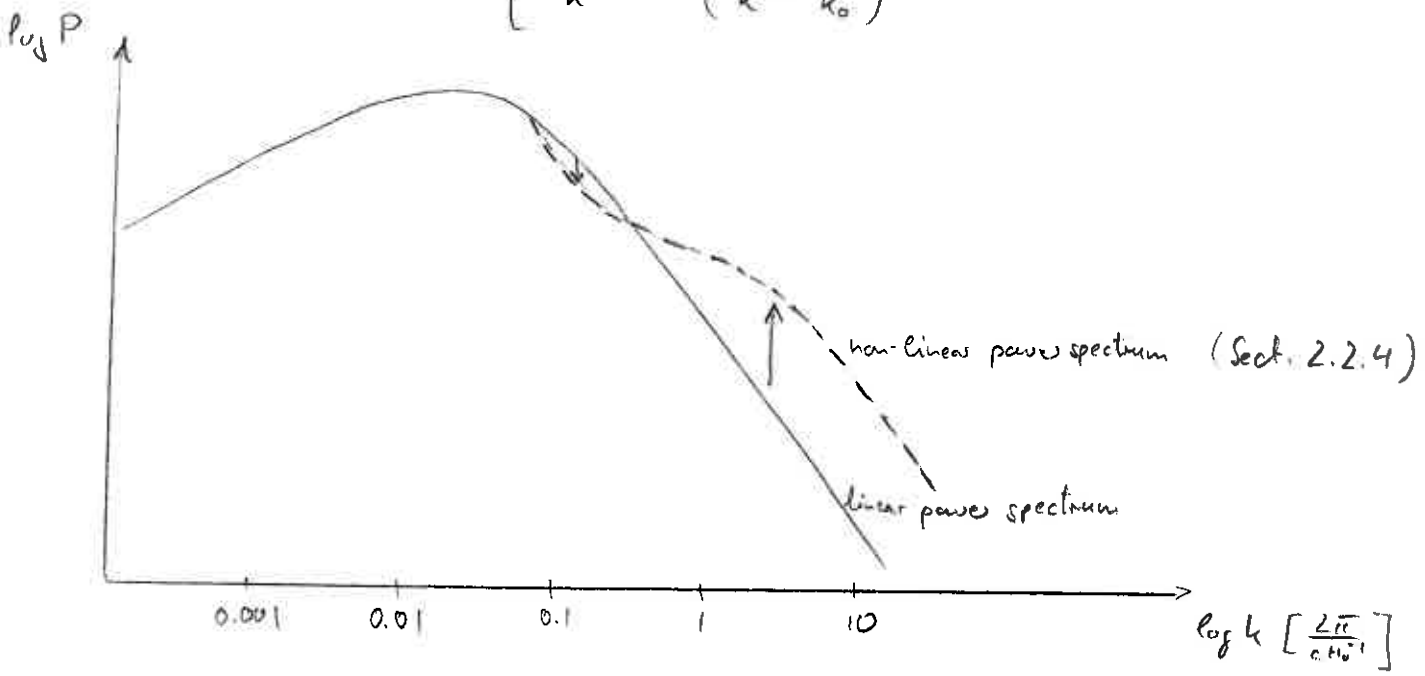
- the total power in density fluctuations on scales $\lambda = \frac{2\pi}{k}$ is $k^3 P(k)$; assuming that the power entering the horizon should not depend on time, the initial power spectrum must obey

$$k^3 P_{enter}(k) = k^3 k^{-4} P_i(k) = \text{const} \Rightarrow P_i(k) \propto k \qquad (2.57)$$

which is the Harrison - Zel'dovich - Peebles spectrum.

- for $k < k_0$, the shape of the spectrum is unchanged because all such modes grew similarly; for $k > k_0$ the suppression $\propto f_{sup}^2 \propto k^{-4}$ sets in; hence we expect the spectrum to scale as

$$P(k) \propto \begin{cases} k & (k < k_0) \\ k^{-3} & (k \gg k_0) \end{cases} \qquad (2.58)$$



## 2.2.3 Hierarchical structure formation

- after studying how perturbations grows and how the spectrum of perturbations evolves as a function of time, we turn to the implications of these findings: hierarchical structure formation in ΛCDM.
- the mass $M_*$ corresponding to the scale $R_*$, on which the variance becomes unity

$$\sigma_*^2 = 4\pi \int_0^{k_*} \frac{k^2 dk}{(2\pi)^3} P(k) \stackrel{!}{=} 1 \qquad (2.56)$$

is the "non-linear" mass.

- assume a power spectrum of the form $P(k) = A k^n$

$$\sigma_*^2 = 4\pi A \int_0^{k_*} \frac{k^{2+n} dk}{(2\pi)^3} = \frac{4\pi A}{(2\pi)^3} \frac{k_*^{n+3}}{n+3} \stackrel{!}{=} 1$$

$$\rightsquigarrow A = \frac{n+3}{k_*^{n+3}} \frac{(2\pi)^3}{4\pi}$$

$$\rightsquigarrow \sigma^2 = \sigma_*^2 \left(\frac{k}{k_*}\right)^{n+3} = \left(\frac{k}{k_*}\right)^{n+3} \qquad (2.57)$$

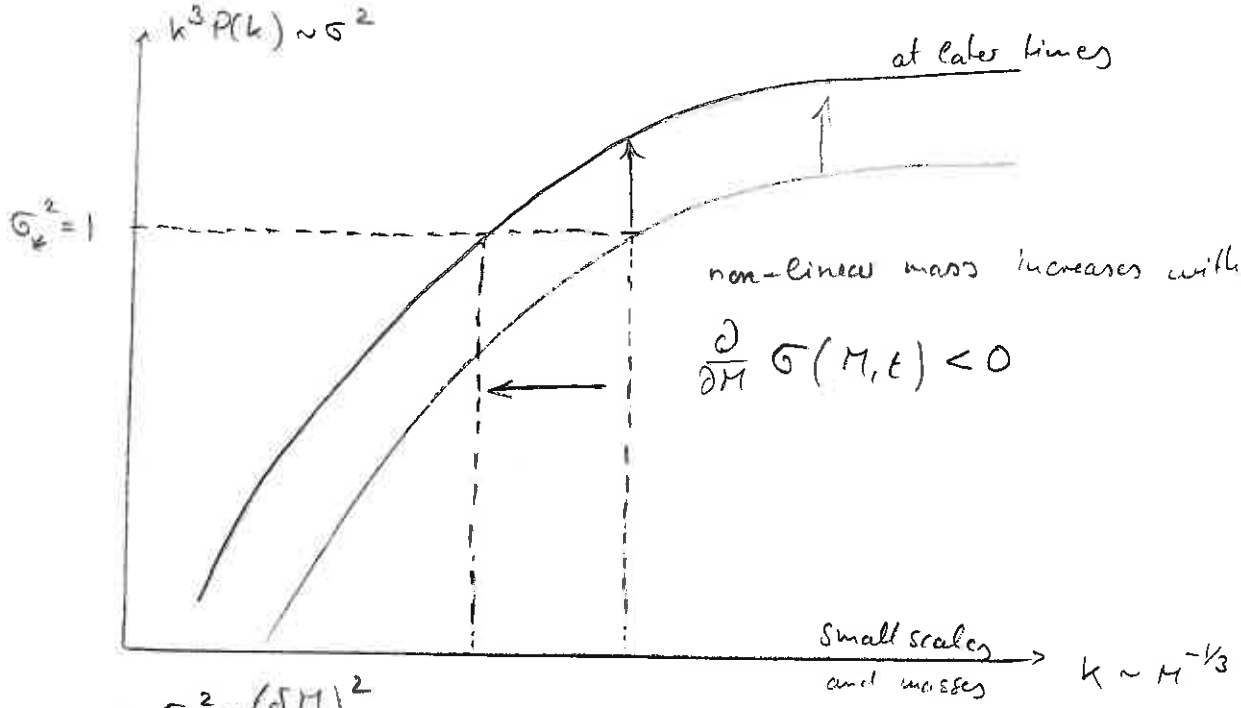
- masses and length scales are related by background density, $M = \frac{4\pi}{3} \bar{\rho} R^3 \propto k^{-3} \rightsquigarrow k \propto M^{-1/3}$, hence

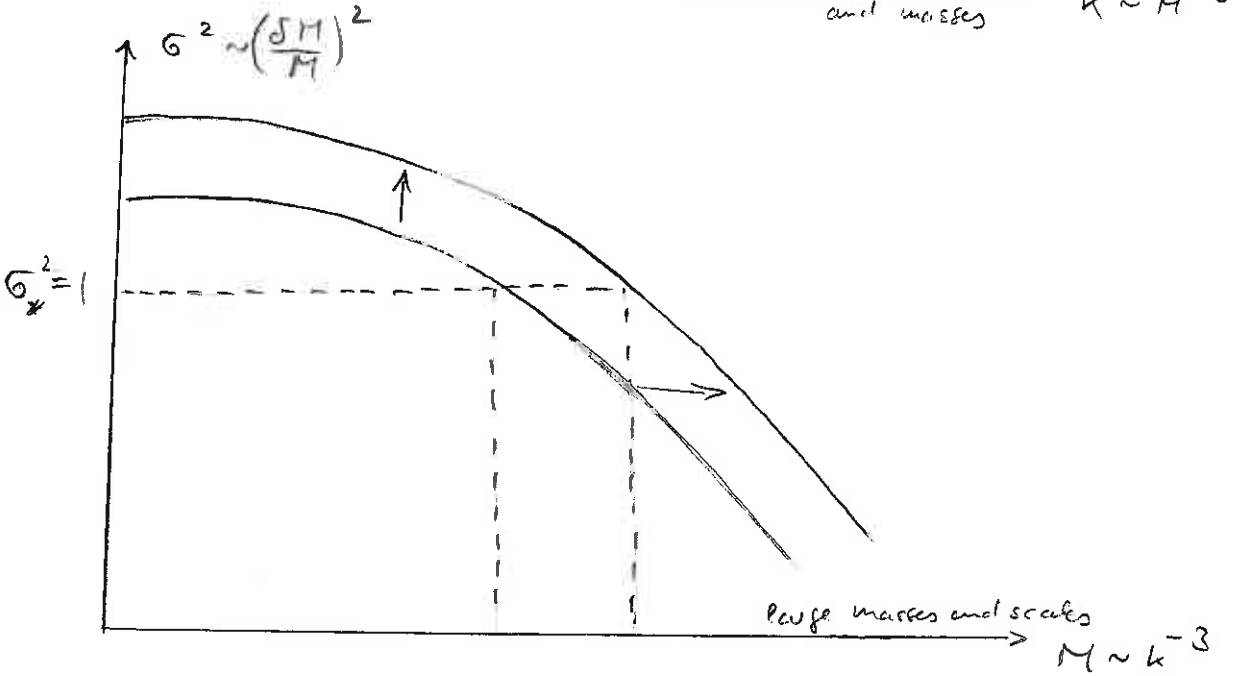
$$\sigma^2 = \left\langle \left(\frac{\delta M}{M}\right)^2 \right\rangle = \left(\frac{M}{M_*}\right)^{-1-\frac{n}{3}} \qquad (2.58)$$

- special cases

$$\sigma^2 = \begin{cases} \left(\frac{M}{M_*}\right)^{-4/3} & \text{for } n = +1 \\ \left(\frac{M}{M_*}\right)^{0} & \text{for } n = -3 \end{cases}$$

## Picture of hierarchical formation:




- k-picture:

  more power on small scales, which collapses first; structure forms "bottom-up" in ΛCDM cosmologies, i.e. hierarchical structure formation

- mass-picture:

  gravity dictates that fluctuations grow with time, growth rate depends on amount of matter, the more matter → the stronger gravity!

discussion:

- another way of looking at this: fluctuations on large scale are more subtle than fluctuations on small scales because the universe is homogeneous on the largest scales.

$$\frac{\partial}{\partial M} G(M,t) < 0 \iff \text{hierarchical formation} \tag{2.60}$$

- Is there a deep reason why fluctuations are smaller on large scales? look at potential fluctuations

$$\delta\phi \sim \frac{GM}{R}\frac{\delta M}{M} \sim G M^{2/3} \bar{\rho}^{1/3} \frac{\delta M}{M} \quad \text{since at any time } \frac{1}{R} \sim \bar{\rho}^{1/3} \tag{2.61}$$

$\sim$ unless $\frac{\delta M}{M} \propto M^{-2/3}$, $\delta\phi$ will diverge: depending on the power-law index of $\left\langle \frac{\delta M}{M} \right\rangle \propto M^{-\alpha}$, $\delta\phi$ will either diverge on large scales ($\alpha < \frac{2}{3}$) or on small scales ($\alpha > \frac{2}{3}$)

$\sim$ most natural fluctuation spectrum $\frac{\delta M}{M} \propto M^{-2/3}$ which corresponds to

Harrison–Zel'dovich–Peebles spectrum:

$\delta\phi \sim G k \delta M$ and $M \propto R^3 \propto k^{-3}$

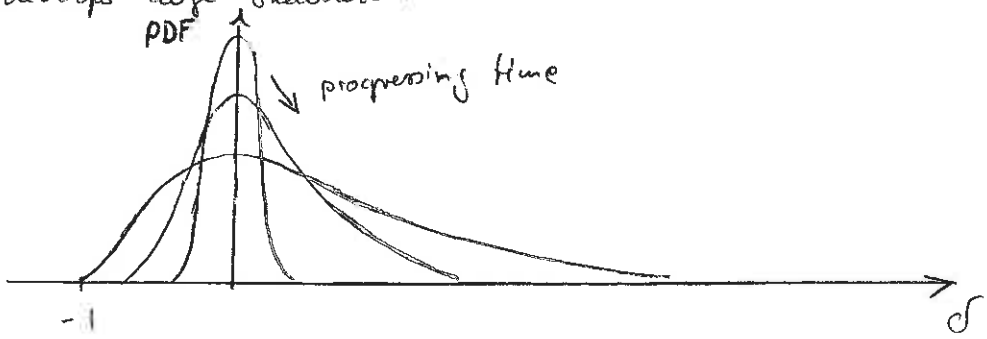
we have $\frac{\delta M}{M} \propto M^{-\left(\frac{n+3}{6}\right)} \sim \delta M \propto M^{-\left(\frac{n-3}{6}\right)} \propto k^{+3\left(\frac{n-3}{6}\right)} \propto k^{\frac{n-3}{2}}$

$\sim$ $\delta\phi \propto k^{\frac{n-3}{2}+1} \propto k^{\frac{n-1}{2}}$ $\Rightarrow$ $n = 1$ special index! (2.62)

## 2.2.4 Nonlinear Evolution

- when $\delta > 1$, perturbation theory breaks down; to fully understand the behavior afterwards (and in particular for collapsed objects) one has to perform numerical simulations; decomposing matter distribution into a large number $N$ particles, velocities in initial conditions slightly perturbed according to assumed power spectrum (at very high-$z$ where linear theory holds); solve Newton's equations of motion to determine evolution
- different numerical techniques available:
  1. direct summation of all gravitational forces of $N-1$ particles on every $N$ particles → too expensive for large $N$ since $\mathcal{O}(N^2)$-scaling of this algorithm
  2. particle-mesh (PM) algorithm: computes gravitational potential of the particle distribution on a mesh using Poisson's equation in Fourier space; gravitational forces are obtained by gradients of potential at the particles' positions. fast technique, but limited spatial resolution in Fourier space; $\mathcal{O}(N \log N)$
  3. particle-particle particle-mesh ($P^3M$): improves over PM by doing direct summation for nearby particles; $\mathcal{O}(N \log N)$ if $\#P^2$ is kept constant!
  4. tree codes combine groups of distant particles and compute the force of the group (that is approximated as a single heavy particle in the center-of-mass of the group) on the given particle; the tree is "opened" into its branches if the group subtends too large a solid angle; tree is updated along evolution; $\mathcal{O}(N \log N)$
- non-linear evolution causes density perturbation modes on different scales to couple: in linear evolution they evolve independently, mode coupling during non-linear evolution moves power from large to small scales as structure collapses; the power spectrum is increased on small scales at the expense of intermediate scales; large scales are still in the linear regime and evolve independently of each other

- even if initial density distribution is Gaussian, it necessarily develops non-Gaussian tails; this can be most easily seen as $\delta \geq -1$ is bounded below but can develop arbitrarily large positive tails; with time, the (almost) perfect Gaussian distribution becomes increasingly skewed and develops large skewness:




- once structure forms, simulations first show the formation of "pancakes", i.e. structure collapses anisotropic along one dimension of the inertial tensor (as predicted by the theory of Gaussian random fields) and then form filaments that are connected; gravitational fragmentation of filaments causes galaxy halos to form; those merge into groups which stream along filaments and are pulled by the large concentrations of mass at the intersections of filaments – galaxy clusters; those are the sites of constructive interference of long waves in the primordial fluctuations and are enhanced by gravitational collapse that yields to continuous merges of galaxy and group halos into clusters. These sit atop the mass hierarchy and present the largest gravitationally collapsed objects!

# 2.3 Spherical Collapse

## 2.3.1 Collapse of a Homogeneous Overdense Sphere

- non-linear evolution results in the collapse of overdense regions ($\delta > 0$) to form dark matter halos; we can obtain an exact analytical solution for the collapse of a spherical density perturbation; realistic density perturbations are not spherical, nevertheless this leads to useful insights into non-linear collapse in more realistic situations:
  1. the relation of the time (or redshift) at which an object collapses to its initial (linear) overdensity
  2. the final density of DM halos formed by collapse, and how this depends on the collapse redshift
- since $\delta \ll 1$ at recombination, as shown by the measured CMB $\frac{\Delta T}{T} \sim 10^{-5}$; non-linear collapse happens at $a \gg a_{rec}$, i.e. in the matter- or vacuum-dominated eras
- assumptions:
  1. spherical perturbation, initially uniform overdensity
  2. fluid has zero pressure and is collisionless (applies to DM, not baryons); later stages of collapse of baryons is different from that of DM, but since they only contribute $1/6$ to the total mass, they do not appreciably change the collapse of DM
  3. for simplicity: $\Omega = \Omega_m = 1$ (flat, matter-dominated universe); this can be generalized to cases with $\Omega_{m0} \neq 1$, $\Omega_\Lambda \neq 0$

- consider a sphere of mass $M$ and proper radius $R$; we assume that the universe outside the sphere remains spherically symmetric, so it exerts no gravitational force on the sphere; since $M = \text{const}$, we have

$$\frac{d^2R}{dt^2} = -\frac{GM}{R^2} \tag{2.63}$$

which integrates to

$$\frac{1}{2}\left(\frac{dR}{dt}\right)^2 - \frac{GM}{R} = E \tag{2.64}$$

- we consider the gravitationally bound case $E<0$, which leads to collapse; setting $R=0$ at $t=0$, we can integrate this equation to give

$$t = \int_0^R \frac{dr}{\sqrt{2\left(\frac{GM}{r}+E\right)}} = \frac{1}{\sqrt{2|E|}} \int_0^R \frac{dr}{\sqrt{\frac{2A}{r}-1}} = \frac{A}{\sqrt{2|E|}} \int_{\vartheta(0)}^{\vartheta(R)} \frac{\sin\vartheta\, d\vartheta}{\sqrt{\frac{2}{1-\cos\vartheta}-1}} \tag{2.63}$$

change of variables: $A = \frac{GM}{2|E|}$, $r = A(1-\cos\vartheta)$, $dr = A\sin\vartheta\, d\vartheta$

* denominator $= \sqrt{\frac{2-1+\cos\vartheta}{1-\cos\vartheta}} = \sqrt{\frac{(1+\cos\vartheta)2}{(1-\cos\vartheta)2}} = \sqrt{\frac{\cos^2\frac{\vartheta}{2}}{\sin^2\frac{\vartheta}{2}}} = \frac{\cos\frac{\vartheta}{2}}{\sin\frac{\vartheta}{2}}$

* $t = \frac{A}{\sqrt{2|E|}} \int \frac{\sin\frac{\vartheta}{2}\,\sin\vartheta\, d\vartheta}{\cos\frac{\vartheta}{2}} = \frac{A}{\sqrt{2|E|}} \int 2\sin^2\frac{\vartheta}{2}\, d\vartheta = \frac{A}{\sqrt{2|E|}} \int (1-\cos\vartheta)\, d\vartheta$

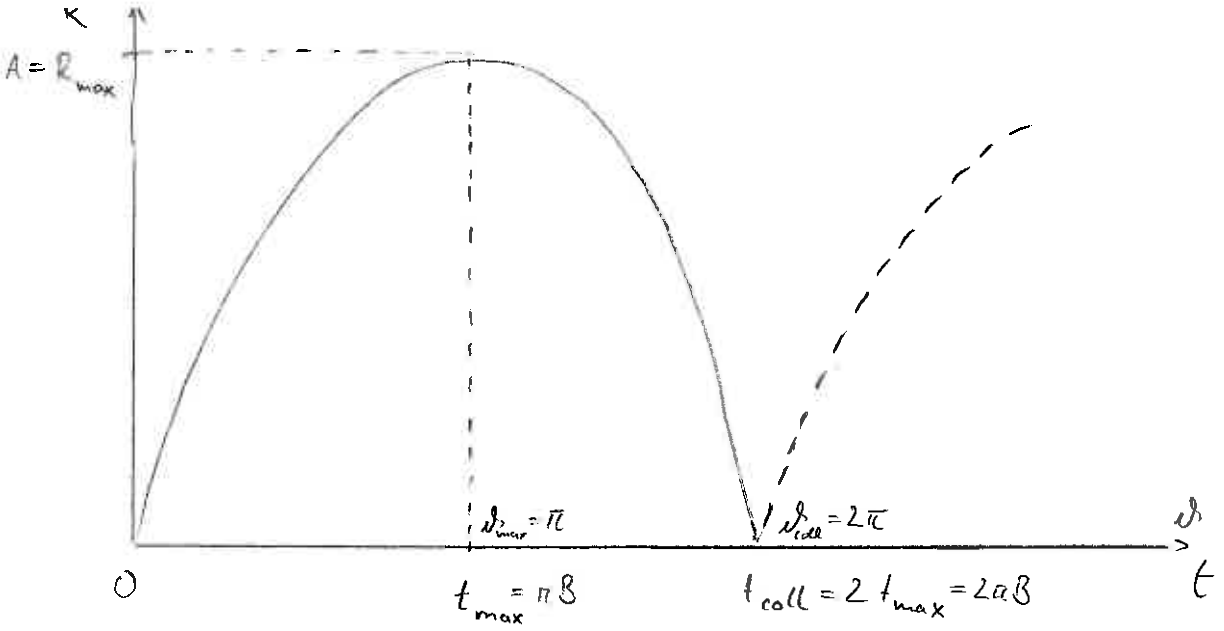
using $\sin\vartheta = 2\sin\frac{\vartheta}{2}\cos\frac{\vartheta}{2}$

$$= \frac{A}{\sqrt{2|E|}}(\vartheta - \sin\vartheta) \equiv B(\vartheta - \sin\vartheta) \quad \text{with} \quad B = \frac{A}{\sqrt{2|E|}} = \frac{GM}{(2|E|)^{3/2}}$$

$\Rightarrow$ parametric solution (cycloide):

$$R = A(1-\cos\vartheta)$$
$$t = B(\vartheta - \sin\vartheta)$$

$$A = \frac{GM}{2|E|}, \quad B = \frac{GM}{(2|E|)^{3/2}} \tag{2.64}$$




- the sphere expands from $R=0$ at $\vartheta = 0$; it reaches a maximum radius $R_{max} = A$ at $\vartheta_{max} = \pi$; it then collapses back to $R=0$ at $\vartheta_{coll} = 2\pi$; in principle it re-expands for $\vartheta > 2\pi$, but in practice other physical effects kick in and complicate things
- the corresponding times are $t_{max} = \pi B$ for the maximum (turnaround) radius and $t_{coll} = 2\pi B = 2\, t_{max}$ for collapse at $R = 0$

## 2.3.2. Connection to Linear Perturbation Theory

- The mean density inside the sphere is (using 2.64):

$$\delta = \frac{M}{\frac{4\pi}{3} R^3} = \frac{3M}{4\pi A^3} \frac{1}{(1-\cos\vartheta)^3} \qquad (2.65)$$

- In contrast, the mean density for the background $\Omega_m = 1$ universe is

$$\bar{\rho} = \frac{3H^2}{8\pi G} = \frac{3 \cdot 4}{8\pi G \cdot 9 t^2} = \frac{1}{6\pi G \cdot t^2} = \frac{1}{6\pi G B^2} \frac{1}{(\vartheta - \sin\vartheta)^2} \qquad (2.66)$$

with $H = \frac{2}{3t}$

- The actual overdensity of the sphere (in general non-linear) is given by combining these

$$1 + \delta = \frac{\rho}{\bar{\rho}} = \frac{9}{2} \frac{(\vartheta - \sin\vartheta)^2}{(1-\cos\vartheta)^3} \qquad (2.67)$$

- To make the connection to linear perturbation theory, we consider the behavior of the collapse at small $t$, which corresponds to small $\vartheta$; performing a Taylor series expansion of $\sin\vartheta$ and $\cos\vartheta$ for $\vartheta \ll 1$:

$$\sin\vartheta = \vartheta - \frac{\vartheta^3}{3!} + \frac{\vartheta^5}{5!} - \dots$$

$$\cos\vartheta = 1 - \frac{\vartheta^2}{2!} + \frac{\vartheta^4}{4!} - \dots$$

hence we obtain for $\delta(\vartheta)$ in (2.67) and $t(\vartheta)$:

$$1 + \delta = 1 + \frac{3}{20}\vartheta^2 + O(\vartheta^4) \qquad (2.68)$$

$$t = \frac{B}{6}\vartheta^3 + O(\vartheta^5) \qquad (2.69)$$

Solving for $\vartheta$ gives (using $t_{max} = \pi B$)

$$\vartheta = \left(\frac{6t}{B}\right)^{1/3} + \dots = (6\pi)^{1/3}\left(\frac{t}{t_{max}}\right)^{1/3} + \dots \qquad (2.70)$$

so that $\vartheta \ll 1$ corresponds to $t \ll t_{max}$. Substituting (2.70) into (2.68) yields

$$\delta \approx \frac{3}{20}(6\pi)^{2/3}\left(\frac{t}{t_{max}}\right)^{2/3} \qquad (t \ll t_{max}) \qquad (2.71)$$

Hence we have $\delta \propto t^{2/3} \propto a$ (since $t \propto a^{3/2}$ in EdS), which is exactly the same behavior of the growing mode of linear perturbation theory.
(The decaying mode is absent due to our choice of $R = 0$ at $t = 0$)

- For some early time $t_i \ll t_{max}$, we have $\delta = \delta_i$ with

$$\delta_i = \frac{3}{20}(6\pi)^{2/3}\left(\frac{t_i}{t_{max}}\right)^{2/3} \ll 1 \qquad (2.72)$$

<u>Corollary</u>: If sphere has uniform initial overdensity $\delta_i$ at $t_i$, then <u>all</u> interior spheres will have the same $t_{max}$, hence the sphere remains uniform as it collapses!

- there is an important distinction between
  (1) the actual overdensity $\delta$,
  (2) the overdensity <u>extrapolated</u> according to linear theory, $\delta_L$, with

$$\delta_L(t) = \delta_i \left(\frac{t}{t_i}\right)^{2/3} = \frac{3}{20}(6\pi)^{2/3}\left(\frac{t}{t_{max}}\right)^{2/3} \quad \text{for } \underline{\text{all}}\ t\ ! \tag{2.73}$$

- <u>maximum expansion (turnaround)</u> at $t = t_{max}$:

$$\delta_L(t_{max}) = \frac{3}{20}(6\pi)^{2/3} = 1.062 \tag{2.74}$$

while the actual overdensity is

$$1+\delta(t_{max}) = \frac{9\pi^2}{16} = 5.55 \tag{2.75}$$

- <u>Collapse</u> at $t = t_{coll} = 2\,t_{max}$:

$$\delta_{L,coll} \equiv \delta_L(t_{coll}) = \frac{3}{20}(12\pi)^{2/3} = 1.686 \tag{2.76}$$

In terms of the initial overdensity $\delta_i$, collapse happens at time

$$t_{coll} = t_i \left(\frac{\delta_{L,coll}}{\delta_i}\right)^{3/2} \propto \delta_i^{-3/2} \tag{2.77}$$

or redshift $1+z_{coll} = (1+z_i)\left(\frac{\delta_i}{\delta_{L,coll}}\right) \propto \delta_i$ since $t \propto a^{3/2} \propto (1+z)^{-3/2}$.

$\Rightarrow$ perturbations that are initially more overdense collapse earlier.

Generally $\delta_{L,coll} = \delta_{L,coll}(\Omega_m, \Omega_\Lambda)$ but dependence on $\Omega_m$ and $\Omega_\Lambda$ is weak.

## 2.3.3 Final Density of a Collapsed Dark Matter Halo

- According to spherical top hat collapse model, a uniform sphere collapses down to a point, i.e. infinite density, and then re-expands; realistically the sphere contains inhomogeneities, which generate random velocities in the DM during collapse → equilibrium configuration where DM velocity dispersion $\sigma$ balances its gravity; this dynamical relaxation process is called "virialization"

- We assume that final DM halo is in dynamical equilibrium and obeys the virial theorem

$$2K_f + U_f = 0 \quad (f \text{ denotes final virialized state}) \tag{2.78}$$

with $K$ denoting the total kinetic energy in random motions and $U$ is the total gravitational binding energy. We have

$$K_f = \frac{M}{2}\sigma^2 \quad \text{and} \quad U_f = -\frac{3}{5}\frac{GM^2}{R_f} \tag{2.79}$$

assuming a uniform sphere of radius $R_f$ ( this assumption is reasonable for mean density).

Hence

$$E_f = K_f + U_f = \frac{1}{2}U_f = -\frac{3}{10}\frac{GM^2}{R_f} \tag{2.80}$$

- at maximum expansion, the sphere is at rest, so $K_{max} = 0$; the total energy at maximum expansion is

$$E_{max} = U_{max} = -\frac{3}{5}\frac{GM^2}{R_{max}} \tag{2.81}$$

conservation of total energy during collapse (since DM is collisionless), so $E_f = E_{max}$

$$\rightsquigarrow \quad R_f = \frac{1}{2}R_{max} \tag{2.82}$$

- The final density is thus $\rho_f = 8\rho(t_{max})$; assuming that virialization happens at $t \approx t_{coll}$ and since $\bar{\rho} \propto t^{-2}$ and $t_{coll} = 2t_{max}$, the overdensity of the final halo is

$$1 + \delta_{vir} \equiv 1 + \delta_{coll} = \frac{\rho_{coll}}{\bar{\rho}_{max}\left(\frac{t_{coll}}{t_{max}}\right)^{-2}} = 8 \cdot 4 \cdot \left[1 + \delta(t_{max})\right]$$
$$= 18\pi^2 = 178 \tag{2.83}$$

- hence the final halo density is

$$\rho_f = (1+\delta_{vir})\,\bar{\rho}(t_{coll}) = 18\pi^2\,\bar{\rho}(t_{coll}) \qquad (2.84)$$

- $\delta_L$ and $\delta_{vir} = \Delta_{vir}$ are widely used used in cosmology to characterise DM halos; other popular choices are $\Delta_{vir} = 100, 200, 500$ (where each definition has its merits and shortcomings).

## 2.3.4 Press-Schechter Mass Function

- an important property of halos is their distribution over mass, the so-called mass function, which is sensitive to cosmological parameters and especially well suited to test for deviations from our standard model such as Gaussian initial conditions using halos of the cluster mass range; the mass function gives the number density of halos at redshift $z$ within the mass interval from $M$ to $M+dM$
- a characteristic length scale $R(M)$ can be assigned to a halo of mass $M$, which is defined as the radius of a homogeneous sphere filled with the mean cosmic matter density,

$$\frac{4}{3}\pi R^3 \rho_{cr}\Omega_m = M \Rightarrow R(M) = \left(\frac{3M}{4\pi\rho_{cr}\Omega_m}\right)^{1/3} \qquad (2.85)$$

and $\Omega_m = \Omega_m(a)$ and $\rho_{cr} = \rho_{cr}(a)$

- to obtain a prescription for a halo of mass $M$ emerging from the fluctuating density field, we filter the density contrast field on a scale $R(M)$; to this end, we use $\bar{\delta}$ defined by equation (2.47), i.e. $\delta$ convolved with a window function $W_R$ where $R = R(M)$
- we define a "nonlinear filtered mass" $M_*$ with a characteristic length scale $R(M_*) = R_*$ that has variance of $\delta_L^2$,

$$\sigma^2_{R_*} = 4\pi \int_0^\infty \frac{k^2 dk}{(2\pi)^3} P(k)\, \hat{W}^2_{R_*}(k) = \delta_L^2 \qquad (2.86)$$

- for a Gaussian random field, the probability of finding a filtered density contrast $\bar{\delta}(\vec{x})$ at a given point $\vec{x}$ in space is

$$p(\bar{\delta}, a) = \frac{1}{\sqrt{2\pi\,\sigma_R^2(a)}} \exp\left(-\frac{\bar{\delta}^2(\vec{x})}{2\,\sigma_R^2(a)}\right) \qquad (2.87)$$

where the variance depends on time (or the scale factor) through the linear growth factor $\sigma_R(a) = \sigma_R\, D_+(a)$

- Press & Schechter suggested that the probability of finding $\bar{\delta}$ at or above the linear density contrast for spherical collapse, $\bar{\delta} > \delta_L$, is equal to the fraction of the cosmic volume filled with halos of mass $M$,

$$F(M,a) = \int_{\delta_L}^{\infty} d\bar{\delta}\, p(\bar{\delta}, a) = \frac{1}{2}\,\frac{2}{\sqrt{\pi}} \int_{\frac{1}{\sqrt{2}}\frac{\delta_L}{\sigma_R(a)}}^{\infty} dx\, e^{-x^2} = \frac{1}{2}\,\mathrm{erfc}\left(\frac{\delta_L}{\sqrt{2}\,\sigma_R(a)}\right) \qquad (2.88)$$

substitution $x = \frac{1}{\sqrt{2}}\frac{\bar{\delta}}{\sigma_R(a)}$, $dx = \frac{1}{\sqrt{2}\,\sigma_R(a)}\, d\bar{\delta}$

where $\mathrm{erfc}(x)$ is the complementary error function; as we can see, this shows that the fraction of cosmic volume filled with halos of a fixed mass $M$ depends sensitively on $\sigma_R(a)$, i.e. the strength of mass/density fluctuations on a certain scale

- the distribution of halos over masses $M$ is simply $\frac{\partial F(M)}{\partial M}$, so we have to relate $M$ to $\sigma_R$ by means of the characteristic radius $R(M)$,

$$\frac{\partial}{\partial M} = \frac{d\sigma_R(a)}{dM}\,\frac{\partial}{\partial \sigma_R(a)} = \frac{d\sigma_R}{dM}\,\frac{\partial}{\partial \sigma_R} \qquad (2.89)$$

where we inserted $\sigma_R$ at the present epoch (as this expression is valid for any epoch)

- using the identity $\frac{d}{dx}\,\mathrm{erfc}(x) = -\frac{2}{\sqrt{\pi}}\, e^{-x^2}$ we obtain (2.90)

$$\frac{\partial F(M)}{\partial M} = \frac{d\sigma_R}{dM}\,\frac{\partial F}{\partial \sigma_R} = \sigma_R \frac{d\ln\sigma_R}{dM}\,\frac{1}{\sqrt{\pi}} \exp\left(-\frac{\delta_L^2}{2\,\sigma_R^2\, D_+^2(a)}\right) \cdot \left(-\frac{\delta_L}{\sqrt{2}\,\sigma_R\, D_+(a)}\right)$$

$$= +\frac{1}{\sqrt{2\pi}}\,\frac{\delta_L}{\sigma_R\, D_+(a)} \left|\frac{d\ln\sigma_R}{dM}\right| \exp\left(-\frac{\delta_L^2}{2\,\sigma_R^2\, D_+^2(a)}\right) \qquad (2.91)$$

and the "+" arises by inserting the condition for hierarchical structure growth,

$$\frac{d\ln\sigma}{dM} < 0$$

- but there is a serious problem with the normalization of the mass function,

$$\int_0^\infty \frac{\partial F(M)}{\partial M}\, dM = \frac{\delta_L}{\sqrt{2\pi}\, D_+(a)} \int_0^\infty \frac{d\sigma_R}{\sigma_R^2} \exp\left(-\frac{\delta_L^2}{2\sigma_R^2 D_+^2(a)}\right)$$

substitution: $\sigma_R^{-1} = x$, $dx = -\frac{d\sigma_R}{\sigma_R^2}$

$$= \frac{\delta_L}{\sqrt{2\pi}\, D_+(a)} \int_0^\infty dx \exp\left(-\frac{\delta_L^2}{2 D_+^2(a)} x^2\right) = \frac{\delta_L}{\sqrt{2\pi}\, D_+(a)} \frac{\sqrt{\pi}}{2} \sqrt{\frac{2 D_+^2(a)}{\delta_L^2}} = \frac{1}{2} \qquad (2.92)$$

i.e. it is too small by a factor 2 (for reasons quite subtle). For the moment, we will arbitrarily multiply $F(M)$ by 2 and return to this problem later

- to get the mass function, we divide by comoving volume, $\frac{M}{\rho_0}$ and obtain the comoving number density of halos within $[M, M+dM]$,

$$n(M,a)\, dM = \frac{\rho_0}{M} \frac{\partial F}{\partial M}\, dM = \sqrt{\frac{2}{\pi}} \frac{\rho_0 \delta_L}{\sigma_R D_+(a)} \left|\frac{d\ln\sigma_R}{dM}\right| \exp\left(-\frac{\delta_L^2}{2\sigma_R^2 D_+^2(a)}\right) \frac{dM}{M} \qquad (2.93)$$

- the Press-Schechter mass function generally describes the abundance of halos well; only at the high-mass end there are significant deviation to large-scale cosmological simulations that can be modeled by taking into account ellipsoidal halo collapse, which is more realistic as opposed to spherical collapse

- for power-law power spectra the mass function can be considerably simplified; recall that the variance can then be written (2.58)

$$\sigma^2 = \sigma_*^2 \left(\frac{M}{M_*}\right)^{-(1+\frac{n}{3})} \equiv \sigma_*^2\, m^{-(1+\frac{n}{3})}$$

where we introduced $m \equiv \frac{M}{M_*}$

- first
$$\frac{d\ln\sigma_R}{dM} = \frac{1}{2}\frac{d\ln\sigma_R^2}{dM} = \frac{1}{2\sigma_R^2}\frac{d\sigma_R^2}{dM} = \frac{-1}{2\sigma_R^2}\sigma_*^2\left(1+\frac{n}{3}\right)\frac{1}{M_*}\left(\frac{M}{M_*}\right)^{-2-\frac{n}{3}}$$

$$= -\frac{\sigma_*^2}{2\sigma_R^2}\left(1+\frac{n}{3}\right)\frac{1}{M}\left(\frac{M}{M_*}\right)^{-(1+\frac{n}{3})} = \frac{-1}{2M}\left(1+\frac{n}{3}\right) \qquad (2.94)$$

hence the Press-Schechter mass function can be written as

$$n(M,a) = \sqrt{\frac{2}{\pi}}\frac{\rho_0\delta_L}{D_+(a)}\frac{1}{\sigma_*}\left(\frac{M}{M_*}\right)^{(1+\frac{n}{3})\frac{1}{2}}\cdot\frac{\left(1+\frac{n}{3}\right)}{2M^2}\exp\left(-\frac{\delta_L^2}{2D_+^2(a)\sigma_*^2 m^{-(1+\frac{n}{3})}}\right)$$

$$= \sqrt{\frac{2}{\pi}}\frac{\rho_0\delta_L}{D_+(a)}\frac{1}{\sigma_*}\, m^{\frac{1}{2}+\frac{n}{6}-2}\cdot\frac{1+\frac{n}{3}}{2M_*^2}\exp\left(-\frac{\delta_L^2}{2D_+^2(a)\sigma_*}\, m^{1+\frac{n}{3}}\right)$$

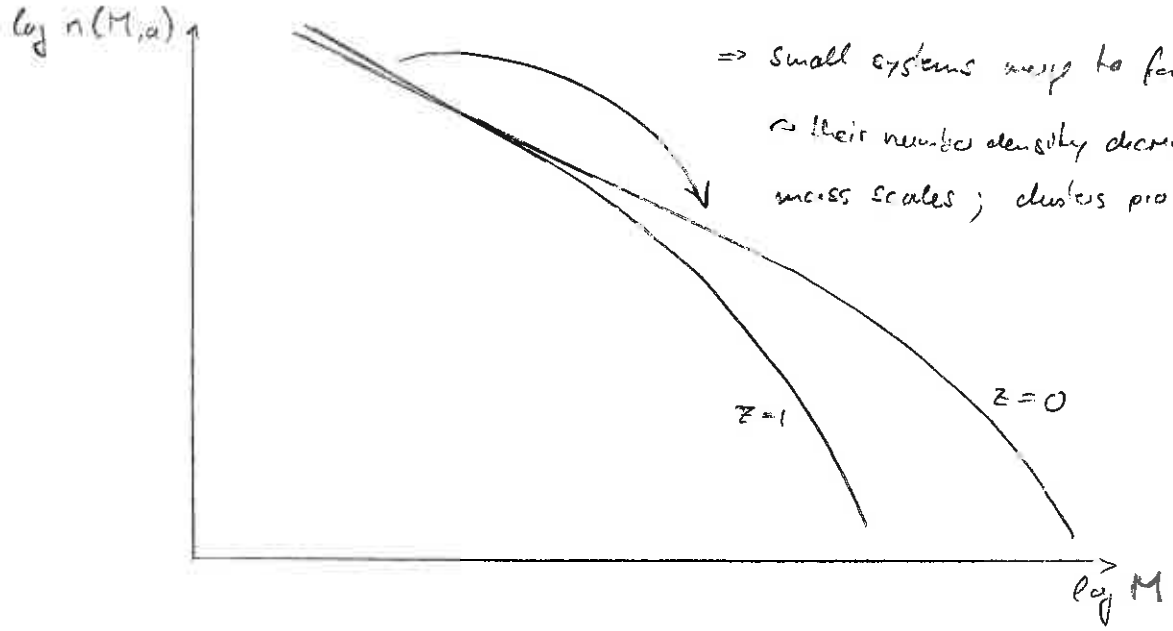
with $\alpha = \left(1+\frac{n}{3}\right)\frac{1}{2}$

$$n(m,a)\,dm = \sqrt{\frac{2}{\pi}}\frac{\rho_0\delta_L\,\alpha}{D_+(a)\sigma_*M_*}\cdot m^{\alpha-2}\exp\left(-\frac{\delta_L^2}{2D_+^2(a)\sigma_*}\, m^{2\alpha}\right)dm \qquad (2.95)$$

---

hence $n(M) \propto M^{\alpha-2}\, e^{M^{2\alpha}}$

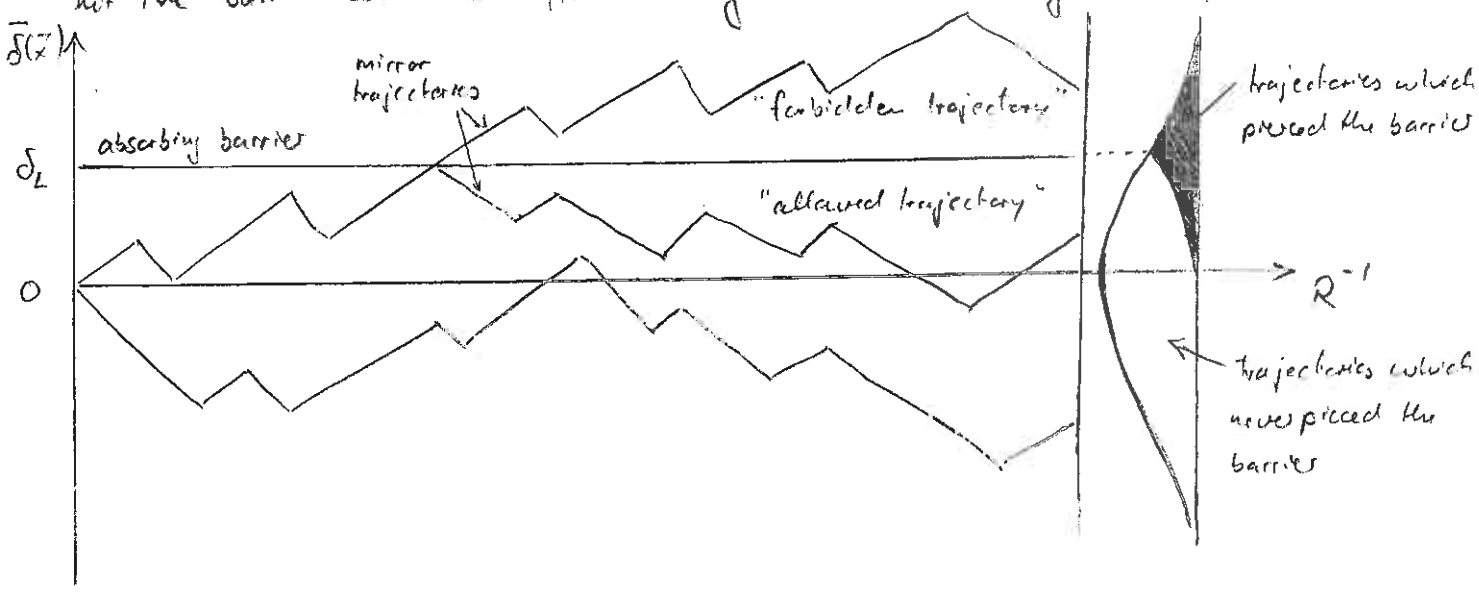
- there are two regimes: (1) a power-law mass function at low masses
with $n(M) \propto M^{\alpha-2} \propto M^{\frac{n}{6}-\frac{3}{2}} \sim M^{-2}$ for $n\sim -3$, $\alpha\sim 0$
(2) an exponential cutoff at large (cluster masses)




⇒ small systems merge to form larger systems
↝ their number density decreases at (dwarf) galaxy mass scales; clusters probe high-M tail!

## 2.3.5 Correct Normalization of the Press-Schechter Mass Function

- the solution to the normalization problem is given by interpreting the statistics of halo formation in terms of a random walk
- given the density contrast field $\delta(\vec{x})$, we center a large sphere on a point $\vec{x}$ and decrease the radius $R$ of the sphere; the density contrast $\bar{\delta}$ averaged within $R$ is monitored as a function of $R$; by choosing a window function $W_R$ in the definition (2.47) of $\bar{\delta}$ such that its Fourier transform has a sharp cutoff in $k$ space, $\bar{\delta}$ will undergo a random walk because decreasing $R$ corresponds to adding shells in $k$ space which are independent of the modes that are already included
- $\bar{\delta}$ is thus following a random trajectory; a halo is expected to be formed at $\vec{x}$ if $\bar{\delta}(\vec{x}) \geq \delta_L$ for some $R$; if $\bar{\delta}(\vec{x}) < \delta_L$ at $R$ it may well exceed $\delta_L$ at some smaller $R$ or conversely, if $\bar{\delta}(\vec{x}) > \delta_L$ at $R$, it may drop below $\delta_L$ at a smaller $R$
- to correctly determine the halo number, we must account for all points in space which are parts of halos; in particular, $\vec{x}$ is counted as being part of a halo if there is some $R$ for which $\bar{\delta}(\vec{x}) \geq \delta_L$
- we have to introduce an "absorbing barrier" at $\delta_L$ in our random walks such that points $\vec{x}$ with trajectories $\bar{\delta}(\vec{x})$ versus (decreasing) $R$ which hit the barrier are removed from counting them as not being parts of halos



- a trajectory meeting the boundary has equal probability for propagating above or below; we call a trajectory "forbidden" in the boundary of a random walk with an absorbing barrier, if it continues above it; for every such a "forbidden" trajectory there will be an "allowed" mirror trajectory that continues below it; hence, for every trajectory reaching a point $\bar{\delta} < \delta_L$ exclusively along "allowed" trajectories, there is a path reaching its mirror point with respect to the line $\bar{\delta} = \delta_L$ exclusively along "forbidden" trajectories, and vice versa (see sketch)
- thus, the probability for reaching a point $\bar{\delta} < \delta_L$ along "allowed" traj. exclusively below the barrier is the probability for reaching it along any trajectory minus the probability for reaching its mirror point $\delta_L + (\delta_L - \bar{\delta}) = 2\delta_L - \bar{\delta}$ along "forbidden" trajectories,

$$p_S(\bar{\delta})\,d\bar{\delta} = \frac{1}{\sqrt{2\pi}\,\sigma_R}\left[\exp\left(-\frac{\bar{\delta}^2}{2\sigma_R^2}\right) - \exp\left(-\frac{(2\delta_L-\bar{\delta})^2}{2\sigma_R^2}\right)\right]d\bar{\delta} \qquad (2.96)$$

where $\sigma_R = \sigma_R(a)$ is the variance of $\bar{\delta}$ on the scale R.

- equation (2.96) is the probability distribution for the averaged density contrast to lie within $[\bar{\delta}, \bar{\delta}+d\bar{\delta}]$ and not to exceed $\delta_L$ when averaged on any scale; we are interested in the complement, i.e. the probability for $\bar{\delta}$ to exceed $\delta_L$ on some scale which is thus

$$1 - P_S = 1 - \int_{-\infty}^{\delta_L} d\bar{\delta}\, p_S(\bar{\delta}) =$$

$$= 1 - \frac{1}{\sqrt{2\pi}\,\sigma_R}\left[\int_{-\infty}^{\delta_L} d\bar{\delta}\,\exp\left(-\frac{\bar{\delta}^2}{2\sigma_R^2}\right) - \int_{-\infty}^{\delta_L} d\bar{\delta}\,\exp\left(-\frac{(2\delta_L-\bar{\delta})^2}{2\sigma_R^2}\right)\right]$$

substitution $y = \frac{\bar{\delta}}{\sqrt{2}\sigma_R}$, $dy = \frac{d\bar{\delta}}{\sqrt{2}\sigma_R}$; $\quad x = \frac{2\delta_L - \bar{\delta}}{\sqrt{2}\sigma_R}$, $\quad dx = -\frac{d\bar{\delta}}{\sqrt{2}\sigma_R}$

$$= 1 - \frac{1}{2}\frac{2}{\sqrt{\pi}}\left[\int_{-\infty}^{\delta_L/(\sqrt{2}\sigma_R)} dx\, e^{-x^2} + \int_{-\infty}^{\delta_L/(\sqrt{2}\sigma_R)} dx\, e^{-x^2}\right]$$

$$1 - P_S = 1 - \mathrm{erf}\left(\frac{\delta_L}{\sqrt{2}\,\sigma_R}\right) = \mathrm{erfc}\left(\frac{\delta_L}{\sqrt{2}\,\sigma_R}\right) \qquad (2.97)$$

without the factor $\frac{1}{2}$ in (2.88); the remaining part of the derivation proceeds as before

## 2.3.6 Halo Density Profiles

- generally, a self-gravitating system of particles does not have an equilibrium state; the virial theorem demands that its total energy $E$ is minus half its potential energy $U$,

$$E = T + U \quad , \text{ where } T \text{ is the kinetic energy}$$

$$2T + U = E + T = 0$$

$$\rightsquigarrow T = -E = -\frac{U}{2} \qquad (2.98)$$

$\rightsquigarrow U < 0$ for self-bound systems and any inevitable energy loss, e.g. through ejection of a body by means of three-body encounters makes the potential energy become more negative, i.e. the halo more tightly bound, which in turn increases its energy loss since the dynamical timescales are also reduced by this contraction according to

$$v^2 = \frac{GM}{R_g} \quad \text{where } R_g \text{ is the gravitational radius.} \qquad (2.99)$$

$$t_{dyn} \sim \left(\frac{R_g^3}{GM}\right)^{1/2} \sim (G\rho)^{-1/2} \qquad (2.100)$$

$\Rightarrow$ any halo density profile must thus reflect a potentially long-lived, but transient state

## 2.3.6.1 Isothermal Sphere

- complementary to properties like halo mass and their distributions in mass and redshift are internal halo properties such as the DM density profile; since DM dominates the halo mass, its internal distribution is crucial to shape the halo potential, which then shapes the gas distribution and the related observables
- we start with a simple model, the "isothermal sphere", which is a spherically-symmetric, self-gravitating system of non-interacting particles whose kinetic energy is characterized by a constant "temperature" $T = \frac{m}{k}\sigma^2$ (where $\sigma$ denotes the 3D velocity dispersion)
- Eulers equation of hydrostatic equilibrium (HSE) $\left[p = n\,kT = \frac{\rho}{m}\,kT\right]$:

$$\vec{\nabla} p = -\rho\,\nabla\phi \quad \Rightarrow \quad \frac{dp}{dr} = \frac{kT}{m}\frac{d\rho}{dr} = -\rho\,\frac{G\,M(r)}{r^2} \tag{2.101}$$

$$\text{and} \quad M(r) = \int_0^r 4\pi r^2 \rho\, dr \quad \Rightarrow \quad \frac{dM}{dr} = 4\pi r^2 \rho \tag{2.102}$$

- multiply by $\frac{r^2 m}{\rho\, kT}$ to get

$$r^2 \frac{1}{\rho}\frac{d\rho}{dr} = -\frac{Gm}{kT}\,M(r) \tag{2.103}$$

- differentiate to get

$$\frac{d}{dr}\left(r^2 \frac{d\ln\rho}{dr}\right) = -4\pi\,\frac{Gm}{kT}\,r^2\rho \tag{2.104}$$

- one solution is singular, the so-called "singular isothermal sphere" (SIS):

assume $\rho = C\,r^{-\alpha}$:

$$\ln\rho = \ln C - \alpha \ln r \quad \Rightarrow \quad \frac{d\ln\rho}{dr} = -\frac{\alpha}{r}$$

$$\frac{d}{dr}(-\alpha r) = -\alpha = -4\pi\,\frac{Gm}{kT}\,C\,r^{2-\alpha}$$

$$\sim \alpha = 2 \quad \text{and} \quad C = \frac{kT}{2\pi G m}$$

$$\Rightarrow \quad \rho(r) = \frac{kT}{2\pi G m}\,r^{-2} = \frac{\sigma^2}{2\pi G\, r^2} \tag{2.105}$$

- the mass interior to $r$ is

$$M(r) = \frac{2\sigma^2}{G} r \qquad (2.106)$$

and the circular velocity is

$$v_c^2 = \frac{GM(r)}{r} = 2\sigma^2 \qquad (2.107)$$

- the solution to (2.104) depends on boundary conditions; it turns out that there is a second solution, which has a finite central density $\rho_0$

- to find this solution, first we have to identify a characteristic length scale so that we can obtain a general solution in terms of dimensionless variables; the dimensional quantities in (2.104) are $G$, $\rho$ and the combination $\sigma^2 = \frac{kT}{m}$; the only combination of these quantities that yields a length scale is $\frac{\sigma}{\sqrt{G\rho_0}}$; physically, this represents the typical distance a particle travels in the central dynamical time; we define the King radius (a.k.a. core radius)

$$r_0 \equiv \left( \frac{9\sigma^2}{4\pi G \rho_0} \right)^{1/2} \qquad (2.108)$$

and the dimensionless variables

$$\tilde{\rho} = \frac{\rho}{\rho_0} \quad , \quad \tilde{r} = \frac{r}{r_0}$$

HSE equation (2.104) becomes

$$\frac{d}{d\tilde{r}} \left( \tilde{r}^2 \frac{d \ln \tilde{\rho}}{d\tilde{r}} \right) = -9 \tilde{r}^2 \tilde{\rho} \qquad (2.109)$$

- the (numerical) solution is obtained by integrating outwards from $\tilde{r} = 0$ with the central boundary conditions

$$\tilde{\rho}(0) = 1 \quad , \quad \frac{d\tilde{\rho}}{d\tilde{r}} = 0$$

(the second condition is necessary since $M(\tilde{r})$ vanishes at $\tilde{r} = 0$.)

- at $\tilde{r} \lesssim 1$, $\tilde{\rho}$ is constant

for $\tilde{r} \lesssim 3$, the King profile, $\tilde{\rho} = (1 + \tilde{r}^2)^{-3/2}$ provides a good approximation (2.110)

for $\tilde{r} \gtrsim 3$, $\tilde{\rho} = \frac{2}{9} \tilde{r}^{-2}$, i.e. the SIS is the asymptotic solution at large $\tilde{r}$

- note that by defining characteristic dimensionless variables, we have reduced the family of solutions with different densities and temperatures to a single solution for appropriately scaled variables

- however, the cored isothermal sphere is unphysical as it has an infinite mass; to get a halo of finite mass, we must truncate it at radius $r_t$ by confining it with an external "pressure" that in reality is provided by accretion of mass

## 2.3.6.2 Navarro-Frenk-White (NFW) density profile

- numerical simulations of halo formation in the CDM model consistently give profiles that are well fitted by the NFW profile

$$\rho(r) = \frac{\rho_s}{x(1+x)^2}, \quad x = \frac{r}{r_s} \tag{2.111}$$

at small scales $\rho \propto r^{-1}$, while at scales $r \gg r_s$ we have $\rho \propto r^{-3}$

- the enclosed mass is given by

$$M(r) = \int_0^r 4\pi r'^2 \rho \, dr' = 4\pi \rho_s r_s^3 \int_0^{r/r_s} x(1+x)^{-2} dx$$

$\frac{x}{(1+x)^2} = \frac{1}{1+x} - \frac{1}{(1+x)^2}$, hence we have

$$M(r) = 4\pi \rho_s r_s^3 \left[ \ln\left(1+\frac{r}{r_s}\right) + \frac{1}{1+\frac{r}{r_s}} - 1 \right]$$

$$= 4\pi \rho_s r_s^3 \left[ \ln\left(1+\frac{r}{r_s}\right) + \frac{\frac{r}{r_s}}{1+\frac{r}{r_s}} \right] \tag{2.112}$$

$M(r)$ rises as $r^2$ for small $r < r_s$ and diverges logarithmically for $r \to \infty$; the divergence is not a fundamental problem because the halo profile becomes invalid once $\rho(r) < \rho_0$, the background density; in practice, the assumption of spherical symmetry starts to break down earlier, around the virial radius

- the virial radius is often defined as the radius enclosing a mean overdensity of 200 times the critical density, i.e. the contribution of $\Lambda$ to the critical density is included in the estimate of the reference density; since $\Lambda$ is constant, this has the advantage that the mass of a halo that decoupled from the cosmic expansion can only grow by accretion but remains constant otherwise (for $a > a_{eq,\Lambda}$); the factor 200 is a rough approximation to the density contrast of $18\pi^2 \approx 178$ expected at virialization in the spherical collapse model; this implies

$$M_{200}\left(\frac{4\pi}{3} r_{200}\right)^{-1} = 200\,\rho_{cr} = 200\,\frac{3H^2(a)}{8\pi G}$$

$$\Rightarrow \quad r_{200} = \left(\frac{G M_{200}}{100\,H^2(a)}\right)^{1/3} \qquad (2.113)$$

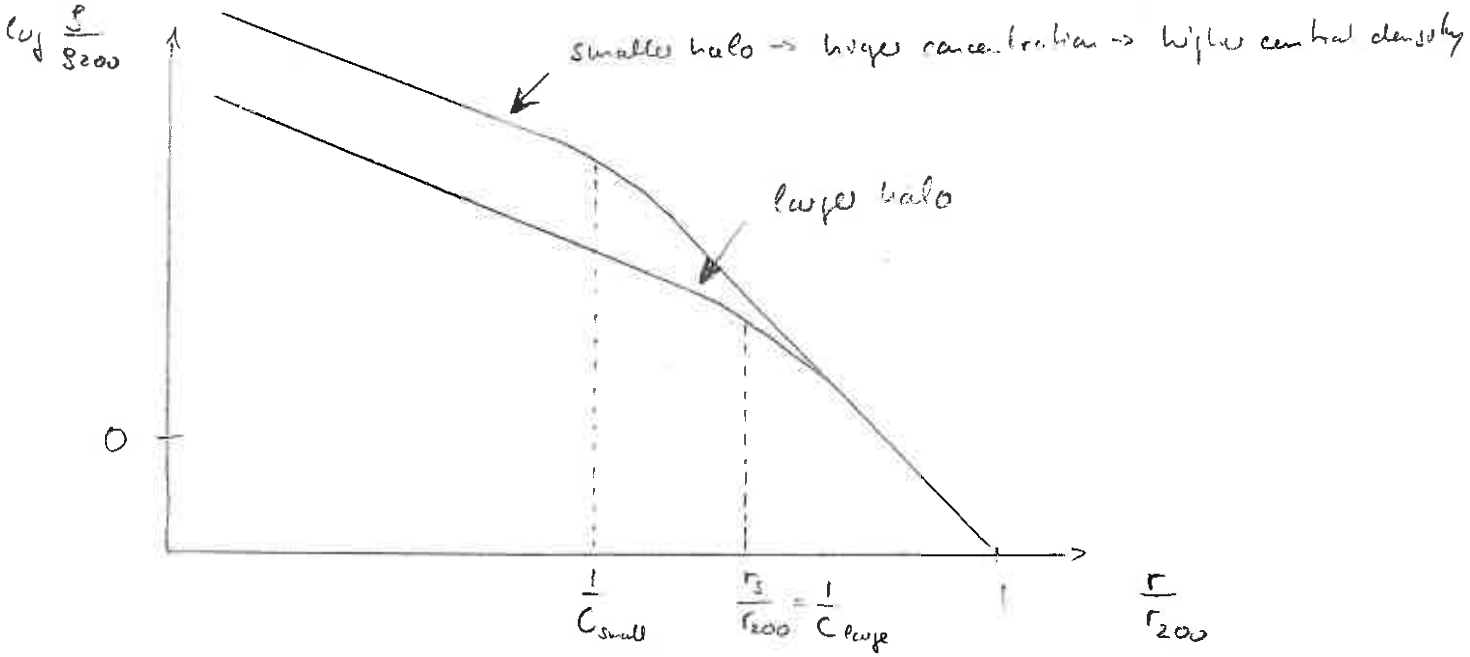
- other frequent definitions consider the radius enclosing a mean overdensity of 200 times the mean density (i.e. without the $\Lambda$-contribution),

$$M_{200m}\left(\frac{4\pi}{3} r_{200m}\right)^{-1} = 200\,\rho_{cr}(a)\,\Omega_m(a) \qquad (2.114)$$

this definition has the advantage that halos of the same mass, but at different redshifts show the same amount of kinetic pressure contribution, or velocity anisotropy as a function of radius, i.e. this definition is close to a dynamical definition of the "virial" radius; but it requires the knowledge of the ab initio unknown cosmological parameter $\Omega_{m,0}$ and has the property that the halo mass increases at late times because of the redshift dilution of the mean matter density as $\frac{\rho_{m0}}{a^3}$

- sometimes, people use a redshift dependent $\Delta(a)$ from the solution of a spherical top-hat perturbation at the time of collapse; while this property is easily calculatable in simulations, the collapse time of a cluster is inaccessible in observations which jeopardizes detailed comparisons of observations and theory

- the ratio $c := \frac{r_{200}}{r_s}$ is called concentration of the halo; it turns out to be a function of halo mass and redshift and to depend on cosmological parameters; generally, c is higher in the earlier halos; this reflects the hierarchical growth of halos and implies that smaller halos form earlier when the mean background density was higher and hence is reflected by the higher central densities of these "higher concentrated" halos:




- given the halo mass M, the (virial) radius is given by (2.113), the concentration parameter gives $r_s = \frac{r_{200}}{c}$ and the scale density, $\rho_s$, is then determined from (2.112) by the requirement that $M(r_{200}) = M$; thus the profile (2.111) is essentially a simple-parameter family, determined by its mass M
- the exact cause of the shape is not clear, but the inner slope seems to be related to the slope of the correlation function of the initial conditions and arises from a regime of "fast" accretion, while the outer slope corresponds to a slower accretion regime and reflects the property of a transient, i.e. non-equilibrium profile

# 3 Evolution of the Baryonic Component

## 3.1 Non-radiative Physics

### 3.1.1 Adiabatic Processes and Entropy

- before we enter a detailed discussion on the evolution of the baryonic component, we review a few basic concepts; first, we start with the first law of thermodynamics (energy conservation):

$$d\varepsilon = -P dV + dq \quad \Rightarrow \quad dq = d\varepsilon + P dV \tag{3.1}$$

$V \equiv \rho^{-1}$ = specific volume $[\mathrm{cm^3\,g^{-1}}]$

$dq \equiv T ds$ = heat change per unit mass $[\mathrm{erg\,g^{-1}}]$

$s \equiv$ specific entropy $[\mathrm{erg\,g^{-1}\,K^{-1}}]$

$\varepsilon =$ specific internal energy $[\mathrm{erg\,g^{-1}}]$

- the specific heat at constant volume

$$c_V \equiv \left(\frac{\partial q}{\partial T}\right)_V \qquad [\mathrm{erg\,g^{-1}\,K^{-1}}] \tag{3.2}$$

is the amount of heat that must be added to raise the temperature of 1g of gas by 1K; at constant volume, $dq = d\varepsilon$, and if $\varepsilon$ depends only on temperature (and not density), $\varepsilon(V,T) = \varepsilon(T)$. Then

$$c_V \equiv \left(\frac{\partial q}{\partial T}\right)_V = \left(\frac{\partial \varepsilon}{\partial T}\right)_V = \frac{d\varepsilon}{dT} \tag{3.3}$$

Implying

$$dq = c_V dT + P dV \tag{3.4}$$

- the equation of state of an ideal gas is given by

$$\varepsilon = \frac{1}{\gamma - 1}\frac{P}{\rho} \tag{3.5}$$

- in the case of a mono-atomic gas (which is to a very good approximation the ionised plasma of the intra-cluster medium, ICM), we have $\gamma = \frac{5}{3}$

$$\varepsilon = \frac{3}{2}\frac{nkT}{\rho} = \frac{3}{2}\frac{kT}{m} \quad (m = \mu m_p) \tag{3.6}$$

hence

$$c_v = \frac{d\varepsilon}{dT} = \frac{1}{(\gamma-1)}\frac{k}{m} \xrightarrow[\text{gas}]{\text{mono-atomic}} \frac{3}{2}\frac{k}{m} \tag{3.7}$$

- the total differential of the EOS of an ideal gas is

$$d\varepsilon = \frac{1}{\gamma-1}\left(\frac{dP}{\rho} - \frac{P}{\rho^2}d\rho\right) \tag{3.8}$$

for adiabatic ($dq = ds = 0$) changes, we can combine this with the first law of thermodynamics

$$d\varepsilon = -PdV = \frac{P}{\rho^2}d\rho \tag{3.9}$$

and find (after multiplying by $\frac{\rho}{P}$)

$$\frac{1}{\gamma-1}\left(\frac{dP}{P} - \frac{d\rho}{\rho}\right) = \frac{d\rho}{\rho}$$

implying

$$\frac{dP}{P} = \gamma\frac{d\rho}{\rho} \Rightarrow P = P_0\left(\frac{\rho}{\rho_0}\right)^\gamma \equiv K\rho^\gamma \tag{3.10}$$

a polytropic equation of state ($P \propto \rho^\gamma$), which defines the quantity

$$K \equiv \frac{P}{\rho^\gamma} = \frac{nkT}{mn\rho^{\gamma-1}} = \frac{kT}{m\rho^{\gamma-1}} \quad \text{and} \quad K_e \equiv \frac{kT_e}{n_e^{\gamma-1}} \propto K \tag{3.11}$$

that is constant upon adiabatic changes and in the context of galaxy clusters often referred to as "entropy"; in particular its cousin $K_e$ can be conveniently computed from the X-ray "observables" $kT_e$ and $n_e$

and has typical values of

$$K_e \sim 100\left(\frac{kT_e}{1\,\text{keV}}\right)\left(\frac{n_e}{10^{-3}\,\text{cm}^{-3}}\right)^{-2/3}\,\text{keV cm}^2 \tag{3.12}$$

- to relate this dusties "entropy" to the proper thermodynamic entropy we start with

$$d\mathcal{E} = -P\,dV + T\,dS$$

and consider adding or removing heat at constant $\rho$ ($dV = 0$)

$$T\,dS = d\mathcal{E} = c_V\,dT$$

implying

$$dS = c_V \frac{dT}{T} \quad \Rightarrow \quad S = c_V \ln T + \text{const.}$$

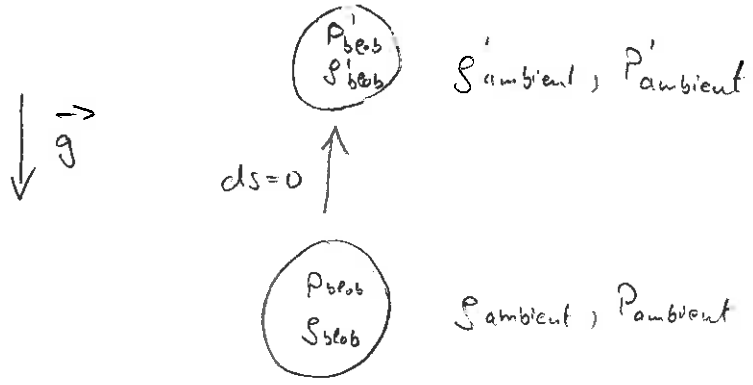
- since $P \propto T$ at constant $\rho$, this implies that $S = c_V \ln P + \text{const.}$; we just derived that adiabatic changes keep $P\rho^{-\gamma}$ constant, so these must be lines of constant entropy (in the $P$-$\rho$ plane); hence for a single species gas, we have

$$S = c_V \ln\left(P\rho^{-\gamma}\right) + \text{const.} = \frac{k}{(\gamma-1)m} \ln K + \text{const} \qquad (3.13)$$

$$S = c_V \ln\left(\frac{K}{K_0}\right) \quad \Rightarrow \quad K = K_0 \exp\left(\frac{S}{c_V}\right) \qquad (3.14)$$

## 3.1.2 Schwarzschild Criterion for Convective Instability (no Magnetic Fields)

- fluid motion driven by thermal gradients (thermal convection) is a common and important phenomenon in clusters; rather than doing a full perturbation analysis of the hydrodynamic equations, we employ simple arguments




we are considering an upward displacement of a small volume element (a "blob") in a gravitational potential (pointing downwards); the ambient gas

is in hydrostatic equilibrium which causes the ambient pressure and density to fall off at larger heights; since the motion of the blob is adiabatic (isentropic), the change in density of the blob is related to the change of its surroundings by

$$\left(d\rho\right)_{blob} = \left(\frac{\partial \rho}{\partial p}\right)_s dp \tag{3.15}$$

the corresponding change in density of the ambient medium is

$$\left(d\rho\right)_{ambient} = \left(\frac{\partial \rho}{\partial p}\right)_s dp + \left(\frac{\partial \rho}{\partial s}\right)_p ds \tag{3.16}$$

where $dp$ and $ds$ are the difference of the pressure and the specific entropy of the medium at the new position that arises as a result of mechanical and thermal balance; the blob will continue to move under the influence of buoyancy, i.e. it will continue to fall (rise) if its density increase (decrease) exceeds (falls short of) that of the ambient medium

$$\begin{aligned} \left(d\rho\right)_{blob} &> \left(d\rho\right)_{ambient} \quad \text{for instability of a downward displacement} \\ \left(d\rho\right)_{blob} &< \left(d\rho\right)_{ambient} \quad \text{-"- upward -"-} \end{aligned} \tag{3.17}$$

Since we assume small blobs, we can identify the $dp$ in (3.15) and (3.16) and we obtain an instability condition for the ambient medium,

$$\begin{aligned} \left(\frac{\partial \rho}{\partial s}\right)_p ds &< 0 \quad \text{for instability of a downward displacement} \\ \left(\frac{\partial \rho}{\partial s}\right)_p ds &> 0 \quad \text{-"- upward -"-} \end{aligned} \tag{3.18}$$

To simplify this we have to apply the Maxwell relation for thermodynamic potentials, in particular that for enthalpy $H(S,P) = U + PV$, $dH = T\,dS + V\,dP$, hence

$$\left(\frac{\partial H}{\partial s}\right)_p = T \;, \left(\frac{\partial H}{\partial p}\right)_s = V = \rho^{-1}$$

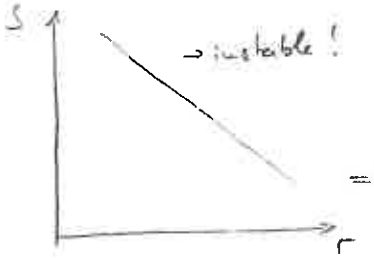
Permutability of the second partial derivatives implies

$$\frac{\partial^2 H}{\partial s\,\partial p} = \left(\frac{\partial T}{\partial p}\right)_s = \left(\frac{\partial \rho^{-1}}{\partial s}\right)_p = -\frac{1}{\rho^2}\left(\frac{\partial \rho}{\partial s}\right)_p \tag{3.19}$$

- all thermodynamic stable substances have temperatures that increase upon adiabatic compression,

$$\left(\frac{\partial T}{\partial p}\right)_S > 0 \quad \Longleftrightarrow \quad \left(\frac{\partial S}{\partial S}\right)_p < 0 \tag{3.20}$$

Hence we obtain the criterion for instability (3.18)




$$ds > 0 \quad \text{in the direction of gravity}$$

$$ds < 0 \quad \text{in the direction opposite of gravity} \tag{3.21}$$

$$\Rightarrow \underline{\frac{ds}{dr} < 0 \quad \text{for convective instability}} \quad (\text{centre of gravity is } r = 0)$$

- to remember this result, we associate specific entropy with buoyancy; if the entropy of the cluster atmosphere increases inwards, then more buoyant material underlies less buoyant material and the medium has the tendency to overturn; the overturning mixes entropy — high entropy (i.e. hotter than average) material rising and low entropy material (colder than average) sinking, so there is a net convective transport of heat outward with an approximately constant entropy distribution as a result; if we assemble the intra-cluster gas with pockets of different entropy (which is expected if we have frequent mergers of objects of different mass and characteristic entropy, i.e. galaxies and groups) then the gas will convectively rearrange itself such that it obeys an increasing entropy profile outwards:


→ stable gas atmosphere as a result of assembly gas with different entropies

- note that the derivation assumes no magnetic fields, i.e. that heat conduction is strictly isotropic and that temperatures can be compared along every direction; the presence of magnetic fields modifies this instability criterium and states that the temperature gradient is the source of free energy (rather than the entropy gradient); it is not clear whether this is of practical relevance for clusters!

## 3.1.3 Cluster Scaling Relations

- in order to use clusters as cosmological probes, we need to relate the different observables to a functional that is sensitive to cosmology; traditionally this is obtained by using the mass function:

observables at different frequencies → (scaling relations) → mass → mass function → (redshift measurement) → abundances vs. $z$ → constraints on cosmological parameters

- assumptions: choose average density of cluster so that this implicitely defines a cluster "radius" by

$$M_\Delta = \frac{4}{3}\pi R_\Delta^3 \cdot \Delta \rho_{cr}(z), \quad \Delta = 100 \ldots 500$$

which also relates the temperature to this definition, $T \sim T_\Delta$

- Cautionary remarks: when considering X-ray emission, we encounter $\rho_X$, $T_X$ which is not necessarily identical to $\bar{\rho} = \Delta \cdot \rho_{cr}$ and $T_\Delta$ as it is degenerate with observational biases; not accounting for these would break self-similarity as e.g., the presence of a clumped multiphase medium may bias $T_X$ towards the dense, cooler phase with a higher X-ray emissivity; on the other hand, we encounter similar problems when defining a 3D "radius" from a projected, non-isotropic density distribution (or X-ray emissivity $j_X \propto \rho^2$) => careful mocks are necessary in either case

### 3.1.3.1 Cosmologist's Ideal Cluster

- assume hydrostatic equilibrium (and mass-dependent kinetic pressure distribution):

$$\frac{kT}{\mu m} \sim v^2 \sim \frac{G M_\Delta}{R_\Delta} \propto M_\Delta^{2/3} \rho_{cr}^{1/3}$$

$$\Rightarrow T_\Delta \propto M_\Delta^{2/3} \rho_{cr}^{1/3} \propto \left(M_\Delta E(z)\right)^{2/3} \Leftrightarrow M_\Delta \propto T_\Delta^{3/2} \rho_{cr}^{-1/2} \propto T_\Delta^{3/2} E(z)^{-1} \quad (3.22)$$

- assume no mass scaling of $f_{gas}(<R_\Delta) = \frac{M_{gas}}{M_{tot}}$, $f_{star}(<R_\Delta) = \frac{M_{star}}{M_{tot}}$,

where $M_{tot} = M_{tot}(<R_\Delta) = M_{DM} + M_{gas} + M_{star}$ is the gravitational mass:

$$M_{gas} = \int_{}^{R_\Delta} dV \, \rho_{gas} \sim M_\Delta \, f_{gas} \sim M_\Delta \qquad (3.23)$$

$$M_{star} \propto M_\Delta \, f_{star} \sim M_\Delta \quad \Rightarrow \quad N_{gals} \propto M_\Delta$$

especially $N_{gals} \propto M_\Delta$ assumes a fair sampling of the luminosity function which is not anymore the case on group scales with $\mathcal{O}(10)$ galaxies

- Sunyaev-Zel'dovich scaling relation

$$Y = \int_{cluster} dA \, y = \frac{\sigma_T}{m_e c^2} \int_{cluster} dV \, n_e \, kT \propto M_{gas} \, T_\Delta \sim M_\Delta^{5/3} \, E(z)^{2/3} \qquad (3.24)$$

- X-ray scaling relation; assuming free-free emission (2-body process):

$$L_X \propto \int_{cluster} dV \, n_e \, n_{ion} \sqrt{kT} \sim M_{gas} \, \rho_{cr} \, T^{1/2} \sim M_\Delta^{4/3} \, E(z)^{1/3}$$

$$\sim T_\Delta^2 \, E(z)^{-1} \qquad (3.25)$$

### 3.1.3.2 Real Clusters

- observational scaling relations do not follow the self-similar predictions above; one finds (for $M$ as some observational proxy for $M_\Delta$):

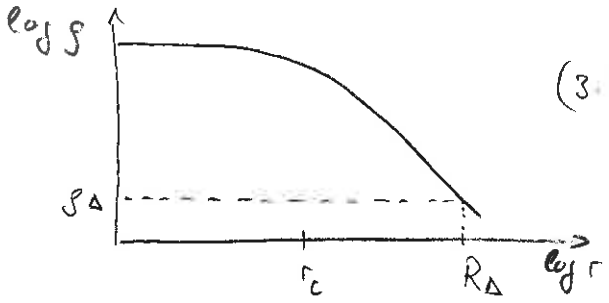
$$L_X \propto T_X^3 \quad , \quad \frac{d}{dM}\left(\frac{M_{gas}}{M}\right) > 0 \quad , \quad \frac{d}{dM}\left(\frac{M_{stars}}{M}\right) < 0, \qquad (3.26)$$

$Y(M)$ and $T_X(M)$ roughly in agreement with self-similar prediction

→ gas physics seems to modify these simple scale-invariant models and to impose new scales to the otherwise scale-free gravity!

- consider a simple toy model for the gas distribution of the ICM:

$$\rho(r) = \begin{cases} \rho_\Delta \left(\frac{r}{R_\Delta}\right)^{-2} & \text{for } r > r_c \\ \text{const.} & \text{for } r < r_c \end{cases} \qquad (3.27)$$




we define the cluster concentration parameter $c = \frac{R_\Delta}{r_c}$, implying

$$\rho(r_c) \equiv \rho_c = c^2 \rho_\Delta$$

- for bremsstrahlung (free-free) emission, the emissivity per unit volume is proportional to $\rho^2 T^{1/2}$; hence we obtain

$$L_X \propto \int_0^R r^2 dr\, \rho^2 T^{1/2} \quad \Rightarrow \quad \frac{dL_X}{d\log r} \propto r^3 \rho^2 T^{1/2} \sim \begin{cases} r^3, & r < r_c \\ r^{-1} T^{1/2}, & r > r_c \end{cases} \qquad (3.28)$$

i.e. the contribution to the X-ray luminosity per logarithmic bin in radius increases steeply towards $r_c$ and then drops beyond $r_c$ (after realizing that $T \sim r^{-1/2}$ in the peripheral cluster parts); the radii around $r_c$ dominate $L_X$ and thus, we expect

$$L_X \propto \rho_c^2 T^{1/2} r_c^3 \sim \rho_\Delta^2 c^4 T^{1/2} c^{-3} R_\Delta^3 \sim \rho_\Delta^2 c R^3 T^{1/2} \qquad (3.29)$$

using (3.22) and (3.23), we obtain

$$L_X \propto c\, M_\Delta^{4/3} E(z)^{1/3} \sim c\, T_\Delta^2 E(z)^{-1} \qquad (3.30)$$

- if $c$ is independent of mass, we recover (3.25), however gas physics modifies $c$ so that it assumes a mass-dependence; there have been 3 classes of models suggested to explain the deviation from scale invariance:

(1) "pre-heating" the gas by supernova winds or some other feedback mechanism <u>before</u> falling into clusters imprints an "entropy floor" onto the gas – a minimum level $K_{min}$ below which it cannot fall; the clusters' central entropy is $K_0 \sim T \rho_0^{-2/3} \sim T c^{-4/3}$; if all clusters have $K_0 = K_{min} = const.$, then

$$c \sim T^{3/4} K_{min}^{-3/4} \Rightarrow L_x \propto T^{2.75} \qquad (3.31)$$

$\Rightarrow$ thus, an entropy floor leads to larger cores (relative to $R_\Delta$), $r_c = \frac{R_\Delta}{c} \sim \left(\frac{K_{min}}{T}\right)^{3/4} R_\Delta$, which is more pronounced for smaller clusters (lower $T_\Delta$), and thus to a steeper $L_x$–$T$ relation closer to the observations

(2) an alternative possibility is that the gas gets heated after falling into the clusters, potentially through feedback by active galactic nuclei (AGN); this is however energetically much more expensive: to reach the same entropy, one needs more energy by a factor

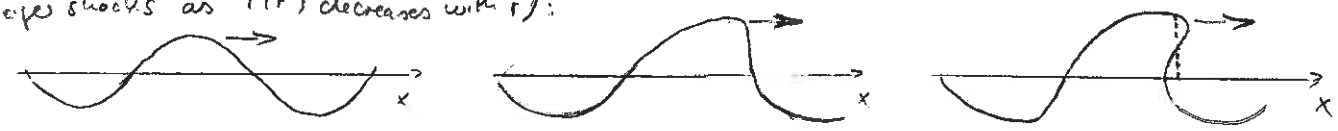
$$\frac{kT(\text{center})}{kT(\text{pre-heated})} \sim \frac{K(\text{center})}{K(\text{pre-heated})} \left(\frac{n(\text{center})}{n(\text{pre-heated})}\right)^{2/3} \sim 10^2$$

since $n_e(\text{center}) \sim 2 \cdot 10^{-3}\,\text{cm}^{-3}$ and $n_e(\text{pre-heated}) \sim 10\,\bar{n} \sim 2 \cdot 10^{-6}\,\text{cm}^{-3}$, but AGNs could provide the energy (Sect. 3.2) if the energy can be efficiently coupled into the ICM

(3) cooling out the low-entropy gas at the clusters center and fueling central star formation selectively removes the low-entropy gas; the gas at larger radii (and on higher adiabates) flows in adiabatically and replaces the condensed gas $\rightarrow$ formation of an entropy floor;
this process is observed to happen, but the star formation rate is only $\sim 1/10$ of what would be needed to explain the steeper $L_x$–$T$ slope

## 3.1.4 Shocks

- general considerations: imagine the propagation of a sound wave with finite amplitude; the sound speed is higher at higher temperature as $c_s \propto \sqrt{kT}$, so that the crest of the wave gradually overtakes the trough ($T \propto \rho^{\gamma-1}$); when faster moving gas overtakes slower moving gas, we get a discontinuous change of density and velocity, a shock; sketch of a such a behavior (important for outwards propagating waves shocks as $T(r)$ decreases with $r$):




- shocks can also be produced by any supersonic compressive disturbance, such as accretion of the intergalactic gas onto a galaxy cluster; in general, a shock wave is
  1) a pressure driven compressive disturbance;
  2) propagating faster than the "signal speed" for compressible waves, $c_s$;
  3) producing an irreversible change of the fluid state, i.e. increase in entropy;

  in most cases, a shock involves a "discontinuous" change of fluid properties over a scale proportional to the effective mean-free path $\lambda_{eff}$
- in "collisional" shocks, $\lambda$ is set by Coulomb-force mediated particle-particle collisions; in a plasma (which is of relevance for clusters) electromagnetic viscosities mediate interactions between charged particles and thus reduce $\lambda_{eff}$ by many orders of magnitude, $\lambda_{eff} \ll \lambda_{Coulomb}$, so that we are dealing here with "collisionless" shocks

## 3.1.4.1 Basic Equations

- we will start with the conservation laws for mass, momentum, and total energy; those are easily derived from the Boltzmann equation,

$$\frac{d}{dt} f(\vec{x},\vec{v},t) = \frac{\partial f}{\partial t} + \sum_i \dot{x}_i \frac{\partial f}{\partial x_i} + \sum_i \dot{v}_i \frac{\partial f}{\partial v_i} = \left.\frac{df}{dt}\right|_c \quad (\leftarrow \text{collision term}) \qquad (3.32)$$

which is the evolution equation for the phase space distribution function $f(\vec{x},\vec{v},t)$; conservation laws are obtained by taking the appropriate moments of the Boltzmann equation and integrating over velocity space, $d^3v$; we identify the mass density

$$\rho = \rho(\vec{x},t) = \int m\, f(\vec{x},\vec{v},t)\, d^3v \qquad (3.33)$$

- <u>mass conservation</u> (continuity equation): multiply (3.32) by $m$ and integrate over $d^3v$:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla}\cdot(\rho\vec{u}) = 0 \quad , \quad \vec{u} \equiv \langle \vec{v} \rangle \text{ mean fluid velocity} \qquad (3.34)$$

- <u>momentum conservation</u>: multiply (3.32) by $m\vec{v}$ and integrate over $d^3v$:

$$\frac{\partial \vec{u}}{\partial t} + (\vec{u}\cdot\vec{\nabla})\vec{u} = \vec{g} - \frac{1}{\rho}\vec{\nabla}P + \frac{1}{\rho}\vec{\nabla}\cdot\hat{\pi} \qquad (3.35)$$

with the viscosity tensor of a "Newtonian fluid" (implying that $\hat{\pi} \propto \frac{\partial u_i}{\partial x_j}$):

$$\pi_{ij} = \mu D_{ij} + \beta \delta_{ij} (\vec{\nabla}\cdot\vec{u}) \qquad (3.36)$$

$$D_{ij} \equiv \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3}\delta_{ij}(\vec{\nabla}\vec{u}) \quad \text{"deformation tensor"} \qquad (3.37)$$

$$\left\{\begin{matrix}\mu \\ \beta\end{matrix}\right\} \equiv \text{coefficients of } \left\{\begin{matrix}\text{shear} \\ \text{bulk}\end{matrix}\right\} \text{ viscosity} \quad [\,g\,cm^{-1}\,s^{-1}\,]$$

- <u>energy conservation</u>: multiply (3.32) by $mv^2$ and integrate over $d^3v$:

$$\rho\left(\frac{\partial \varepsilon}{\partial t} + \vec{u}\cdot\vec{\nabla}\varepsilon\right) \overset{(3.34)}{=} \frac{\partial}{\partial t}(\rho\varepsilon) + \vec{\nabla}\cdot(\rho\varepsilon\vec{u}) = -P\vec{\nabla}\cdot\vec{u} - \vec{\nabla}\cdot\vec{F} + \psi \qquad (3.38)$$

with $\varepsilon \equiv$ specific internal energy $= \frac{1}{2}\langle|\vec{v}-\vec{u}|^2\rangle = \frac{1}{2}\langle|\vec{w}|^2\rangle$ $[\,erg\,g^{-1}\,]$

$\vec{F} \equiv$ conduction heat flux $= \frac{1}{2}\rho\langle\vec{w}|\vec{w}|^2\rangle$ $[\,erg\,cm^{-3}\,s^{-1}\,cm\,]$

$\psi \equiv$ viscous dissipation rate $= \sum_{i,j} \pi_{ij}\frac{\partial u_i}{\partial x_j}$ $[\,erg\,cm^{-3}\,s^{-1}\,]$

- since we are interested how the total energy changes in a given volume, $\frac{\rho}{2}u^2 + \rho\varepsilon$, we are supplementing (3.88) with a conservation laws of $\frac{\rho}{2}u^2$; to this end we consider: $\frac{\partial}{\partial t}\left(\frac{\rho u^2}{2}\right) = \frac{u^2}{2}\frac{\partial \rho}{\partial t} + \rho\vec{u}\cdot\frac{\partial \vec{u}}{\partial t}$ and substitute (3.84) and (3.85) to get

$$\frac{\partial}{\partial t}\left(\frac{\rho u^2}{2}\right) = -\frac{u^2}{2}\vec{\nabla}(\rho\vec{u}) - \rho\vec{u}(\vec{u}\cdot\vec{\nabla})\vec{u} - \vec{u}\cdot\vec{\nabla}p - \rho\vec{u}\cdot\vec{g} + \vec{u}\cdot\vec{\nabla}\vec{\Pi} - \rho\frac{\vec{u}}{2}\vec{\nabla}\cdot u^2$$

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho u^2\right) + \vec{\nabla}\left(\frac{1}{2}\rho u^2\vec{v}\right) = -\vec{u}\cdot\vec{\nabla}p - \rho\vec{u}\vec{g} + \vec{u}\,\vec{\nabla}\hat{\Pi} \qquad (3.89)$$

"conservation of kinetic energy"

### 3.1.4.2 Jump conditions

- consider a propagating fluid discontinuity in its restframe; we denote
  - upstream conditions by $\rho_1, u_1, T_1$
  - downstream conditions by $\rho_2, u_2, T_2$
- we would like to derive the relations (a.k.a. "jump conditions") between $\rho_1, u_1, T_1$ and $\rho_2, u_2, T_2$ for a steady state, plane-parallel geometry of a fluid discontinuity such as a shock ($\vec{u}$ perpendicular to the discontinuity); as we will see, there are two types of discontinuities:
  1) shocks that are characterized by a mass flux through their interface
  2) contact discontinuities which have no mass flux through their interface
- within the shock front (a.k.a. "transition layer") on a scale of $\lambda_{eff}$, viscous effects are important and cause the shock in the first place, i.e. dissipate energy into heat and generate entropy; however, outside this layer, viscous effects are small on scales $L \gg \lambda_{eff}$; we will derive conservation equations of the form

$$\frac{d}{dx}Q(\rho, u, P) = 0 \Rightarrow Q(\rho, u, P) = \text{const.}$$

and although Q involves viscous terms, we can ignore these outside the shock zone and can derive jump conditions from equations without viscosity terms;

- we assume steady state, $\frac{\partial}{\partial t} = 0$, plane-parallel, $\frac{\partial}{\partial y} = \frac{\partial}{\partial z} = 0$, $\frac{\partial}{\partial x} = \frac{d}{dx}$, and ignore gravity; the equations (3.34) – (3.39) become

$$\frac{d}{dx}(\rho u) = 0 \tag{3.40}$$

$$u \frac{du}{dx} = -\frac{1}{\rho}\frac{dP}{dx} + \frac{1}{\rho}\frac{d}{dx}\left(\frac{4}{3}\mu \frac{du}{dx}\right) \tag{3.41}$$

$$\frac{d}{dx}(\rho \varepsilon u) = -P \frac{du}{dx} \tag{3.42}$$

$$\frac{d}{dx}\left(\frac{1}{2}\rho u^2 u\right) = -u \frac{dP}{dx} + \frac{u}{\rho}\frac{d}{dx}\left(\frac{4}{3}\mu \frac{du}{dx}\right) \tag{3.43}$$

- equation (3.40) gives

$$\rho u = \text{const.} \Rightarrow \rho_1 u_1 = \rho_2 u_2 \Rightarrow [\rho u] = 0 \tag{3.44}$$

where brackets indicate differences between up- and downstream quantities

- using

$$\frac{d}{dx}(\rho u^2) = \rho u \frac{du}{dx} + u \frac{d}{dx}(\rho u) \overset{(3.40)}{=} \rho u \frac{du}{dx}$$

allows (3.41) to be rewritten as

$$\rho u \frac{du}{dx} + \frac{dP}{dx} - \frac{d}{dx}\left(\frac{4}{3}\mu \frac{du}{dx}\right) = \frac{d}{dx}\left(\rho u^2 + P - \frac{4}{3}\mu \frac{du}{dx}\right) = 0$$

$$\Rightarrow \left[\rho u^2 + P - \frac{4}{3}\mu \frac{du}{dx}\right] = 0 \tag{3.45}$$

This demonstrates that within the transition zone, where $\mu$ and $\frac{du}{dx}$ are non-zero, $\rho u^2 + P$ is not constant; however, in the pre- and post-shock zones, $\mu$ and $\frac{du}{dx}$ are negligible, implying

$$\left[\rho u^2 + P\right] = 0 \tag{3.46}$$

In principle, we could use (3.45) to follow the behavior in the transition zone, i.e., to understand how entropy is generated, but on scales $L < \lambda$ the fluid description breaks down and we have to reside to the kinetic theory (or use plasma particle-in-cell codes to understand the non-linear behavior); from now on neglect viscosity effects in the bulk

- adding equations (3.42) and (3.43) gives

$$0 = \frac{d}{dx}\left( u\left(\tfrac{1}{2}\rho u^2 + \rho\varepsilon\right) + Pu \right) = \frac{d}{dx}\left( \rho u \left( \tfrac{1}{2}u^2 + \varepsilon + \frac{P}{\rho} \right)\right)$$

$$= \left( \tfrac{1}{2}u^2 + \varepsilon + \frac{P}{\rho} \right) \frac{d}{dx}(\rho u) + \rho u \frac{d}{dx}\left( \tfrac{1}{2}u^2 + \varepsilon + \frac{P}{\rho} \right)$$

since $\frac{d}{dx}(\rho u) = 0$ and $\rho u \neq 0$ we obtain

$$\frac{d}{dx}\left( \tfrac{1}{2}u^2 + \varepsilon + \frac{P}{\rho} \right) = 0 \quad \Rightarrow \quad \left[ \tfrac{1}{2}u^2 + \varepsilon + \frac{P}{\rho} \right] = 0 \qquad (3.47)$$

note, that we neglect viscosity and heat conduction that are only important in the transition zone

- summarizing, we have the Rankine-Hugoniot jump conditions for a plane-parallel shock:

$$[\rho u] = 0 \qquad (3.48)$$

$$[\rho u^2 + P] = 0 \qquad (3.49)$$

$$\left[ \tfrac{1}{2}u^2 + \varepsilon + \frac{P}{\rho} \right] = 0 \qquad (3.50)$$

Independent of the complicated physics within the transition layer, these conditions simply follow from the conservation laws; the first follows from mass conservation, the second from mass and momentum conservation and the third from mass and energy conservation; using $\varepsilon_i = \frac{1}{\gamma_i - 1}\frac{P_i}{\rho_i}$, we can rewrite (3.50) to get

$$\frac{1}{2}u_1^2 + \frac{\gamma_1}{\gamma_1 - 1}\frac{P_1}{\rho_1} = \frac{1}{2}u_2^2 + \frac{\gamma_2}{\gamma_2 - 1}\frac{P_2}{\rho_2} \qquad (3.51)$$

for a single-species gas described by a polytropic EOS (in principle, $\gamma_1 \neq \gamma_2$, e.g. if a shock dissociates molecules, or raises $T$ so that previously inaccessible degrees of freedom became accessible).

### 3.1.4.3 Contact Discontinuities

- $[\rho u] = 0$ allows for two types of solutions; the first type is clearly $\rho_1 u_1 = \rho_2 u_2 = 0$ and since $\rho_1$ and $\rho_2$ are non-zero, we have

$$u_1 = u_2 = 0 \tag{3.52}$$

$$p_1 = p_2 \quad \Rightarrow \quad [p] = 0$$

for (3.49); i.e. there is no mass flux through a contact discontinuity (CD)

- at a CD, there can be an arbitrary jump of the density, that however needs to be compensated by the same jump of $T$, but in the opposite direction!

### 3.1.4.4 Shock Mach Number

- the other type of solution requires $\rho_1 u_1 \neq 0$ so that we have a mass flux through this type of discontinuity that we call "shock"
- we define a dimensionless number that characterizes the shock strength, the Mach number as the ratio of shock speed to upstream sound speed $c_1^2 = \gamma \frac{p_1}{\rho_1}$,

$$M_1 \equiv \frac{u_1}{c_1} = \sqrt{\frac{\rho_1 u_1^2}{\gamma p_1}} = \sqrt{\frac{m u_1^2}{\gamma k \bar{T}_1}} \tag{3.53}$$

which can be interpreted as ratio of ram pressure $\rho_1 u_1^2$ to thermal pressure in pre-shock gas (or kinetic-to-thermal energy density)

- we can rewrite the Rankine-Hugoniot conditions in terms of $M_1$

$$\frac{\rho_2}{\rho_1} = \frac{u_1}{u_2} = \frac{(\gamma+1) M_1^2}{(\gamma-1) M_1^2 + 2} \tag{3.54}$$

$$\frac{p_2}{p_1} = \frac{\rho_2 k \bar{T}_2}{\rho_1 k \bar{T}_1} = \frac{2\gamma M_1^2 - (\gamma-1)}{\gamma+1} \tag{3.55}$$

implying

$$\frac{T_2}{T_1} = \frac{[(\gamma-1) M_1^2 + 2][2\gamma M_1^2 - (\gamma-1)]}{(\gamma+1)^2 M_1^2} \tag{3.56}$$

- those simplify for strong shocks ($M_1 \gg 1$), yielding

$$\frac{\rho_2}{\rho_1} = \frac{u_1}{u_2} \approx \frac{\gamma+1}{\gamma-1} = 4 \qquad (3.57)$$

$$P_2 \approx \frac{2\gamma}{\gamma+1} M_1^2 P_1 = \frac{2}{\gamma+1} \rho_1 u_1^2 = \frac{3}{4} \rho_1 u_1^2 \qquad (3.58)$$

$$k\bar{T}_2 \approx \frac{2\gamma(\gamma-1)}{(\gamma+1)^2} k T_1 M_1^2 = \frac{2(\gamma-1)}{(\gamma+1)^2} m u_1^2 = \frac{3}{16} m u_1^2 \qquad (3.59)$$

where we used $\gamma = \frac{5}{3}$ in the last equalities

- in the shock-rest frame, the post-shock kinetic and thermal energies are ($\gamma = \frac{5}{3}$)

$$\frac{1}{2} u_2^2 \approx \frac{1}{32} u_1^2$$

$$\frac{3}{2} \frac{k\bar{T}_2}{m} \approx \frac{9}{32} u_1^2 = \frac{9}{16} \cdot \frac{1}{2} u_1^2$$

so roughly half of the pre-shock kinetic energy is converted to thermal energy (in the shock rest frame); the total energy of the post-shock gas,

$$K.E._2 + U_2 = \frac{1}{2} u_2^2 + \frac{3}{2} \frac{k\bar{T}_2}{m} \approx \frac{10}{16} \cdot \frac{1}{2} u_1^2 = \frac{5}{8} K.E._1$$

is lower (in the shock rest frame) because of the work done on the post-shock gas by the pressure and viscosity in the shock; i.e. the $pdV$ work done on the post-shock gas in expanding its volume; note that this $pdV$ term is absent in the rest frame of the post-shock gas (see Sect. 3.1.5)

- the post-shock Mach number is

$$M_2 \equiv \frac{u_2}{c_2} = \frac{u_1}{c_1} \frac{u_2}{u_1} \frac{c_1}{c_2} = M_1 \frac{u_2}{u_1} \left(\frac{\bar{T}_1}{\bar{T}_2}\right)^{1/2} \qquad (3.60)$$

in the strong-shock limit

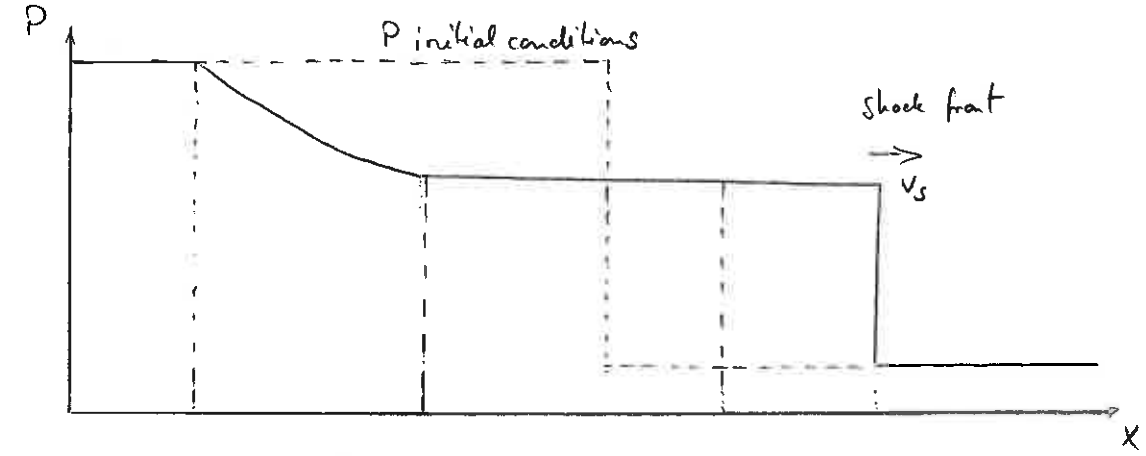
$$M_2 \approx M_1 \frac{(\gamma-1)}{(\gamma+1)} \left[\frac{(\gamma+1)^2}{2\gamma(\gamma-1)M_1^2}\right]^{1/2} = \left(\frac{\gamma-1}{2\gamma}\right)^{1/2} \approx 0.45 \qquad (3.61)$$

$\Rightarrow$ a shock converts supersonic gas into denser, slower moving, higher pressure, subsonic gas.

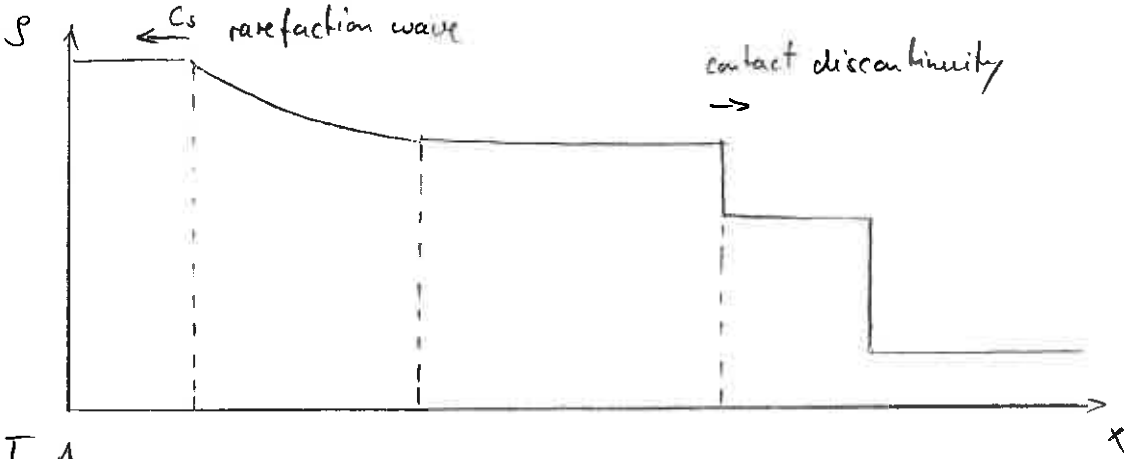
- the shock increases the specific entropy of the gas by an amount

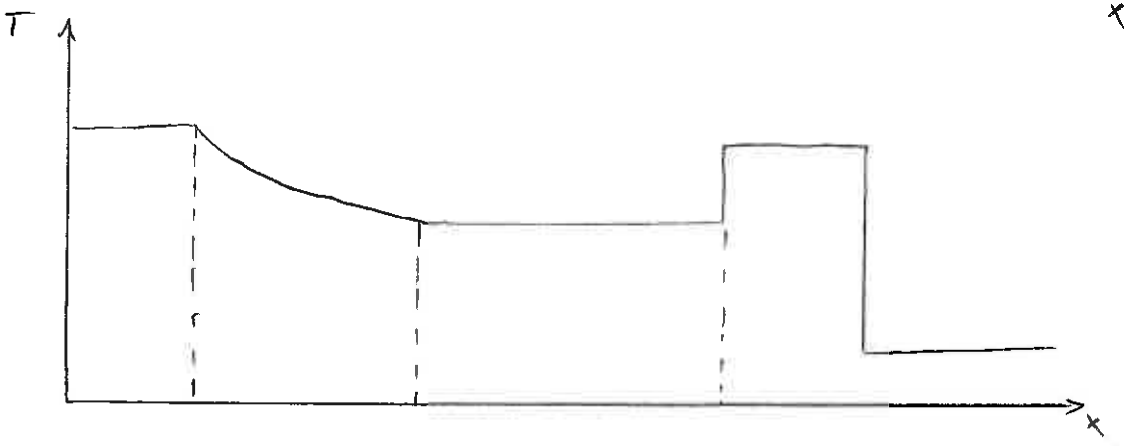
$$s_2 - s_1 = c_v \ln\left(\frac{p_2}{\rho_2^\gamma}\right) - c_v \ln\left(\frac{p_1}{\rho_1^\gamma}\right) = c_v \ln\left(\frac{p_2}{p_1}\right) - c_v \gamma \ln\left(\frac{\rho_2}{\rho_1}\right) = c_v \ln\left(\frac{u_2}{u_1}\right)$$

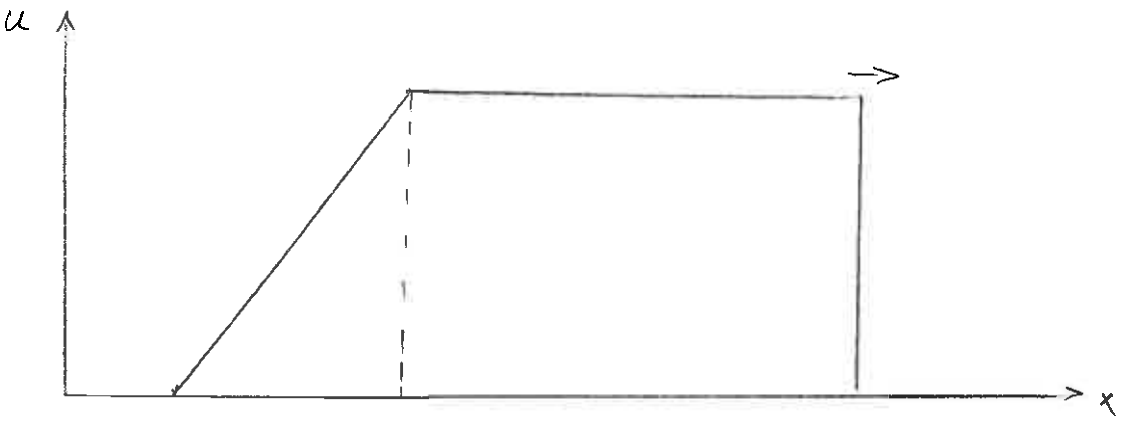
hence, the shock shifts gas to a higher adiabat (gas can move adiabatically along an adiabat, while changes in entropy move it from one adiabat to another)




solution to the "Riemann problem"[4]



## 3.1.5 Entropy Generation by Accretion

- philosophical detour: the Uncertainty Principle is $\Delta p_x \Delta x = h$, and statistical mechanics counts the number of states with $h^{-3} d^3x d^3p$; hence the phase space density of cluster gas is

$$f \sim \frac{h^3 d^6N}{d^3x\, d^3p} \sim n (m_p v)^{-3} h^3 \sim \frac{10^{-3}}{cm^3} \left(\frac{3}{2} \cdot 10^{-24} \text{ g } 10^8 \frac{cm}{s}\right)^{-3} \cdot \left(6 \cdot 10^{-27} \text{ ergs}\right)^3$$

$$\sim 6 \cdot 10^{-35} \propto K^{-3/2} \quad \left[\text{since } K \sim \frac{kT}{n^{2/3}} \sim \frac{v^2}{n^{2/3}}\right] \qquad (3.63)$$

if this was unity we would deal with degenerate gas; instead, this is extremely small, making it the least degenerate gas in the universe or equivalently the highest entropy gas (of equilibrium systems)!

### 3.1.5.1 Smooth Accretion

- one way to approach the problem of gravitationally driven entropy generation is through spherically symmetric models of smooth accretion, in which gas passes through an accretion shock as it enters the cluster; if the incoming gas is cold, then the cluster accretion shock is the sole source of cluster entropy; if, instead, the incoming gas has been heated before passing through the accretion shock, the Mach number is smaller and the cluster entropy level reflects both the amount of preheating and entropy production at the shock
- we first consider the case of cold accretion (P and K of incoming gas are negligible), which implies the strong shock regime; conveniently, we transform our Rankine-Hugoniot jump conditions to the rest frame of the post-shock gas, i.e. the cluster rest frame


$$\Rightarrow u_1 = \frac{4}{3} v_{ac}$$

$$kT_2 = \frac{3}{16} m u_1^2 = \frac{1}{3} m v_{ac}^2$$

- note that the location of the accretion shock does not move outwards with $\frac{1}{4} v_{ac}$, as the gravitational attraction of the cluster potential causes it to fall onto the cluster; depending on that infall rate, the accretion shock can either be stationary or move (slowly) outwards
- suppose that mass accretes in a series of concentric shells, each with baryon fraction $f_b$, that initially comove with the Hubble flow as in the spherical collapse model of Sect. 2.3; in this simple model, a shell that initially encloses total mass $M$ reaches zero velocity at the turnaround radius $r_{ta}$ and falls back through an accretion shock at radius $r_{ac}$ in the neighborhood of the virial radius $\frac{r_{ta}}{2}$; the system of governing equations are

$$\dot{M}_g = 4\pi r_{ac}^2 \rho_1 v_{ac} = f_g \dot{M} \qquad (3.64)$$

$$v_{ac}^2 = \frac{2GM}{r_{ta}} \qquad (\text{assuming } \Omega_\Lambda = 0) \qquad (3.65)$$

$$kT_2 = \frac{1}{3} m v_{ac}^2 \quad , \quad m = \mu m_p \qquad (3.66)$$

$$\rho_2 = 4 \rho_1 \qquad (3.67)$$

here, $\rho_1$ denotes pre-shock density, $\rho_2$ and $T_2$ is post-shock density and temperature, $r_{ac} = \frac{r_{ta}}{2}$ is the accretion radius; in the post-shock frame the post-shock thermal energy equals the pre-shock ram pressure (+ initial thermal energy that we neglect here) and (3.66) implies

$$\frac{3}{2} \frac{kT_2}{m} = \frac{3}{2} \frac{1}{3} v_{ac}^2 = \frac{v_{ac}^2}{2}$$

- the post-shock entropy produced by smooth accretion is thus

$$P = K \rho_g^{5/3}$$

$$\rightsquigarrow K_2 \equiv \frac{k T_2}{m \rho_2^{2/3}} \overset{(3.66),(3.67)}{=} \frac{v_{ac}^2}{3 (4 \rho_1)^{2/3}} \qquad (3.68)$$

- combining (3.64) and (3.65) yields

$$\dot{M}_g = 4\pi r_{ac}^2 \rho_1 \left( \frac{GM}{r_{ac}} \right)^{1/2} \rightsquigarrow \rho_1 = \frac{\dot{M} f_g}{4\pi r_{ac}^{3/2} \sqrt{GM}}$$

- Plugging this into (3.68) leads to

$$K_2 = \frac{v_{ac}^2}{3 (4\rho_1)^{2/3}} = \frac{GM \quad (4\pi)^{2/3} \quad r_{ac}^{3/2 \cdot 2/3} \quad (GM)^{1/3}}{r_{ac} \; 3 \cdot (4 f_g \dot{M})^{2/3}} =$$

$$= \frac{1}{3} \left( \frac{4\pi (GM)^2}{4 f_g \dot{M}} \right)^{2/3} = \frac{1}{3} \left( \frac{\pi G^2}{f_g} \right)^{2/3} \left( \frac{M}{\frac{t}{t} \frac{d \ln M}{dt}} \right)^{2/3}$$

$$K_{2,\mathrm{smooth}} = \frac{1}{3} \left( \frac{\pi G^2}{f_g} \right)^{2/3} \left( \frac{d \ln M}{d \ln t} \right)^{-2/3} (Mt)^{2/3} \propto t^{2/3} \qquad (3.69)$$

=> because the entropy generated at the shock front increases monotonically with time, such an ideal, smooth accreting cluster never connects, but rather accretes shells of baryons as if they were onion skins; it is useful to cast (3.69) into dimensionless form using a characteristic cluster entropy $K_{200}$:

$$K_{200} \equiv \frac{k T_{200}}{m (200 f_b \rho_{cr})^{2/3}} = \frac{1}{2} \left( \frac{2\pi}{15} \frac{G^2 M_{200}}{f_b H(z)} \right)^{2/3} \qquad (3.70)$$

with $\rho_{cr} = \frac{3H^2(z)}{8\pi G}$, $k T_{200} \overset{\mathrm{SIS}}{=} \frac{G M_{200} m}{2 r_{200}} = \frac{m}{2} \left[ 10 G M_{200} H(z) \right]^{2/3}$

and $M_{200} = \frac{4}{3} \pi r_{200} \cdot 200 \rho_{cr} \rightsquigarrow r_{200} = \left( \frac{GM_{200}}{100 H^2} \right)^{1/3}$

- note that we adopted here the characteristic temperature of a singular isothermal sphere (Sect. 2.3.6.1, Equ. (2.106))

- we define an effective radial coordinate corresponding to the amount of gas accreted at $t$ divided by that at $t_0$:

$$\eta \equiv \frac{M_g(t)}{f_b M_{200}(t_0)} \qquad (3.71)$$

and cast (3.69) into dimensionless form

$$\frac{K_{2,\mathrm{smooth}}}{K_{200}} = \frac{2}{3}\left(\frac{15}{2}\frac{H(z)}{M_{200}}\right)^{\frac{2}{3}}\left(\frac{d\ln M}{d\ln t}\right)^{-\frac{2}{3}}(Mt)^{\frac{2}{3}} = \frac{2}{3}\left(\frac{15}{2}H_0 t_0\right)^{\frac{2}{3}}\left(\frac{d\ln M}{d\ln t}\right)^{-\frac{2}{3}}\left(\frac{Mt}{M_{200}t_0}\right)^{\frac{2}{3}}$$

$$\rightsquigarrow \quad \frac{K_{2,\mathrm{smooth}}}{K_{200}} = \frac{2}{3}\left(\frac{15}{2}H_0 t_0\right)^{2/3}\left(\frac{d\ln\eta}{d\ln t}\right)^{-2/3}\left(\frac{\eta\, t(\eta)}{t_0}\right)^{2/3} \qquad (3.72)$$

- hence the entropy profile due to smooth accretion of cold gas depends entirely on the mass accretion history $M(t)$, and the entropy profiles of objects with similar $M(t)$ should be self-similar with respect to $K_{200}$
- extended Press-Schechter theory or numerical simulations show that for clusters in the mass range $10^{14} - 10^{15}\ M_\odot$ grow roughly as $M(t) \propto t$ to $M(t) \propto t^2$ in the concordance model [from Voit et al. ApJ, 593, 272 (2003)]
- plugging these growth times $t \sim M^{0.5\ldots1}$ into Eqn. (3.72) yields

$$K_{\mathrm{smooth}} \propto M_g^{1\ldots4/3} \qquad (3.73)$$

- throughout a cluster, $M_g$ encompassed by a given radius is approximately $M \sim r$ (which is exact for the SIS, Eqn. (2.106)), we have

$$K_{\mathrm{smooth}} \propto r^{1\ldots4/3}, \qquad (3.74)$$

which compares well with the numerical results

$$K_{\mathrm{sim}} \propto r^{1.1}$$

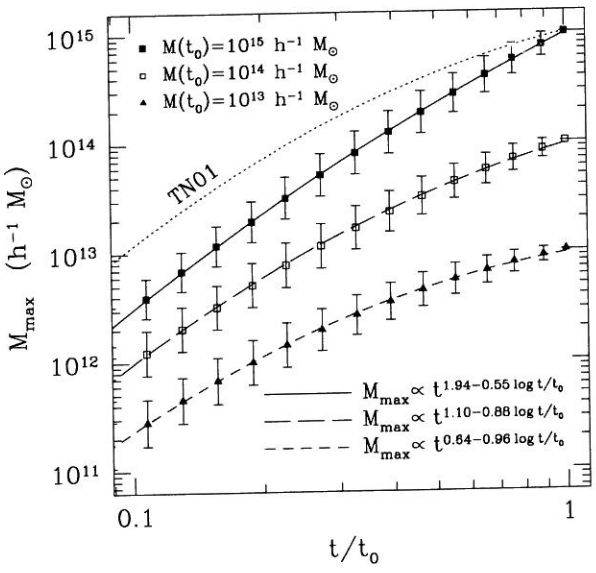



FIG. 1.—Accretion histories for halos of total mass $M(t_0)$ at the present time $t_0$. Filled squares give the logarithmic mean of the maximum progenitor mass $M_{\mathrm{max}}$ in the merger history of a halo of mass $M(t_0) = 10^{15}$ $h^{-1}$ $M_\odot$, computed from 1000 realizations of the Lacey & Cole (1993) algorithm. Open squares and filled triangles show the corresponding quantity for $M(t_0) = 10^{14}$ and $10^{13}$ $h^{-1}$ $M_\odot$. Error bars indicate one standard deviation. The solid line shows the best-fitting parabola in $\log M(t)$–$\log t$ space for $M(t_0) = 10^{15}$ $h^{-1}$ $M_\odot$. The long- and short-dashed lines show best fits for $M(t_0) = 10^{14}$ and $10^{13}$ $h^{-1}$ $M_\odot$, respectively. The dotted line indicates the accretion history used by Tozzi & Norman (2001) for their $10^{15}$ $h^{-1}$ $M_\odot$ cluster in a ΛCDM cosmology.

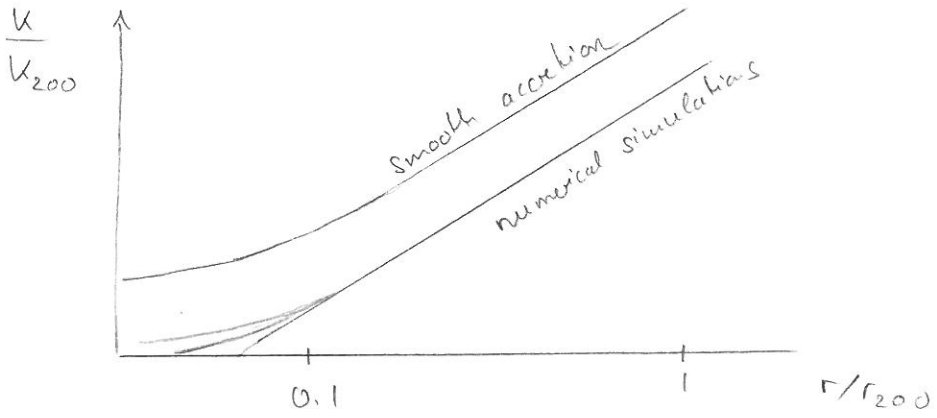
## 3.1.5.2 Hierarchical Merging

- in real clusters, the accreting gas is lumpy, not smooth, which transforms the nature of entropy generation; incandy gas that is bound to accreting sublumps of matter enters the cluster with a wide range of densities; there is no well-defined accretion shock, but rather a complex network of shocks as different lumps of infalling gas mix with the intracluster medium (ICM) of the main cluster
- numerical codes employing different numerical techniques (Eulerian grid codes or Lagrangian "particle" codes) all agree on the baseline profile in non-radiative gas simulations for $r > 0.1\ r_{200}$:

$$K_{sim} \simeq 1.32\ K_{200} \left(\frac{r}{r_{200}}\right)^{1.1} \qquad (3.75)$$

for $r < 0.1\ r_{200}$, there is more dispersion among the simulated clusters and the answer depends somewhat on the numerical technique, with grid codes showing an elevated entropy core due to efficient "entropy mixing"

- despite the complexity of shock structure in hierarchical accretion, the simulated entropy profiles resemble that from smooth accretion models, but with an important distinction: the normalization of the smooth models is higher

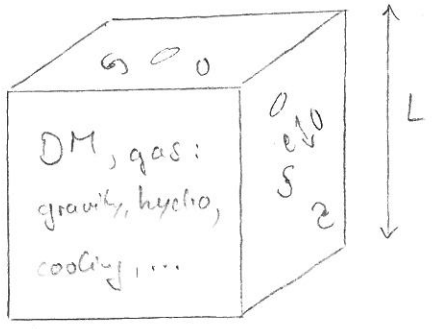



- the likely reason is that smooth accretion maximizes the entropy production as smoothing does not change the accretion velocity but reduces the mean density of accreting gas lumps; since the post-shock entropy scales as $K_2 \propto v_{ac}^2\ \rho_1^{-2/3}$, a smaller (smoothed) density implies lower entropy everywhere

- this may be of physical relevance : consider the case of pre-heating gas before it falls into clusters; heating causes the gas to adiabatically expand and smoothes the density field of the pre-shock gas ; this could then explain elevated entropy profiles of low-temperature clusters relative to the baseline profile

## 3.2 Radiative Physics

- observed cluster scaling relations do not obey self similar predictions; hence we have to take more realistic physics such as cooling and star formation into consideration ; the non-linearity of the problem requires numerical simulations that represent a formidable computational challenge, requiring codes that simulate 3D hydrodynamics spanning an enormous range in scales and track a plethora of physical processes :




periodic box of a representable volume of the universe

$L \gtrsim 300\ \mathrm{Mpc} \sim 10^{27}$ cm : for a few massive clusters

$l \sim 30\ \mathrm{kpc} \sim 10^{23}$ cm : diameter of galaxies :
$\Rightarrow$ grid refinement or Lagrangian code possible

$l_* \sim 3\ \mathrm{pc} \sim 10^{19}$ cm : for star forming region
$\Rightarrow$ sub-grid approximation necessary

$\Rightarrow$ overall $\left(\frac{L}{l_*}\right)^3 \sim 10^{8^3} \sim 10^{24}$ dynamic (3D) range!

- first, we will turn to the physics of cooling and condensation of baryons into stars (that will happen inevitably once the gas gets sufficiently dense) and transformed a fraction $\frac{f_*}{f_{gas}} \sim 20\%-40\%$ into stars without accounting for energy feedback ; since this is 4-5 times as much as observed in a cluster, we will then look at various "feedback processes" that were suggested to solve the "central problem" of galaxy formation!

## 3.2.1 Radiative Cooling

- at high temperatures ($kT \gtrsim 2$ keV) the light- and intermediate-mass elements of the ICM are fully ionized, so the only cooling process for them is free-free emission (thermal bremsstrahlung); below ~2 keV, recombination-line cooling of heavy elements (Fe, ...) starts to dominate the cooling (and associated X-ray emission, assuming typical heavy element abundances relative to hydrogen, which are $\sim 0.3$ times those found in the Sun)

- the physics of bremsstrahlung emission is simple: electrons scatter off ions and are deflected in the Coulomb field of the ions; they radiate because of their acceleration and thus loose energy – they "cool"

- the spectral X-ray emissivity $j_\nu$ is defined as the amount of energy emitted in photons of frequency $\nu$ per unit frequency interval $d\nu$, per unit time and per unit plasma volume, $j_\nu = \frac{d^3E}{d\nu\, dt\, dV}$; it must scale with the product of electron and ion number density (2-body process); with the time available for the scattering process, $t \sim \frac{l}{\Delta v} \sim \frac{l\, m}{\sqrt{kT}}$, where $\Delta v$ is the relative velocity of electron and ion, and the Boltzmann factor for the distribution of energy at a given temperature; hence we get

$$j_\nu = \frac{d^3E}{d\nu\, dt\, dV} = \tilde{C}\, \frac{n^2}{\sqrt{kT}}\, e^{-\frac{h\nu}{kT}}\,, \quad \tilde{C} = \text{const.} \qquad (3.76)$$

- the volume emissivity is the integral of $j_\nu$ over frequency $\nu$:

$$j = \frac{d^2E}{dV\,dt} = \int_0^\infty d\nu\, \frac{d^3E}{d\nu\, dt\, dV} = \frac{\tilde{C}}{h}\, n^2\, \frac{kT}{\sqrt{kT}} \underbrace{\int_0^\infty dx\, e^{-x}}_{=1} = C\, n^2 \sqrt{kT} \qquad (3.77)$$

$$\sim 2.5 \times 10^{-23} \left(\frac{n_H}{1\,cm^{-3}}\right)^2 \left(\frac{T}{10^8 K}\right)^{1/2} \frac{erg}{cm^3\, s}\,, \text{ for } 0.3 \text{ solar metallicity } Z_\odot$$

- comparing the thermal energy content to the total (frequency-integrated) X-ray emissivity defines the cooling time

$$t_{cool} = \frac{\varepsilon_{th}}{\dot{\varepsilon}_{Brems}} = \frac{3nkT}{2j} \propto \frac{\sqrt{kT}}{n}$$ , with $n = \frac{\rho}{\mu m_p}$ and $n_H = \frac{\rho}{m_p} X_H$

$\mu = 0.588$, $X_H = 0.76$ (sect 1.1.2)

$$t_{cool} \simeq 2\ \mathrm{Gyr} \left(\frac{kT}{\mathrm{keV}}\right)^{1/2} \cdot \left(\frac{n_e}{10^{-2}\,\mathrm{cm}^{-3}}\right)^{-1} \qquad (3.78)$$

→ hence in the centers of galaxy clusters, $t_{cool}$ is smaller than the Hubble time: if gas in pressure equilibrium cools, it becomes denser and cools even faster; this is a runaway process that should lead to large amounts of cold gas and star formation – in conflict with observations; this is the famous "cooling flow problem"

- it is enlightening to rewrite $t_{cool}$ in terms of the cluster entropy $K_e$

$$t_{cool} = t_0 \left(\frac{kT}{kT_0}\right)^{1/2} \frac{n_0}{n_e} \quad \text{with } t_0 = 2\,\mathrm{Gyr},\ kT_0 = \mathrm{keV},\ n_0 = 10^{-2}\,\mathrm{cm}^{-3}$$

and using $K_e = \frac{kT}{n_e^{2/3}} \rightsquigarrow K_e^{3/2} = \frac{(kT_e)^{3/2}}{n_e}$, we get

$$t_{cool} = t_0 \frac{(kT)^{3/2}}{n_e} \frac{kT_0}{kT} \frac{n_0}{(kT_0)^{3/2}} = t_0 \left(\frac{K_e}{K_0}\right)^{3/2} \frac{kT_0}{kT} \qquad (3.79)$$

with $K_0 \simeq 21.5\ \mathrm{keV\,cm^2}$ as typical values of the central entropy in cool-core clusters that need to be stabilized by some energy feedback; this demonstrates that clusters with similar temperatures (~ potential depths) have longer cooling times if the central entropy is large → critical entropy $K_c$:

$$K_c(T) \approx 80\ \mathrm{keV\,cm^2} \left(\frac{t_{cool}}{14\,\mathrm{Gyr}}\right)^{2/3} \left(\frac{kT}{\mathrm{keV}}\right)^{2/3}, \qquad (3.80)$$

that is large enough to avoid the cooling catastrophe in groups of $kT \sim$ keV

## 3.2.2 Energy Feedback

- we have seen that the cooling time in the core region of cool core clusters is smaller than the Hubble time, which would imply a cooling catastrophe if not countered by energy feedback; to see how much feedback is needed let's first compute the cooling rate and redefine the X-ray emissivity as an energy cooling rate $\Lambda(T)$ according to

$$j = C\, n_H^2 \sqrt{kT} \equiv \Lambda_0(T)\, n_H^2 \qquad (3.81)$$

with $\Lambda_0\left(T = 10^8 K\right) = 2.5 \times 10^{-23} \frac{erg\, cm^3}{s} \left(\frac{T}{10^8 K}\right)^{1/2}$, $\left(Z = 0.3\, Z_\odot\right)$

- we adopt a typical gas density profile as found in X-ray observation, the so-called beta-profile which is simply a King profile with the outer slope parametrized by $\beta \sim \frac{2}{3} \ldots 1$:

$$n(r) = n_0 \left[1 + \left(\frac{r}{r_c}\right)^2\right]^{-\frac{3}{2}\beta} \qquad (3.82)$$

- hence the X-ray luminosity ~ cooling luminosity is given by

$$L_X = \int dV\, j = \Lambda_0 \sqrt{\frac{kT}{kT_0}}\; 4\pi \int n^2(r)\, r^2\, dr \qquad \left[x = \frac{r}{r_c}\right]$$

$$= \frac{4\pi}{3} r_c^3\, n_0^2\, \Lambda_0 \sqrt{\frac{kT}{kT_0}}\; 3 \int_0^\infty \frac{x^2\, dx}{\left(1 + x^2\right)^{3\beta}}$$

$$= \frac{4\pi}{3} r_c^3\, n_0^2\, \Lambda_0 \sqrt{\frac{kT}{kT_0}} \begin{cases} 3 \cdot \frac{\pi}{16} \left(\sim 0.6\right) & \text{for } \beta = 1 \\ 3 \cdot \frac{\pi}{4\sqrt{1}} \left(\sim 2.4\right) & \text{for } \beta = \frac{2}{3} \end{cases}$$

$$\sim 10^{44} \frac{erg}{s} \left(\frac{r_c}{100\, kpc}\right)^3 \left(\frac{n_0}{10^{-2} cm^{-3}}\right)^2 \left(\frac{kT}{3\, keV}\right)^{1/2} \quad \text{and } \beta = 1 \qquad (3.83)$$

→ note that to order of magnitude, it suffices to assume a homogeneous sphere with radius $r_c$ and the central (core) density for calculating $L_X$; this corresponds to our finding $\frac{dL_X}{d\ln r} \propto r^3 n^2 \sqrt{T}$ (Sect. 3.1.3.2) that radii around $r_c$ dominate $L_X$

## 3.2.2.1 Feedback from Supernovae

- the first obvious candidate for energy feedback are supernovae, i.e. exploding stars at the end of their lifetimes; there are two types of supernovae (SNe) 1) core-collapse SNe (of type SN Ib,c; II) and 2) thermonuclear SNe (SN Ia)

1) core collapse SNe:

- as a massive star ($M \geq 10 M_\odot$) has used up its "fuel" (H, He, ...) the energy generated by nuclear burning cannot anymore balance the gravitational attraction; the core collapses and forms a black hole or neutron star (pulsar); the envelope also collapses to nuclear densities which triggers a shock that unbinds the envelope and ejects it => enrichment of the surroundings with intermediate-mass elements, so-called "$\alpha$-elements", which can be built from $\alpha$-particle nuclei, $^4He$, such as $^{16}O$, $^{20}Ne$, $^{24}Mg$, $^{28}Si$, $^{32}S$, $^{36}Ar$, $^{40}Ca$, $^{48}Ti$

- to estimate the effect of SNe heating on the ICM, we make two simplifying assumptions: 1) metals are fully mixed within the ICM and neglect radiative losses, and 2) assume solar abundances. Since the metallicity $Z$ of clusters is typically $0.3\, Z_\odot$ and radiative losses cause a large fraction of this SNe energy to be radiated away, these numbers represent the absolute upper limit that SNe can contribute to the heating:

$$\frac{M_\alpha}{M_{gas}} \sim 0.02 \qquad \text{(solar abundance)}$$

$$\frac{E_{SN}\, m_p}{M_\alpha} \sim \frac{10^{51}\,\text{erg}}{10\, M_\odot}\, m_p \sim \frac{10^{51-24-34}}{2}\ \frac{\text{erg}}{\text{nucleon}} \sim \frac{1}{2}\, 10^{-7}\ \frac{\text{TeV}}{\text{nucleon}} \sim 50\ \frac{\text{keV}}{\text{nucleon}}$$

$$\rightsquigarrow \frac{E_{SN}\, m_p}{M_{gas}} \sim 1\ \frac{\text{keV}}{\text{nucleon}} \qquad \text{(neglecting radiative losses)} \qquad (3.84)$$

2) Thermonuclear SNe:

- progenitor system consists of a binary system with at least one massive ($\sim 1 M_\odot$) C-O white dwarf; in the "single-degenerate" scenario, the companion star is a red giant that accretes onto the C-O white dwarf that triggers a thermonuclear explosion as it approaches the Chandrasekhar mass of $\sim 1.4 M_\odot$; alternatively 2 C-O white dwarfs can merge and also cause a thermonuclear runaway burning of C+O $\Rightarrow$ type Ia SN explosion that generates $\sim 1 M_\odot$ $^{56}Ni$ which decays radioactively into $^{56}Co$ and eventually $^{56}Fe$

- same assumptions as above yield

$$\frac{M_{Fe}}{M_{gas}} \sim 0.001 \qquad \text{(solar abundance)}$$

$$\frac{E_{SN}\, m_p}{M_{Fe}} \sim \frac{10^{51}\,\text{erg}\; m_p}{1 M_\odot} \sim 500 \frac{\text{keV}}{\text{nucleon}}$$

$$\approx \frac{E_{SN}\, m_p}{M_{gas}} \sim 0.5 \frac{\text{keV}}{\text{nucleon}} \qquad (3.85)$$

$\Rightarrow$ 2 problems: 1) energetics not sufficient and 2) rad. losses too strong to solve the "cooling flow problem"

1) for comparison, estimate thermal energy of a Milky Way-type galaxy and a massive cluster:

$$E_{gal} \sim \frac{m_p}{2} v_{gal}^2 \sim 0.25\ \text{keV} \left(\frac{v_{gal}}{220\ \text{km s}^{-1}}\right)^2$$

$$E_{cluster} \sim \frac{m_p}{2} \sigma_{cluster}^2 \sim 8\ \text{keV} \left(\frac{\sigma}{1200\ \text{km s}^{-1}}\right)^2 \qquad (3.86)$$

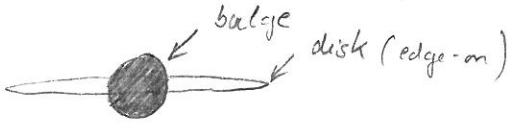
$\Rightarrow$ while SNe feedback can energetically modulate the star formation within galaxies, it is simply too weak in clusters

2) in order to avoid radiative losses, SNe heating has to raise the entropy of the gas it heats to at least $\sim 100\ \text{keV cm}^2$ (equ. 3.80); an evenly distributed thermal energy input of order 1 keV would thus have to go into gas significantly less dense than $10^{-3}\ \text{cm}^{-3}$ to avoid such losses;

but gas near the centers of present-day clusters (not to mention the galaxies, where SNe occur) is denser than that with average densities $\bar{n}_{ISM} \sim 1\,cm^{-3}$, particularly at earlier times when most of the star formation happened; simulations that spread SNe feedback evenly thus produce too many stars in clusters!

### 3.2.2.2 Feedback by active galactic nuclei (AGN)

- observationally, it is known that virtually every galaxy hosts a supermassive black hole (SMBH) at its center with masses $10^6 \leq \frac{M_{SMBH}}{M_\odot} \leq 10^{10}$; those SMBH masses are tightly correlated with the stellar mass in galactic bulges; bulges are defined as the central spheroidal stellar component in a disk galaxy ("late types") or the entire elliptical stellar distribution of ellipticals ("early types"), including the bright central galaxies (BCGs) in clusters:




$$M_{SMBH} \sim 0.005\ M_{bulge}$$
$$M_{*,BCG} \sim 10^{12}\ M_\odot \qquad (3.87)$$
$$M_{SMBH}(BCG) \sim 5 \times 10^9\ M_\odot$$

which compares well with the latest mass measurement of the SMBH in M87 of $6 \times 10^9\ M_\odot$ (M87 is the BCG in Virgo, the closest galaxy cluster to us with $D_{Virgo} \sim 17\ Mpc$)

- the accretion power onto the SMBH can be estimated by the release of gravitational energy with a radiative efficiency $\eta \sim 0.1$,

$$E_{AGN} \sim \eta\, M_{SMBH}\, c^2 \sim 0.1 \cdot 5\,10^9 \cdot 2 \cdot 10^{33}\ 10^{21}\ erg \sim 10^{63}\ erg\ (\sim 10^{12}\ SNe)$$

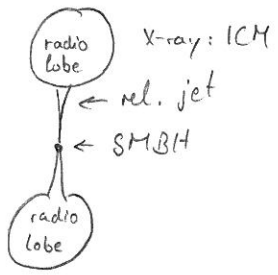
$$\frac{E_{AGN}}{M_{gas}} \sim \frac{10^{63}\ erg\ 10^{-24}\ g}{10^{14} \cdot 2 \cdot 10^{33}\ g} \sim \frac{1}{2}\,10^{-8}\,\frac{TeV}{nucleon} \sim 5\ keV\ nucleon^{-1} \qquad (3.88)$$

→ energetically this is a more promising mechanism!

- <u>emerging picture:</u>

radio-emitting bubbles embedded in X-ray emitting ICM with $n_{ambient} \sim 10^{-2}\,cm^{-3}$, $kT \sim 3$ keV

bubble radius $r_{bubble} \sim 10-30$ kpc
distance to SMBH $R \sim 20-60$ kpc




- the centers of many cool core clusters with low-entropy gas whose cooling time is less than the age of the universe also contain AGN; cooling gas can form stars and feed the accretion disk of the AGN; accreting gas has to lose its angular momentum which can be used to trigger a relativistic jet that is composed of cosmic rays and magnetic fields; it slows down due to the ram pressure of the ambient ICM and blows lobes of relativistic plasma; as lobes detach from jets they find themselves underneath the heavier thermal plasma of the ICM → convectively unstable → bubble rises buoyantly and subsonically
- the relativistic jets displace the ICM at the location of the cavities, i.e. they do $pdV$ work against ICM, as well as supplying internal energy to the cavities; hence the total energy required to create the cavity is equal to its enthalpy,

$$H = U + pV = \frac{1}{\gamma_b - 1} PV + PV = \frac{\gamma_b}{\gamma_b - 1} PV = 4PV, \qquad (3.89)$$

where we used $\gamma_b = \frac{4}{3}$ (for a relativistic filling); of this $4PV$, only $1PV$ is directly available for mechanical work on the surroundings while $3PV$ are stored as internal energy; hence the work done by the 2 bubbles is

$$W = PV = 2 \cdot \frac{4}{3}\pi r_b^3 \cdot n_a kT \sim 10^{59}\ erg \qquad (3.90)$$

where we used $r_b \sim 20$ kpc, $n_a \sim 10^{-2}\,cm^{-3}$, $kT \sim 3$ keV.

- there are 3 different ways to estimate the bubble's rise time, using 1) the sound crossing time, 2) the buoyant rise time, and 3) the time required for the ambient medium to refill the displaced volume as the bubble rises upward.

1) sound crossing time of distance from cavity center to SMBH using $\gamma_a = \frac{5}{3}$ for the ambient ICM:

$$t_{cs} \sim \frac{R}{\sqrt{\gamma_a \frac{kT}{\mu m_p}}} = R\sqrt{\frac{\mu m_p}{\gamma_a kT}} \sim 4\times 10^7 \text{ yr} \left(\frac{R}{40\text{ kpc}}\right)\left(\frac{kT}{3\text{ keV}}\right)^{-1/2} \quad (3.91)$$

2) to estimate the buoyancy time, we compute the buoyancy force acting upon the bubble,

$$\vec{F}_{buoy} = -\vec{g}\, V(\rho_a - \rho_b) \quad (3.92)$$

where $\vec{g}$ is the gravitational acceleration (assuming hydrostatic equilibrium of the ambient gas), $V$ is the bubble volume, $\rho_a$ and $\rho_b$ denote the mass density of the ambient gas and the bubble; the ram pressure exerts a drag force on the bubble, oppositely directed to rise velocity

$$\vec{F}_d \simeq -\frac{C}{2}\,\sigma\,\rho_a\, v^2\,\frac{\vec{v}}{v} \quad (3.93)$$

where $\sigma$ is the cross section of the bubble, $C$ is the drag coefficient and depends on bubble geometry and Reynolds number (whether the flow is turbulent or laminar); $C \sim 0.6$ for Mach $\simeq 0.7$; in equilibrium, the terminal velocity is obtained by balancing $F_{buoy}$ and $F_d$ yielding

$$v \simeq \sqrt{g\,\frac{V}{\sigma}\,\frac{2}{C}\,\frac{\rho_a - \rho_b}{\rho_a}} \sim \sqrt{g\,\frac{V}{\sigma}\,\frac{2}{C}} \quad (3.94)$$

in the last step, we assumed $\rho_b \ll \rho_a$; for a singular isothermal sphere (SIS), we write down $g \approx \frac{v_c^2}{R} = \frac{2\sigma^2}{R} = \frac{2kT}{\mu m_p R}$; (3.95)

with $\sigma = \pi r^2$ and $V = \frac{4}{3}\pi r^3$, we obtain $\frac{\sigma}{V} = \frac{3}{4r}$; hence

$$t_{buoy} \sim R\sqrt{\frac{C}{2}\,\frac{\sigma}{V}\,\frac{1}{g}} \sim R\sqrt{\frac{C}{2}\,\frac{3}{4r}\,\frac{\mu m_p R\,\gamma_a}{2kT\,\gamma_a}} \sim t_{cs}\sqrt{\frac{3C\gamma_a}{16}\,\frac{R}{r}} \quad (3.96)$$

$$t_{buoy} \sim 0.6\, t_{cs}\left(\frac{R}{2r}\right)^{1/2}$$

3) the time required to refill the volume as the bubble rises upward is

$$t_r \sim 2R\sqrt{\frac{r}{GM(R)\,R}} \sim R\sqrt{\frac{4\pi\mu m_p}{R\,2kT\gamma_a}} \sim t_{cs}\sqrt{\frac{2\gamma_a r}{R}} \sim 1.3\, t_{cs}\left(\frac{2r}{R}\right)^{1/2} \quad (3.97)$$

using the potential of the SIS, $\frac{GM}{R} = \frac{2kT}{\mu m_p}$

=> all 3 estimates provide similar results (albeit with a different scaling in the ratio $\frac{r}{R}$)

=> we finally obtain the AGN heating rate (combining (3.90) and (3.96)):

$$L_{AGN} \sim \frac{PV}{t_{buoy}} \sim \frac{10^{59}\,erg}{10^{15}\,s} \sim 10^{44}\,\frac{erg}{s} \sim L_X \quad (3.98)$$

i.e. comparable to the X-ray "cooling" luminosity (3.83)

- open questions (active research):
  - how is accretion output thermalized? suggestions include dissipation of non-linear waves, turbulence, CR streaming losses (that can excite Alfvén waves in the magnetized ICM which get damped)
  - is heating/cooling balance stable? possibly conduction is needed (-> Sect. 3.2.3)
  - how is accretion rate tuned?

Schwarzschild radius of SMBH is

$$r_{SMBH} = \frac{2GM_{SMBH}}{c^2} \sim \frac{2\frac{2}{3}10^{-7}\cdot 5\cdot 10^9\, 2\cdot 10^{33}\,cm}{10^{21}} \sim \frac{4}{3}\cdot 10^{15}\,cm \sim 1\ \text{light-day} \quad (3.99)$$

and cooling occurs on scales of $\sim 30\,kpc \sim 10^{23}\,cm \sim 10^8\, r_{SMBH}$!

=> at least Nature finds a way to solve all of these problems, because observationally SMBH activity accompanies transition to complexity when $t_{cool} \lesssim 1\,Gyr$

## 3.2.3 Transport Processes of Gas and Thermal Stability

- in Sect. 3.1.2, we talked about convection as an effective transport process of gas; here, we discuss 2 other important transport processes: turbulent transport and thermal conduction (which is merely energy but ionic particle transport)

> "Big whirls have little whirls,
> That feed on their velocity;
> And little whirls have lesser whirls,
> And so on to viscosity."
> – Lewis Fry Richardson

### 3.2.3.1 Turbulence

- we start with the Navier-Stokes equation for an incompressible fluid ($\vec{\nabla}\cdot\vec{u}=0$) and obtain from (3.35)–(3.37):

$$\frac{\partial \vec{u}}{\partial t} + \underbrace{\left(\vec{u}\cdot\vec{\nabla}\right)\vec{u}}_{(1)} = \underbrace{\vec{g}}_{(2)} - \underbrace{\frac{1}{\rho}\vec{\nabla}p}_{(3)} + \underbrace{\nu\,\Delta\vec{u}}_{(4)} \qquad (3.100)$$

where we defined the kinematic viscosity $\nu = \frac{\mu}{\rho}$; the terms have the following meaning: 1) advective transport, 2) external force (e.g. gravity), 3) pressure force, 4) dissipation term;

- we compare time scales for advection $\tau_{adv}$ and for viscous dissipation $\tau_{diss}$:

$$\tau_{adv} = \frac{L}{u} \quad \text{and} \quad \tau_{diss} = \frac{L^2}{\nu} \quad \text{where } [\nu] = cm^2\,s^{-1} \qquad (3.101)$$

and define the Reynolds number

$$Re = \frac{\tau_{diss}}{\tau_{adv}} = \frac{Lu}{\nu} = \frac{Lu}{\lambda_{mfp}\,v_{th}} \qquad (3.102)$$

here $L$ and $u$ are characteristic length and velocity scales of the (macroscopic) system, and the kinematic viscosity is the product of thermal velocity and particle mean free path;

note that the assumption of incompressible turbulence

$$\vec{\nabla}\cdot\vec{u}=0 \;\Rightarrow\; \vec{k}\cdot\vec{u}=0, \qquad \vec{u}(\vec{x},t) = \int \vec{\hat{u}}(\vec{k},\omega)\, e^{i(\vec{k}\vec{x}-\omega t)}\, d^3k\, d\omega \qquad (3.103)$$

implies no longitudinal disturbances (sound waves), but only allows for rotational flows – "eddies"

- if $Re \gg 1$, advection is much faster than dissipation, which cannot stabilize the dynamical growth; the vertical fluid motions interact nonlinearly and turbulence sets in; in 2D small eddies merge to form larger eddies and energy flows from small to large scales along an "inverse cascade"; in 3D energy is being transported from large to small scales until it is dissipated through the production of viscous heat on sufficiently small scales of the mean free path
- Let $\lambda$ be the size of an eddy and $u_\lambda$ the typical rotational velocity across the eddy; the energy flow through such an eddy is

$$\dot{\varepsilon} \approx \underbrace{\left(\frac{u_\lambda^2}{2}\right)}_{\text{kinetic energy}} \cdot \underbrace{\left(\frac{\lambda}{u_\lambda}\right)^{-1}}_{\substack{\text{eddy turnover} \\ \text{time scale}}} \approx \frac{u_\lambda^3}{\lambda} \qquad (3.104)$$

- energy is being fed into the turbulent cascade on the macroscopic "injection scale" $L$ with a typical velocity $u$; through non-linear interactions of fluid motions eddies break up into smaller eddies until the energy gets dissipated on the "viscous scale" $\lambda_{visc}$; in between, on scales $\lambda$ of the "inertial range"

$$\lambda_{visc} < \lambda < L \qquad (3.105)$$

the energy flow must be independent of scale because the energy cannot be accumulated anywhere in steady state: the only channel for the energy transfer is through non-linear interactions with other eddies and hence, we obtain the velocity scaling from (3.104) as $u_\lambda \propto \lambda^{1/3}$ or

$$u_\lambda \approx u\left(\frac{\lambda}{L}\right)^{1/3} \qquad (3.106)$$

- the largest eddies assume the largest velocities (and kinetic energies), but the smallest have the highest vorticity

$$\Omega \approx \frac{u_\lambda}{\lambda} \approx \frac{u}{(\lambda^2 L)^{1/3}} \qquad (3.107)$$

- to compute the power spectrum of the eddy velocity $u_\lambda \approx (\dot{\varepsilon}\lambda)^{1/3}$, we write down the correlation function, which scales as

$$\xi_v \propto u_\lambda^2 \propto (\dot{\varepsilon}\lambda)^{2/3} \tag{3.108}$$

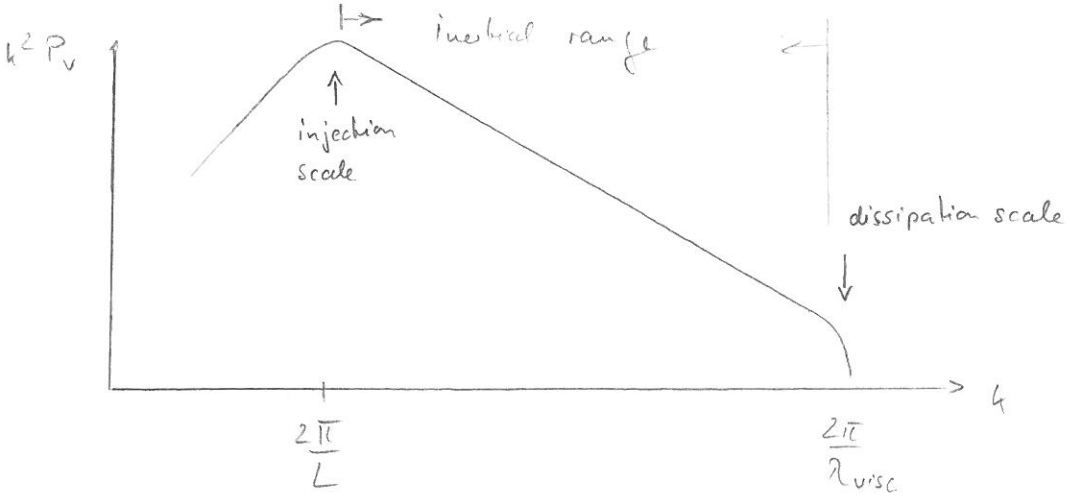
the Fourier transform of $\xi_v$, the power spectrum $P_v$ then scales as

$$P_v \propto \lambda^3 \xi_v \propto k^{-3} (\dot{\varepsilon} k^{-1})^{2/3} \propto \dot{\varepsilon}^{2/3} k^{-11/3} \tag{3.109}$$

the power per logarithmic $k$ interval scales as

$$k^2 P_v \propto \dot{\varepsilon}^{2/3} k^{-5/3} \tag{3.110}$$

which is the Kolmogorov turbulence spectrum of driven turbulence




- in contrast to driven subsonic turbulence, in clusters we encounter decaying turbulence, i.e. after a merger injects the energy on scales $L \sim r_c$ which will decay after a few eddy turnover time scales $\frac{L}{u}$

- Implications of turbulent transport in clusters:
  - mixing of metals
  - shredding AGN bubbles and mixing relativistic components with the ICM
    → possible source of magnetization and $\beta$-amplification of the ICM
  - heating cool cores and arresting the overcooling in them

## 3.2.3.2 Heat Conduction

- a system can be in hydrostatic equilibrium, but out of thermal equilibrium; the entropy equation reads (neglecting velocity gradients)

$$\rho T \frac{ds}{dt} = \vec{\nabla} \cdot (\kappa \nabla T) \tag{3.111}$$

- using $dq = T\,ds$ and $c_p = \left(\frac{dq}{dT}\right)_p$ (3.112)

we get $$c_p\,dT = T\,ds \quad \Rightarrow \quad ds = c_p\,d\ln T \tag{3.113}$$

- hence we can rewrite (3.111)

$$\rho c_p \frac{dT}{dt} = \kappa \vec{\nabla}^2 T \quad \text{or}$$

$$\frac{dT}{dt} = \chi \vec{\nabla}^2 T \quad \text{with} \quad \chi \equiv \frac{\kappa}{\rho c_p} \tag{3.114}$$

$\Rightarrow$ the temperature can only change as a result of thermal conduction if $\vec{\nabla} T \neq 0$ since the temperature gradient is the source of free energy

- we now want to estimate the heat conductivity $\kappa$; to this end, we consider a system in thermal equilibrium with $T$ whose particles are moving randomly in all directions, $\Delta A$ denotes the area of a screen perpendicular to the $x$ axis
- the number of particles, that fly per unit time through the screen from one side to the other is given by

$$\frac{\Delta N}{\Delta t} = \frac{n v \Delta A}{6}$$

where the factor of 6 arises because on average $\frac{1}{3}$ of all particles fly along the $x$ axis and of those, only $\frac{1}{2}$ in either direction

- the particle mean free path is $\lambda = (n\sigma)^{-1}$ with $\sigma$ as the collisional cross section; particles at $x - \lambda$ transport gas properties to $x$ and vice versa; this is particularly important for gradients in gas properties that will be smoothed out as a result of such a transport

- hence in the presence of a density gradient, $\frac{\partial n}{\partial x} \neq 0$, the net number of particles flying from the denser to the more dilute region

$$\frac{\Delta N}{\Delta t} = \frac{\Delta n\,(x+\lambda)\, v\, \Delta A}{\Delta x\, 6} - \frac{\Delta n\,(x-\lambda)\, v\, \Delta A}{\Delta x\, 6} \approx \frac{v\,\Delta A}{6}\, 2\, \frac{\partial n}{\partial x}\, \lambda$$

assuming $\lambda \ll \Delta x$, the typical length scale of the number density gradient

- the diffusion coefficient that relates the particle current $j = \frac{\Delta N}{\Delta t\, \Delta A}$ to the number density gradient is given by

$$\frac{\Delta N}{\Delta t\, \Delta A} \stackrel{!}{=} D\, \frac{\partial n}{\partial x} \qquad \text{with} \qquad D = \frac{v\,\lambda}{3} = \frac{v}{3 n \sigma} \tag{3.115}$$

- if the temperature changes along $x$, $\frac{\partial T}{\partial x} \neq 0$, the particles transport energy

$$\frac{\Delta \varepsilon}{\Delta t\, \Delta A} = \frac{n v\, \Delta A}{6\, \Delta A\, \Delta x} \left[ \varepsilon(x+\lambda) - \varepsilon(x-\lambda) \right]$$

$$= \frac{n v \lambda}{3} \left( \frac{\partial \varepsilon}{\partial T}\, \frac{\partial T}{\partial x} \right) = \frac{n v c_v \lambda}{3}\, \frac{\partial T}{\partial x}$$

where $c_v$ is the heat capacity at constant volume (3.3); hence we find

$$\frac{\Delta \varepsilon}{\Delta t\, \Delta A} \stackrel{!}{=} \kappa\, \frac{\partial T}{\partial x} \qquad \text{with} \qquad \kappa = \frac{n v c_v \lambda}{3} \stackrel{\text{simple compound gas}}{=} \frac{v\, C_v}{3\sigma} = \frac{v\, k}{2\sigma} \tag{3.116}$$

where we used $c_v = \frac{3}{2} k$ (3.7) in the last step (implying an ideal, monoatomic gas)

$\kappa$ = heat conductivity $\quad \left[ \frac{\text{erg}}{\text{cm s K}} \right]$

- heat is conducted by electrons since they move faster than ions by $\frac{v_e}{v_i} = \sqrt{\frac{m_i}{m_e}} \approx 43 \sqrt{Z}$ (assuming $T_e = T_i$ which applies to the ICM except for immediate post-shock regions); the electron mean free path is determined by ion number density and scattering cross section, implying $\lambda = \frac{1}{n_i \sigma}$

interlude: Coulomb Logarithm:

· let's first consider an electron scattering in the Coulomb field of an ion:

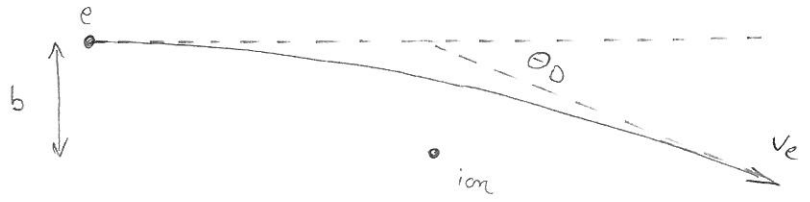
① if the deflection angle is small, $\theta_D \ll 1$, we can approximate its value by computing the perpendicular impulse exerted by the ion's Coulomb field, integrating along the electron's unperturbed straight line trajectory ("Born approximation")

$$m_e v_e \theta_D = \int_{-\infty}^{\infty} \vec{\nabla}_\perp \phi_i \, dt = \int_{-\infty}^{\infty} \frac{\partial}{\partial b} \left( \frac{Ze^2}{\sqrt{b^2 + v_e^2 t^2}} \right) dt = \int_{-\infty}^{\infty} \frac{Ze^2 b \, dt}{(b^2 + v_e^2 t^2)^{3/2}}$$

$$= \frac{Ze^2}{v_e} \int \frac{b^2 \, dx}{b^3 (1+x^2)^{3/2}} \qquad x = \frac{v_e t}{b}$$

$$= \frac{Ze^2}{v_e b} \left. \frac{x}{\sqrt{1+x^2}} \right|_{-\infty}^{\infty} = \frac{2Ze^2}{b v_e} \qquad (3.117)$$

$$\Rightarrow \quad \theta_D = \frac{b_0}{b} \quad \text{for } b \gg b_0 = \frac{2Ze^2}{m_e v_e^2} \qquad (3.118)$$




② if the dominant source of this deflection were a single large-angle scattering event, then the relevant cross section would be $\sigma = \pi b_0^2$ (since all impact parameters $\leq b_0$ produce large-angle scatterings), and the mean deflection frequency $\nu_D$ and time $t_D$ would be

$$\nu_D \equiv \frac{1}{t_D} = n_i \sigma v_e = n_i \pi b_0^2 v_e \quad \begin{pmatrix} \text{single large-angle scattering} \\ \text{of electron by an ion} \end{pmatrix} \qquad (3.119)$$

③ the cumulative, random-walk effects of many small-angle scatterings off ions produce a net deflection of order a radian in a shorter time; as the directions of the individual scatterings are random, the mean deflection angle after many scatterings vanishes, $\langle \theta \rangle = 0$; however $\langle \theta^2 \rangle$ will not vanish, where $\langle \theta^2 \rangle = \sum_{\substack{\text{all} \\ \text{encounters}}} \theta_D^2 = \sum \left( \frac{b_0}{b} \right)^2$

(3) cont'd:

- the number of encounters, that occur with impact parameters between $b$ and $b+db$ during time $t$ is $n_i v_e t\, 2\pi b\, db$, hence the mean square deflection angle accumulates up to

$$\langle \theta^2 \rangle = \int_{b_{min}}^{b_{max}} \left(\frac{b_0}{b}\right)^2 n_i v_e t\, 2\pi b\, db = n_i 2\pi b_0^2 v_e t \ln\left(\frac{b_{max}}{b_{min}}\right) \quad (3.120)$$

- the integral diverges logarithmically, but physics saves the day:

$$b_{min} = \frac{2e^2}{kT} \quad (3.121)$$

is the radius, where the Coulomb energy of the electron in the field of the ion vanishes, $U = \frac{m}{2} v^2 - \frac{Ze^2}{b_{min}} \overset{!}{=} 0$; and $b_{max}$ is given by the maximum distance over which electric fields of individual particles can reach without being screened by the oppositely charged particles in a plasma, the Debye length $\lambda_D$:

$$d_{max} = \lambda_D = \sqrt{\frac{kT}{4\pi n_e Z e^2}} \quad (3.122)$$

- hence we can define the Coulomb logarithm

$$\ln \Lambda = \ln \frac{b_{max}}{b_{min}} = \ln \sqrt{\frac{kT^3}{n\ 4\pi Z^3 e^6}} \quad (3.123)$$

$$\ln \Lambda = 35 - \frac{1}{2} \ln\left(\frac{n}{10^{-2}\,cm^{-3}}\right) + \frac{3}{2} \ln\left(\frac{kT}{keV}\right)$$

- the value of $t$, that implies $\langle \theta^2 \rangle \approx 1$ is the deflection time $t_D$

$$\nu_D^{ei} = \frac{1}{t_D^{ei}} = n_i 2\pi b_0^2 v_e \ln \Lambda = \frac{n_i\, 8\pi Z^2 e^4}{m_e^2 v_e^3} \ln \Lambda \quad (3.124)$$

and $\ln \Lambda \sim 35 \ldots 40$ in the ICM

$\Rightarrow$ this deflection frequency is larger, by a factor $2 \ln \Lambda$ than the frequency in (3.119) for a single large-angle scattering

- back to our heat conductivity of electrons (3.116)

$$\kappa = \frac{n_e v_e c_v \lambda}{3} = \frac{n_e v_e c_v}{3 \sigma n_i} \qquad (3.125)$$

from (3.124), we can read off $\sigma$ by remembering $\nu_D = n_i \sigma v_e$:

$$\sigma = 2\pi b_0^2 \ln \Lambda = \frac{8\pi Z^2 e^4 \ln \Lambda}{m_e^2 v_e^4} \qquad (3.126)$$

this yields the heat conductivity of electrons scattered by ions

$$\kappa = \frac{n_e v_e}{3} c_v \frac{m_e^2 v_e^4}{8\pi n_i Z^2 e^4 \ln \Lambda} = \frac{1}{3}\left(\frac{m_e^2}{8\pi Z^2 e^4}\right)\left(\frac{n_e}{n_i}\right)\frac{c_v v_e^5}{\ln \Lambda} \qquad (3.127)$$

- in a thermal electron gas, $c_v = \frac{3k}{2}$ and

$$v_e = \left(\frac{3kT_e}{m_e}\right)^{1/2}$$

yielding the heat capacity

$$\kappa = \frac{k}{2}\left(\frac{m_e^2}{8\pi Z^2 e^4}\right)\left(\frac{n_e}{n_i}\right)\left(\frac{3kT_e}{m_e}\right)^{5/2}\frac{1}{\ln \Lambda} \qquad (3.128)$$

$$\kappa = 1.7\ 10^{-7}\left(\frac{T}{1K}\right)^{5/2}\frac{erg}{sK\,cm} = 1.7 \cdot 10^{13}\left(\frac{T}{10^8 K}\right)^{5/2}\frac{erg}{sK\,cm}\cdot\left(\frac{\ln \Lambda}{35}\right)^{-1}$$

(using $\ln \Lambda = 35$ appropriate for cool cores)

## 3.2.3.3. Thermal Stability

### Fields length:

- cool star forming clouds should appear only in systems whose size is greater than a critical length scale, known as the "Fields length" below which thermal conduction smoothes out temperature inhomogeneities; one can derive the Field length heuristically by considering thermal balance for a cool cloud of radius $r$ embedded in a medium of temperature $T$
- electron thermal conduction sends energy into the cloud at a rate

$$H_{cond} \sim r^2 \kappa(T) \frac{T}{r} \sim \kappa_0 f_c r \frac{T^{7/2}}{T_8^{5/2}} \qquad (3.129)$$

with the Spitzer conductivity and a magnetic suppression factor $f_c$ (depending on magnetic topology)

$$\kappa(T) = 6 \cdot 10^{13} \left(\frac{T}{T_8}\right)^{5/2} f_c \frac{erg}{s\,K\,cm}, \quad T_8 = 10^8 K \qquad (3.130)$$

- radiative cooling can get rid of the energy at a rate

$$\mathcal{C}(T) \sim r^3 n_H^2 \Lambda_0(T) \sim r^3 n_H^2 \frac{T^{1/2}}{T_8^{1/2}} \Lambda_0 \qquad (3.131)$$

with $\Lambda_0(T) \approx 2.5 \cdot 10^{-29} \frac{erg\,cm^3}{s} \left(\frac{T}{T_8}\right)^{1/2}$ cf. (3.81)

- cooling and conduction are thus in approximate balance for systems with a radius of order the Field length $\lambda_F$

$$H_{cond} \sim \mathcal{C} \Rightarrow \lambda_F \equiv \left[\frac{T\kappa(T)}{n_H^2 \Lambda_0(T)}\right]^{1/2} = \left(\frac{\kappa_0 f_c T_8^{1/2} x_e^2}{\Lambda_0 T_8^{5/2} k_8^2 k_3}\right)^{1/2} \left(\frac{T^{7/2} k_8^3}{T^{1/2} n_e^2}\right)^{1/2} \qquad (3.132)$$

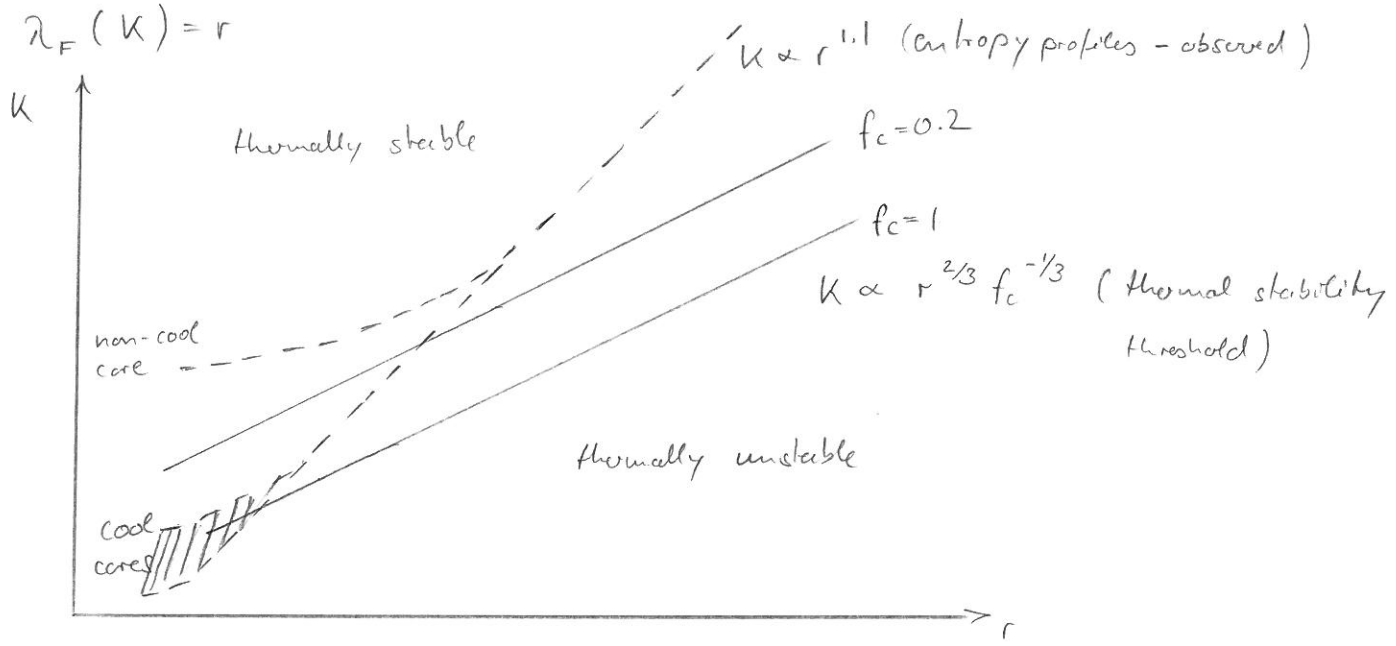
$$\curvearrowright \lambda_F = \left(\frac{\kappa_0 f_c x_e^2}{k_3 \Lambda_0 k_8^2 T_8^2}\right)^{1/2} K_e^{3/2} \approx 4\,kpc \left(\frac{K_e}{10\,keV\,cm^2}\right)^{3/2} f_c^{1/2}$$

where we used $n_H^2 = \frac{S^2}{m_p^2} X_H^2 \frac{x_e^2}{x_e^2} = n_e^2 \frac{1}{x_e^2}$ (pg. 5), $K_e = \frac{k_8 T_e}{n_e^{2/3}}$;

through a coincidence of scaling, the Field length is a function of entropy alone when free-free is the dominant cooling mechanism

- we can translate this criterion (3.182) in the entropy - radius plane by adopting

$\lambda_F(K) = r$




below each line $K \propto r^{2/3} = \lambda_F^{2/3}$, gas within radius r constitutes a subsystem with $r > \lambda_F$ (at constant K); i.e. the amount of entropy in the large cloud is too small to support fast enough conduction to prevent a cooling runaway, allowing multiphase gas to persist and star formation to proceed; above each line is the region of stability, in which conduction leads to evaporation and homogeneity

## Heating vs. Cooling – a Visual Stability Analysis

- formally one would have to do a perturbation analysis of the hydrodynamic equations (Navier-Stokes + entropy equations with thermal conduction), but here, we will only sketch the ideas
- since we allow for thermal conduction, entropy is not anymore conserved in a given fluid element; instead, we consider hydrostatic rearrangements that conserve the thermal pressure $P = n\,kT$ and rewrite the energy deposition and cooling rates as a function of temperature and pressure using $n \propto \frac{P}{T}$; we consider radiative cooling (bremsstrahlung, line-emission), and heating by conduction, turbulence and Coulomb interactions of cosmic rays with thermal electrons:

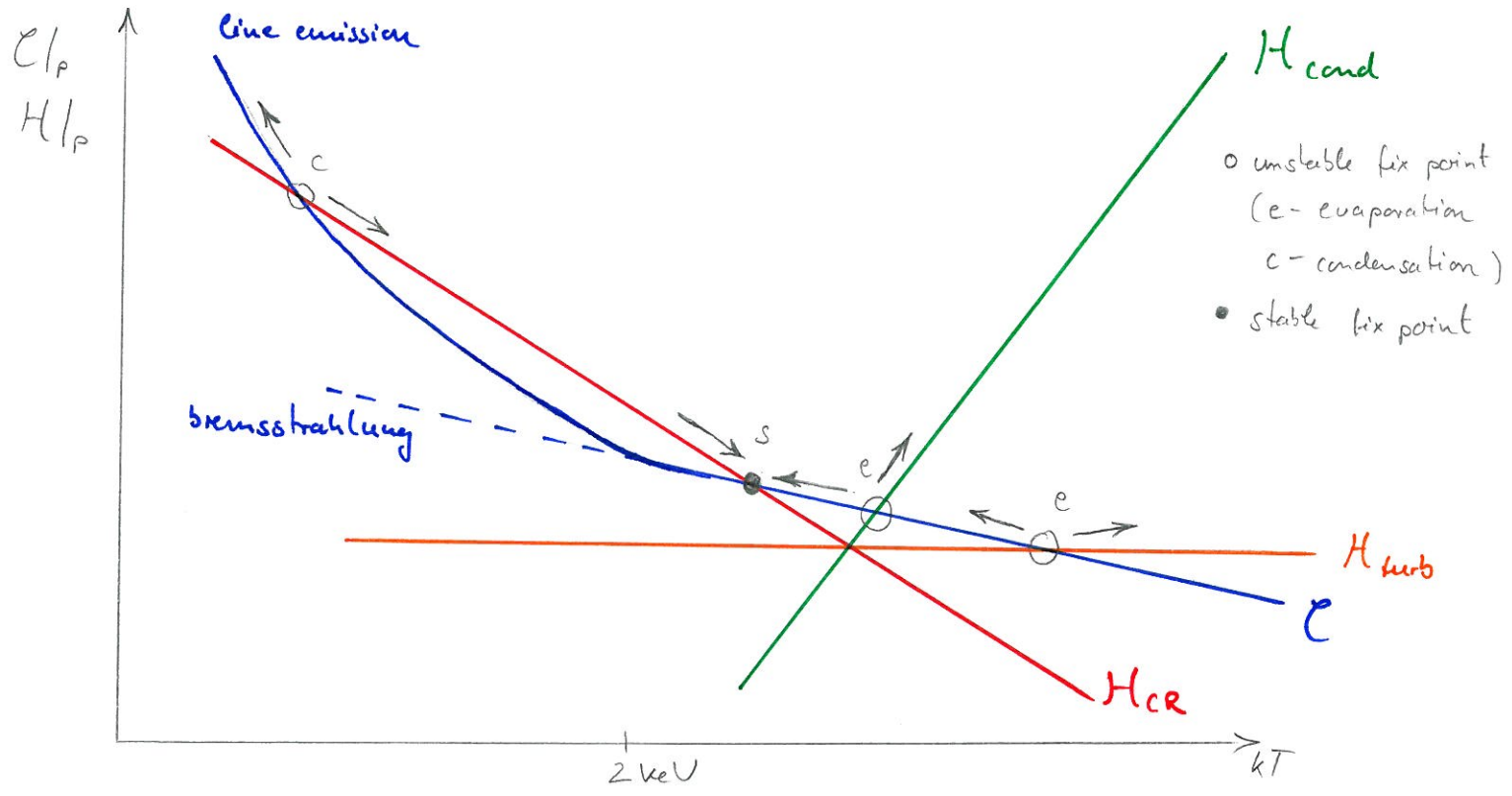
$$\mathcal{C} \propto r^3 n^2 \left[ T^{1/2} + \Lambda_{line}(T) \right] \sim P \left[ T^{-1/2} + \frac{\Lambda_{line}(T)}{T} \right] \qquad (3.132)$$

where $\Lambda_{line}(T) \propto T^{\alpha_{line}}$, $\alpha_{line} < \frac{1}{2}$ for $kT < 2\,keV$

$$\mathcal{H}_{cond} \propto r^2 \kappa(T) \frac{T}{r} \propto r\, T^{7/2} \propto n^{-1/3} \frac{T^{1/3}}{T^{1/3}} T^{7/2} \propto P^{-1/3} T^{23/6} \qquad (3.133)$$

$$\mathcal{H}_{turb} \propto r^3 n \sim T^0 \qquad (3.134)$$

$$\mathcal{H}_{CR} \propto r^3 n\, n_{CR} \propto f_{CR}\, n \propto f_{CR} \frac{P}{T} \quad \left( f_{CR} = \frac{n_{CR}}{n_{th}} \right) \qquad (3.135)$$



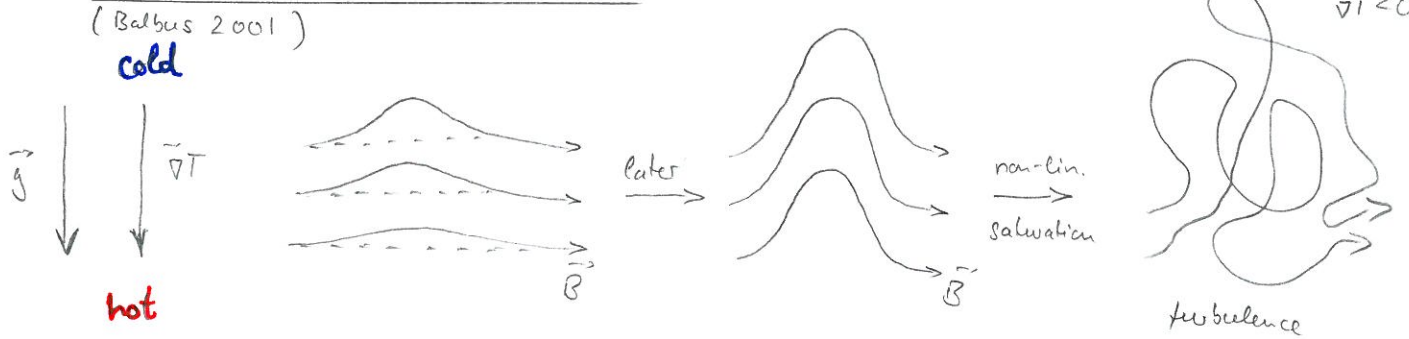
- this demonstrates, that conductive and turbulent heating cannot be in stable equilibrium with radiative cooling; in contrast, a heating mechanism with an energy deposition rate that scales with $T^{-1}$ (such as cosmic ray heating through Coulomb interactions) allows for stable solutions
- of course the final state of the system depends on boundary conditions and conservation laws:
  - conduction: free energy is borrowed from $\vec{\nabla} T$ → endstate has $T$ = const. = $T_{vir}$
  - turbulence: $T$ rises until $3kT \sim m v^2$; or if all turbulent energy has been dissipated
  - cosmic rays: can only dissipate the total CR energy on the corresponding time scale

## 3.2.3.4 Thermal Stability with Magnetic Fields

- in the presence of magnetic fields, electrons cannot anymore move "freely" but are bound to follow and gyrate around a given field line; this modifies the convective stability criterion and the type of instability depends on the sign of $\vec{\nabla} T$:

① magneto-thermal instability (MTI): applicable for outer cluster regions ($r \geq 0.2\, R_{200}$) where $\vec{\nabla} T < 0$

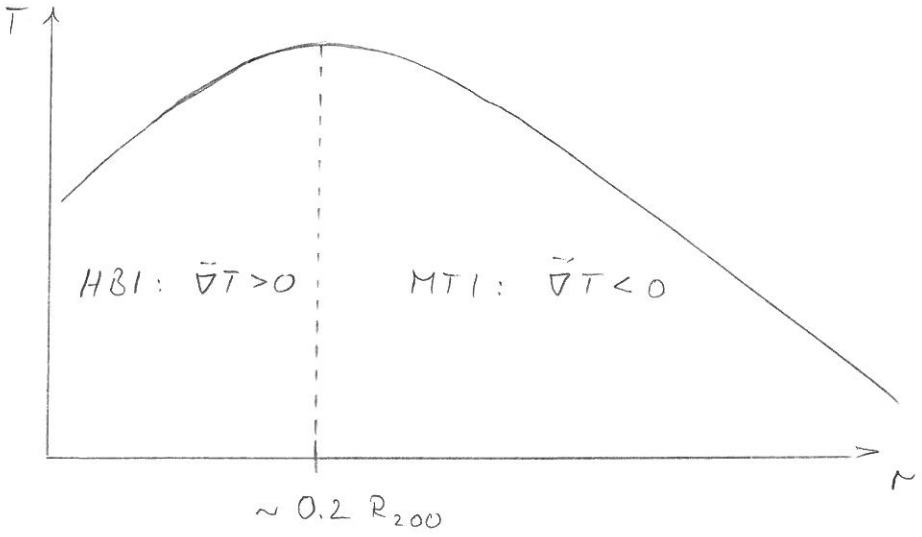
(Balbus 2001)




- displacing a volume element upwards in the gravitational potential would cause it to adiabatically expand and cool in the absence of conduction; instead it is conductively heated from the hotter heat bath below which causes further expansion and dilution; it continues to rise as it remains lighter than the surrounding ICM → instability; it turns out, that the MTI does not quasi-saturate with a radial field but in a turbulent state as such a radial field configuration is overstable, i.e. it always overshoots with a growing amplitude

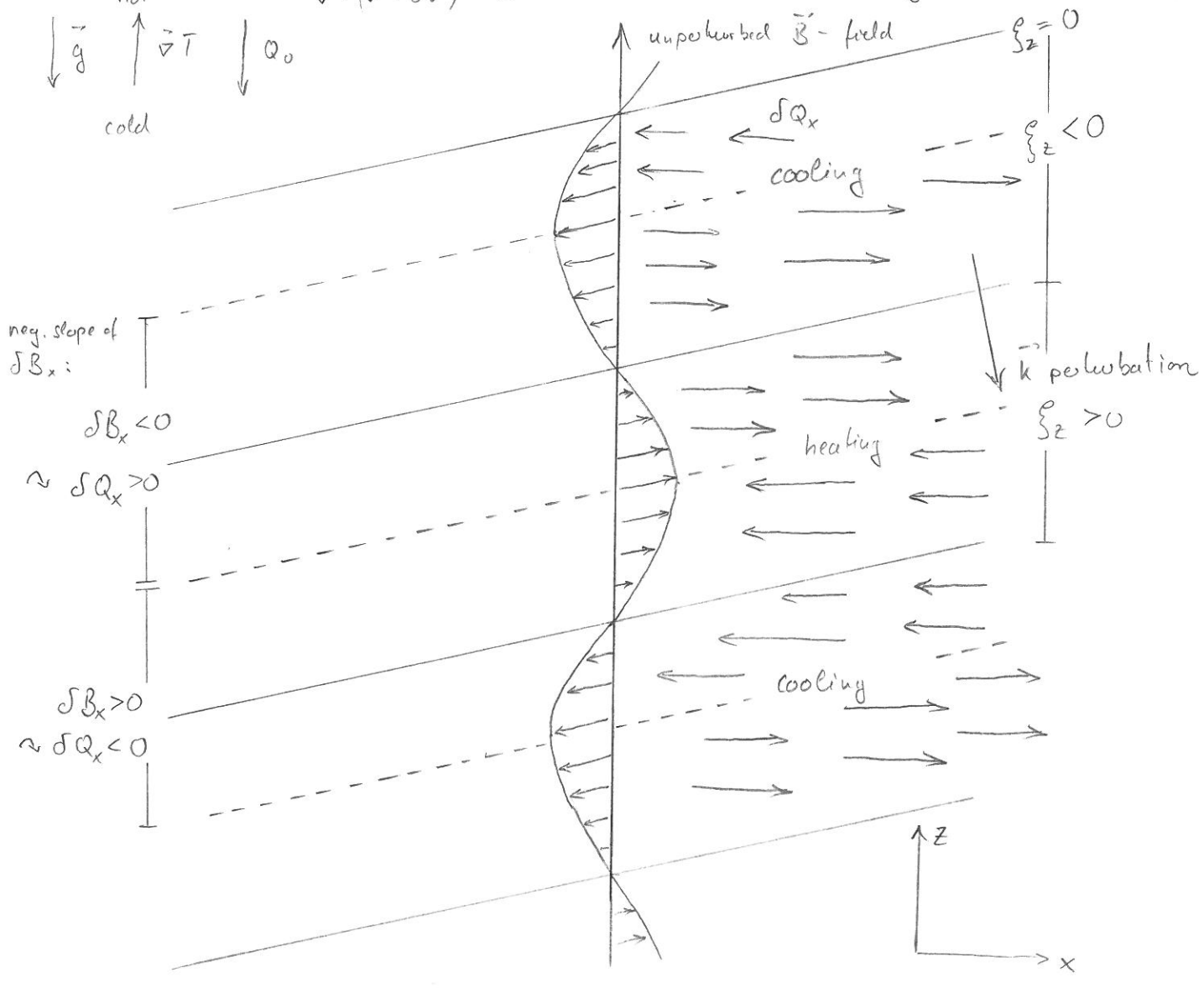
## (2) heat flux driven buoyancy instability (HBI): Quataert (2008)

- first, we look at the temperature profile of a cool core cluster




- we introduce the displacement field $\vec{\xi} \equiv \frac{i\delta\vec{v}}{\omega}$, which is proportional to $\delta\vec{v}$; since we are interested in incompressible gas and Fourier transform background quantities + perturbations:

$$\vec{\nabla}\cdot(\vec{v}+\delta\vec{v})=0 \quad \rightsquigarrow \quad \vec{k}\cdot\delta\vec{v}=0 \quad \rightsquigarrow \quad \vec{\xi}\perp\vec{k} \qquad (3.136)$$



- if the temperature gradient increases outward, there will be a background heat flux, $Q_0$, pointing inward along the unperturbed radial magnetic field lines (our assumed initial state); oblique perturbations with a wave vector $\vec{k}$ at some angle with the $\vec{B}$-field perturbs $\vec{B}$ in perpendicular direction (because of the incompressibility, $\vec{\nabla} \cdot \vec{v} = 0$); in the figure, we work out the perturbations to the heat flux along the x direction, $\delta Q_x$, in response to the displaced x component of the magnetic field (that is flux-frozen into the cluster plasma (as we will show in Sect. 3.3.1); the figure demonstrates that
  1) regions with a positive displacement field $\xi_z > 0$ experience a corresponding heat-flux perturbation $\delta Q_x$ implying heating; this causes the upwards-displaced fluid elements to rise further => instability
  2) regions with $\xi_z < 0$ experience cooling; they get denser and heavier and continue to sink in the gravitational potential => instability

- it turns out that this instability saturates quiescently in the non-linear regime with the field vectors aligning horizontally (as suggested by the discussion of of the linear regime of the instability); i.e. within gravitational equipotential surfaces (shells of constant radius for spherical symmetry)
  => this suppresses the heat flux inward by a large factor and thermally isolates the cooling core, which should enter a cooling catastrophe
  => thermal conduction appears not to be the sole solution to the "cooling flow problem"!

## 3.3 Non-thermal Processes

- the observation of diffuse radio synchrotron emission in the form of radio halos and relics proofs the existence of cluster-filling magnetic fields and relativistic electrons with $E_e \sim (1-30)$ GeV; "(giant) radio halos" are centered on the cluster and similar in morphology as the X-rays, occur in X-ray-luminous clusters (selection effect?), and are unpolarized $\Rightarrow$ volume-filling magnetic fields (or are threading a large fraction thereof) since synchrotron emission is intrinsically polarized (and a superposition of causally uncorrelated radio-emitting patches becomes uncorrelated); "(giant) radio relics" are situated at the cluster outskirts, correlate with merger/accretion shocks, have a high degree of polarization $\Rightarrow$ small emission volume (consistent with acceleration of relativistic electrons at the shock and synchrotron cooling in post-shock regime); to date we know of $\gtrsim 80$ systems of halos/relics
- when polarized radio emission propagates through a magnetized medium, its plane of polarization rotates for a non-zero line-of-sight component of the magnetic field $\vec{B}$ due to the birefringent property of the plasma – Faraday rotation; the Faraday rotation angle is given by

$$\phi_{obs} = \lambda^2 \, RM + \phi_{intrinsic} \qquad (3.137)$$

$$RM(\vec{x}_\perp) = \frac{e^3}{2\pi m_e^2 c^4} \int_0^L n_e(\vec{x}_\perp, \ell) \, \vec{B} \cdot d\vec{\ell} \qquad (3.138)$$

$$\simeq 812 \, \frac{\text{rad}}{\text{m}^2} \, \frac{B}{\mu\text{G}} \, \frac{n_e}{10^{-3}\,\text{cm}^{-3}} \, \frac{L}{\text{Mpc}}$$

Faraday rotation measure (FRM) observations suggest central field strengths of 1–20 $\mu$G (with wider values applicable to cool cores)

- relativistic (cosmic ray, CR) electrons are plausibly accelerated at shocks; those should be equally good (if not better) in accelerating cosmic ray protons that have much longer cooling time than the CR electrons $\rightarrow$ CR protons should accumulate

- this section is meant to introduce the basic concepts of magnetic fields (origin, transport) and CRs (acceleration, transport) in a pedagogical manner and focusses on aspects important for galaxy clusters

## 3.3.1 Magnetic Fields

### 3.3.1.1 Generating Magnetic Fields – Biermann Battery

- generate magnetic field on small scales from inhomogeneities where $\vec{\nabla} n_e$ is not correlated with $\vec{\nabla} T_e$
- we start with Faraday's law,

$$\frac{\partial \vec{B}}{\partial t} = -c\, \vec{\nabla} \times \vec{E} \tag{3.139}$$

- non-relativistic Lorentz force for electrons + protons: (viscous forces are ~0 on large scales)

$$m_e \frac{d\vec{v}_e}{dt} = -e\left( \vec{E} + \frac{\vec{v}_e}{c} \times \vec{B} + \frac{1}{e n_e} \vec{\nabla} P_e \right) - \nu m_e \frac{}{n_e} (\vec{v}_e - \vec{v}_i) \tag{3.140}$$

$$m_p \frac{d\vec{v}_p}{dt} = e\left( \vec{E} + \frac{\vec{v}_p}{c} \times \vec{B} \right) \rightarrow \text{negligible, since protons move slower than electrons by } \sqrt{\frac{m_p}{m_e}} \text{ and can be considered almost at rest for the electrons}$$

- we assume a steady state, which is valid for time scales $\tau > \omega_{Pl}^{-1}$,

$\omega_{Pl}^2 = \frac{4\pi n_e e^2}{m_e}$, the "plasma frequency" (eigenmode of a plasma)

$$\vec{E} = -\frac{\vec{v}_e \times \vec{B}}{c} - \frac{\vec{\nabla} P_e}{e n_e} \qquad \left| \left(\frac{}{}\right) \cdot c \cdot \mathrm{curl} \right.$$

$$\rightsquigarrow \frac{\partial \vec{B}}{\partial t} = \vec{\nabla} \times (\vec{v}_e \times \vec{B}) + \frac{c}{e} \vec{\nabla} \times \left( \frac{\nabla P_e}{n_e} \right) \tag{3.141}$$

- we use $P_e = n_e k T_e$ and the identity

$$\vec{\nabla} \times (f \vec{\nabla} g) = (\vec{\nabla} f) \times (\vec{\nabla} g) \tag{3.142}$$

- 2nd term in (3.141):

$$\frac{1}{k}\vec{\nabla}\times\frac{\vec{\nabla}P}{n} = \vec{\nabla}\times\left(\frac{1}{n}\,\nabla(nT)\right) = \underbrace{\vec{\nabla}\times(\nabla T)}_{=0} + \vec{\nabla}\times\left(\frac{T}{n}\,\nabla n\right)$$

$$\overset{(3.142)}{=} \vec{\nabla}\left(\frac{T}{n}\right)\times\vec{\nabla}n = \frac{1}{n}\vec{\nabla}T\times\vec{\nabla}n - \underbrace{\frac{T}{n^2}\vec{\nabla}n\times\vec{\nabla}n}_{=0} = \frac{1}{n}\vec{\nabla}T\times\vec{\nabla}n$$

- hence we obtain the battery equation:

$$\frac{\partial\vec{B}}{\partial t} = \vec{\nabla}\times(\vec{v}_e\times\vec{B}) - \frac{ck}{e\,n_e}\,\vec{\nabla}n_e\times\vec{\nabla}T_e \qquad (3.143)$$

=> if there is no magnetic to start with (i.e. a vanishing first term on the right-hand side), then the magnetic field can be generated by a baroclinic flow ($\vec{\nabla}n_e\times\vec{\nabla}T_e \neq 0$); this could be achieved in shocks in the intergalactic medium (ISM), ionization fronts, etc.; e.g. seeded by rotational motions at shocks of finite extent (chaotic collapse of protogalaxy) → picture

- typically, field strengths generated this way are very small as the following estimate shows with L as the characteristic gradient length and typical ISM cond's:

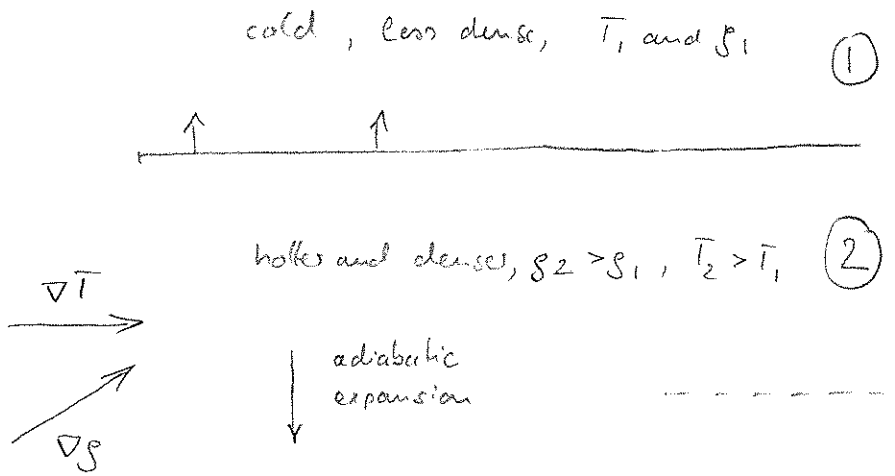
$$B \sim \frac{ckT}{e}\cdot\frac{t}{L^2} \sim \frac{ckT}{e}\,\frac{1}{\sqrt{G\rho}\,L^2} \sim 10^{-20}\,G\left(\frac{T}{10^4 K}\right)\left(\frac{n}{1\,cm^{-3}}\right)^{-1/2}\left(\frac{L}{kpc}\right)^{-2} \qquad (3.144)$$

naively, going to smaller L should increase B, but then we would have to account for adiabatic losses that necessarily accompany the expansion from small to large scales (see Sect. 3.3.1.2)

- hence, the big challenge consists in growing a coherent field from the small scale field → major challenge of dynamo theory; suggestions include
  1) diffusion of small-scale field fills volume of galaxies
  2) shearing and stretching in shocks
  3) blow-out by galactic winds and further dynamo amplification in the ICM

# Physics of Biermann battery

consider a shock of finite extent that propagates into zero-pressure medium




=> shocks of limited spatial extend break the barotropic relation $p_e = p_e(n_e)$ which couples $p_e$ to $n_e$

## 3.3.1.2 Evolution

- generally the evolution of $\vec{B}$ is given by magneto-hydrodynamics (MHD); for an inviscid (no viscosity); ideally conducting fluid, we simply add the Lorentz force to the momentum conservation (Euler equation) and supplement the system of equations by the equation for magnetic induction, i.e. (3.142) without the beraction term:

$$\begin{aligned} \rho\left(\partial_t \vec{v} + \vec{v}\cdot\vec{\nabla}\vec{v}\right) &= -\nabla p + \vec{j}\times\vec{B} \\ \partial_t \rho + \vec{\nabla}\cdot(\rho\vec{v}) &= 0 \\ \partial_t s + \vec{v}\cdot\vec{\nabla}s &= 0 \\ \partial_t \vec{B} - \vec{\nabla}\times(\vec{v}\times\vec{B}) &= 0 \end{aligned} \qquad (3.145)$$

- this is non-linear in $\vec{B}$ and in general has to be solved numerically; we will explore one crucial fact about $\vec{B}$-fields that their flux is "frozen" into the plasma; starting with the induction equation

$$\frac{\partial\vec{B}}{\partial t} = \vec{\nabla}\times(\vec{v}\times\vec{B}) \qquad (3.146)$$

- using $\vec{\nabla}\cdot\vec{B} = 0$ we obtain

$$\frac{\partial\vec{B}}{\partial t} = (\vec{B}\cdot\vec{\nabla})\vec{v} - (\vec{v}\cdot\vec{\nabla})\vec{B} - (\vec{\nabla}\cdot\vec{v})\vec{B}$$

$$\rightsquigarrow \frac{d}{dt}\vec{B} = \frac{\partial\vec{B}}{\partial t} + (\vec{v}\cdot\vec{\nabla})\vec{B} = (\vec{B}\cdot\vec{\nabla})\vec{v} - (\vec{\nabla}\cdot\vec{v})\vec{B} \qquad (3.147)$$

- with the continuity equation, $\frac{d}{dt}\rho = \frac{\partial}{\partial t}\rho + (\vec{v}\cdot\vec{\nabla})\rho = -(\vec{\nabla}\cdot\vec{v})\rho$, we obtain

$$\frac{d\vec{B}}{dt} = (\vec{B}\cdot\vec{\nabla})\vec{v} + \frac{\vec{B}}{\rho}\frac{d\rho}{dt} \qquad (3.148)$$

- multiplying (3.148) by $\frac{1}{\rho}$ and rearranging gives

$$\frac{1}{\rho}\frac{d\vec{B}}{dt} - \frac{\vec{B}}{\rho^2}\frac{d\rho}{dt} = \left(\frac{\vec{B}}{\rho}\cdot\vec{\nabla}\right)\vec{v}$$

$$\rightsquigarrow \underline{\frac{d}{dt}\left(\frac{\vec{B}}{\rho}\right) = \left(\frac{\vec{B}}{\rho}\cdot\vec{\nabla}\right)\vec{v}} \qquad (3.149)$$

→ this is the flux-freezing equation; to see this explicitly, consider the evolution of $\delta\vec{x}$ connecting 2 neighboring points in the fluid

$$\Delta\vec{x}(t) = \delta\vec{x}$$

$$\Delta\vec{x}(t+\Delta t) = \delta\vec{x} + (\delta\vec{x}\cdot\vec{\nabla})\vec{v}\,\Delta t + \mathcal{O}(t^2)$$

$$\frac{d\,\delta\vec{x}}{dt} \equiv \frac{\Delta\vec{x}(t+\Delta t) - \Delta\vec{x}(t)}{\Delta t} = (\delta\vec{x}\cdot\vec{\nabla})\vec{v} \qquad (3.150)$$

$\Rightarrow$ $\frac{\vec{B}}{\rho}$ and $\delta\vec{x}$ satisfy the same ordinary differential equation; hence, if initially $\delta\vec{x} = \varepsilon\frac{\vec{B}}{\rho}$, the same relation will hold for all times; if $\delta\vec{x}$ connects two particles on the same field line they remain on the same field line (neglecting diffusion effects)

- consider a uniform isotropic contraction of the plasma toward the origin; plasma in a sphere of radius $r$ (not necessarily at the origin) conserves mass and magnetic flux $d\Phi = \vec{B}\cdot d\vec{A}$ so that both $\rho r^3$ and $Br^2$ are constant, and hence

$$B \propto \rho^{\alpha_B}, \quad \alpha_B = \frac{2}{3} \qquad (3.151)$$

for isotropic contraction or expansion; flux freezing alone predicts a third relation between $B$ and $\rho$; note, the scaling exponent, $\alpha_B$, depends on the type of symmetry invoked during collapse (isotropic) and can differ for contraction along a homogeneous magnetic field ($\alpha_B = 0$) and perpendicular to it ($\alpha_B = 1$)

## 3.3.1.3 Magneto-hydrodynamic Waves and Turbulence

- this topic is generically subject to plasma physics; here we will give the basic picture while leaving the rigorous proofs for the "plasma physics" lecture
- linearizing the MHD equations (3.145), we have 8 equations for 8 unknowns $\delta\rho$, $\delta p$, $\delta\vec{B}$ and $\delta\vec{v}$; those are subject to the $\operatorname{div}\vec{B}=0$ constraint which reduces the dimensionality by 1; we then carry out a Fourier analysis of these perturbed quantities, e.g.

$$\delta\rho(\vec{x},t) = \int \delta\tilde{\rho}(\vec{k},t)\, e^{i\vec{k}\cdot\vec{r}}\, d^3k \qquad (3.152)$$

⇒ hence, we have seven different wave modes in a magnetized plasma that propagate forward in time (2 more than unmagnetized plasma):

•) 2 polarization states of <u>Alfvén modes</u> are polarized transverse to the unperturbed magnetic field; the group velocity is along the mean magnetic field with $\vec{v}_{gr} = \frac{\partial\omega}{\partial\vec{k}} = \frac{\vec{B}_0}{\sqrt{\mu_0\rho}} \equiv \vec{v}_A$, $v_A \sim 90\,\frac{km}{s}\left(\frac{B}{\mu G}\right)\left(\frac{n}{10^{-2}\,cm^{-3}}\right)^{-1/2}$ (3.153)

•) 2 polarization states of <u>fast magnetosonic modes</u>; equivalent to sound waves in a high-$\beta$ cluster plasma with

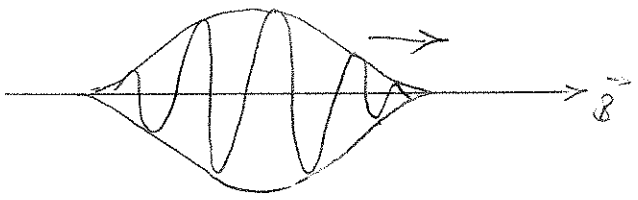
$$\beta = \frac{p_{th}}{p_B} = \frac{n k\bar{T}}{\frac{B^2}{8\pi}} = \frac{2n\cdot k\bar{T}\,\gamma}{\frac{\vec{B}^2}{\mu_0}\cdot \mu_{mp}\, n\,\gamma} = \frac{2}{\gamma}\,\frac{c_s^2}{v_A^2} \gtrsim 100 \qquad (3.154)$$

fast magnetosonic waves do not interact with Alfvén modes

•) 2 polarization states of <u>slow magnetosonic modes</u>

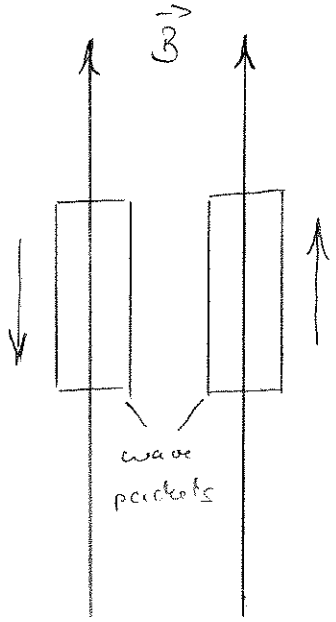
•) <u>the entropy mode</u>: zero-frequency wave with fluctuations in density alone (all other perturbations being zero)

- we now consider Alfvén wave packets (spatially localized) that transport energy and momentum along $\vec{B}$-fields



- MHD wave turbulence deals with the physics of interacting wave packets; Alfvénic turbulence is incompressible:

$$\frac{\delta v_A}{v_A} = \frac{\delta B}{B} \equiv \frac{b_\lambda}{B} \qquad (3.155)$$

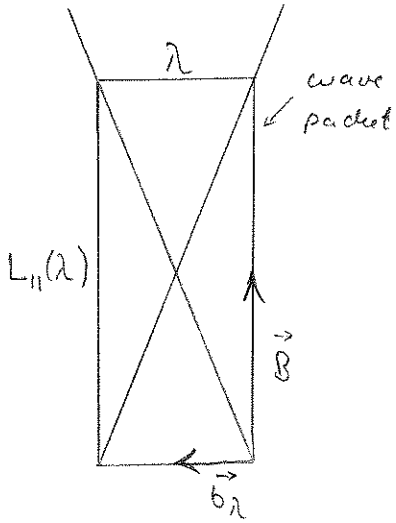



- what happens when the two wave packets are interacting?

  => the downgoing packet causes field line wandering such that the upward going packet is broken apart after a distance $L_\parallel(\lambda)$

  => critical balance condition of Alfvénic turbulence:

$$L_\parallel = \frac{\lambda B}{b_\lambda} \qquad (3.156)$$




geometrical interpretation of critical balance cond.n

- perpendicular to mean field, we have Kolmogorov turbulence; the energy flux of fluctuating field at scale $\lambda$ is constant, $\dot{\varepsilon}_\lambda \simeq \frac{b_\lambda^2}{t_\lambda} = \text{const.}$

$$\rightsquigarrow t_\lambda = \frac{L_\parallel}{v_A} = \frac{\lambda B}{v_A b_\lambda} \propto b_\lambda^2 \text{, and } B \propto v_A = \text{const.} \qquad (3.157)$$

=> we obtain the scaling of Alfvénic turbulence:

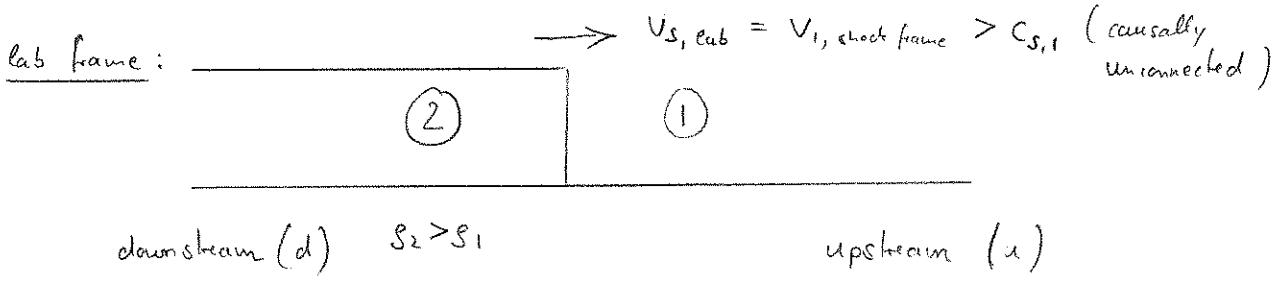
$$b_\lambda \propto \lambda^{1/3} \quad \text{or} \quad L_\parallel \propto \lambda^{2/3} L_{MHD}^{1/3} \qquad (3.158)$$

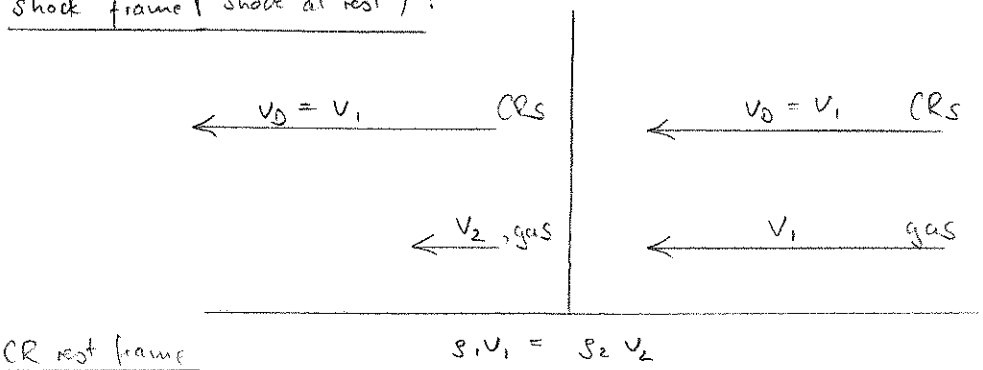
=> the smaller the scale $\lambda$, the more anisotropic is the turbulent scaling, and the more elongated are the eddies with an axis ratio $\left(\frac{L_\parallel}{\lambda} \propto \lambda^{-1/3}\right)$; the long axis of the spaghetti-like eddies is aligned with the local mean field $\langle B \rangle$

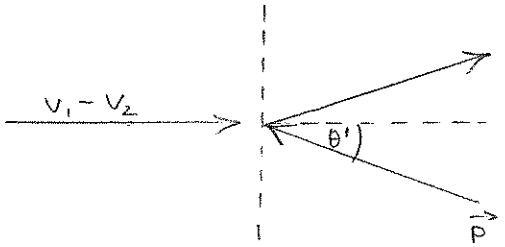
# 3.3.2 Cosmic Rays

## 3.3.2.1 Diffusive Shock Acceleration (First-order Fermi Acceleration)

- at a collisionless shock, electrons and protons (ions) can be accelerated to highly relativistic energies; consider the geometry:




here, the postshock medium acts as an approaching mirror; scattering off this mirror (i.e. macroscopic magnetic irregularities or MHD waves) energizes the particle; we can work out a condition for a particle to scatter downstream and to reach upstream again, where it was initially advected with the flow; to this end, we transform from the lab to the post-shock frame and back to the lab frame; accretion shocks are non-relativistic (NR) shocks with $v_s < c$, while shocks in AGN jets are relativistic

- the Lorentz boost is taken along $c\vec{\beta} = \vec{v}_1 - \vec{v}_2$ ; we define the parallel component of the particle momentum in the lab frame $p_\parallel = \cos\theta\, p = \mu p$ ; in the post-shock frame, the cosmic ray energy and momentum are

$$\begin{pmatrix} E' \\ p' \end{pmatrix} = \begin{pmatrix} \gamma & \pm\beta\gamma \\ \pm\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} E \\ \mu p \end{pmatrix} = \begin{pmatrix} \gamma E \pm \beta\gamma\mu p \\ \pm\beta\gamma E + \gamma\mu p \end{pmatrix} \overset{\text{NR limit}}{\approx} \begin{pmatrix} E + \beta\mu p \\ \beta E + \mu p \end{pmatrix} \tag{3.159}$$

in the last step, we assumed a non-relativistic shock and adopted the "+" sign for a Lorentz transformation in the direction of the cosmic ray's parallel momentum ; after colliding with the "magnetic mirror" $p'$ is reversed to $-p'$ and $E'$ remains unchanged ; transferring back to the lab frame, we find

$$\begin{pmatrix} E'' \\ p'' \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} E' \\ -p' \end{pmatrix} \overset{\text{NR limit}}{\approx} \begin{pmatrix} 1 & -\beta \\ -\beta & 1 \end{pmatrix} \begin{pmatrix} E' \\ -p' \end{pmatrix} = \begin{pmatrix} E + \beta\mu p - \beta(\beta E - \mu p) \\ -\beta(E + \beta\mu p) - \beta E - \mu p \end{pmatrix}$$

$$\overset{\text{NR limit}}{\approx} \begin{pmatrix} E + 2\beta\mu p \\ -2\beta E - \mu p \end{pmatrix} \equiv \begin{pmatrix} E + \delta E \\ p + \delta p \end{pmatrix} \tag{3.160}$$

- for a relativistically moving cosmic ray, we have a condition to be advected downstream, scatter there and reach the upstream again :

$$\delta E = \delta p c = 2(v_1 - v_2)\mu p \quad ; \quad \mu = \cos\theta > 0$$

hence the relative energy increase during an "udu" half cycle is

$$\frac{\delta p}{p} = \frac{2(v_1 - v_2)\mu}{c} > 0 \qquad (\mu > 0) \tag{3.161}$$

- energy is also increased during a "dud" half cycle :

$$\frac{\delta p}{p} = -\frac{2(v_1 - v_2)\mu}{c} > 0 \qquad (\mu < 0) \tag{3.162}$$

$\Rightarrow$ there is a net energy increase of a particle that finishes a full cycle "udud"

- to obtain the energy increase of a population of cosmic rays, we have to integrate this expression (3.161) for individual particles that are distributed according to a distribution function $f(\vec{x},\vec{p})$ over phase space; we conveniently employ spherical coordinates in momentum space $(p, \mu=\cos\theta, \varphi_p)$ and cylindrical coordinates in configuration space $(r, \varphi, z)$ where the first 2 coordinates span an (arbitrary) cross section $\sigma$ of the shock front and the z coordinate is transformed into a time coordinate via $dz = \vec{n}_s \cdot \vec{v}\, dt$ ($\vec{n}_s$ denotes the shock normal); hence, we have the differential volume $d^3x = d\sigma\, \vec{n}_s \cdot \vec{v}\, dt = d\sigma\, v\mu\, dt$; for a half cycle "out", we get

$$\frac{\delta p}{p} = 2\,\frac{v_1 - v_2}{c}\,\frac{\int \mu\, f(\vec{x},\vec{p})\,\frac{d^3x}{2}\,d^3p}{\int f(\vec{x},\vec{p})\,d^3x\,d^3p} = 2\,\frac{v_1-v_2}{c}\,\frac{\int_0^1 \mu\, f\,\sigma\, v\mu\, dt\, 2\pi p^2 dp\, d\mu}{2\int_0^1 f\,\sigma\, v\mu\, dt\, 2\pi p^2 dp\, d\mu}$$

$$= \frac{2(v_1-v_2)}{c}\,\frac{\frac{\mu^3}{3}\Big|_0^1}{\mu^2\Big|_0^1} = \frac{2}{3}\,\frac{v_1-v_2}{c} \qquad (3.162)$$

$\Rightarrow$ the relative energy gain averaged over all particles is then (for a full cycle)

$$\mathcal{E} \equiv \frac{\delta E}{E} = \frac{\delta p}{p} = \frac{4}{3}\,\frac{v_1 - v_2}{c} \qquad (3.163)$$

- the escape probability is the ratio of particle flux carried by downstream flow at $v_2$ over the particle flux that crosses the shock front at speed c (assuming again relativistic particles)

$$P_{esc} = \frac{f_2\, 4\pi\, p^2\, dp\, \sigma\, v_2\, dt}{\int_0^1 f_0\, 2\pi\, p^2 dp\, \sigma c\mu\, dt\, d\mu} = \frac{4v_2}{c}\,\frac{f_2}{2f_0\,\frac{\mu^2}{2}\Big|_0^1} \simeq \frac{4v_2}{c} \qquad (3.164)$$

assuming that the cosmic ray distribution function is rapidly homogenized behind the shock, i.e. $f_2 = f_0$

- after $n$ cycles, the particle has the energy

$$E = E_0 (1+\varepsilon)^n \quad \Longleftrightarrow \quad n = \frac{\log \frac{E}{E_0}}{\log(1+\varepsilon)} \tag{3.165}$$

at each acceleration cycle, there is the escape probability $P_e$; after $n$ cycles, the particle has the probability $(1-P_e)^n$ still to participate in the process; hence, the number of particles with energy larger than $E$ is given by

$$N(>E) \propto \sum_{m=n}^{\infty} (1-P_e)^m = (1-P_e)^n \sum_{m=0}^{\infty} (1-P_e)^m = (1-P_e)^n \frac{1}{1-(1-P_e)} = \frac{(1-P_e)^n}{P_e} \tag{3.166}$$

(↑ geometrical series)

- using $a^{\ln b} = \left(e^{\ln a}\right)^{\ln b} = b^{\ln a}$, we can rewrite this $\left[a \equiv 1-P_e\,;\; b \equiv \frac{E}{E_0}\right]$

$$N(>E) \propto \frac{1}{P_e} \left(\frac{E}{E_0}\right)^{-\tilde{\alpha}}, \quad \tilde{\alpha} = \frac{\log \frac{1}{1-P_e}}{\log(1+\varepsilon)} \tag{3.167}$$

- using (3.163) and (3.164) and Taylor expanding $\log(1+x) = x - \frac{x^2}{2} + \mathcal{O}(x^3)$, we get

$$\tilde{\alpha} = \frac{-\log\left(1-\frac{4v_2}{c}\right)}{\log\left(1+\frac{4}{3}\frac{v_1-v_2}{c}\right)} \approx \frac{4\frac{v_2}{c}}{\frac{4}{3}\frac{v_1-v_2}{c}} = \frac{3\,v_2}{v_1-v_2} \tag{3.168}$$

$$= \frac{3}{\frac{v_1}{v_2}-1} = \frac{3}{r-1}$$

where we introduced the density compression factor $r = \frac{\rho_2}{\rho_1} = \frac{v_1}{v_2}$

- the cumulative particle spectrum due to diffusive shock acceleration is

$$N(>E) = \int N(E)\,dE \propto \frac{c}{4v_2} \left(\frac{E}{E_0}\right)^{-\frac{3}{r-1}} \tag{3.169}$$

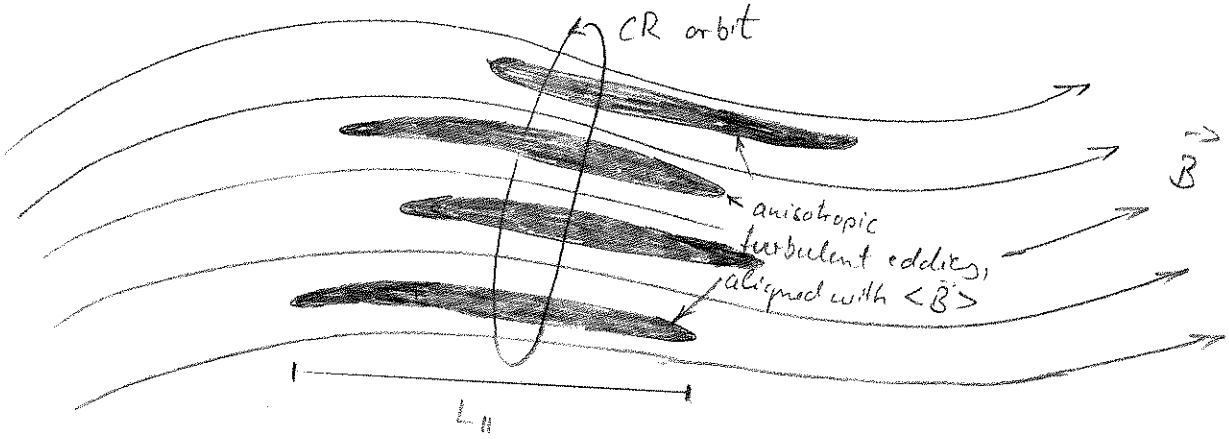
- the differential spectrum $N(E)$ has a spectral index $\alpha$ that connects to $\tilde{\alpha}$ via

$$\tilde{\alpha} = 1-\alpha = -\frac{3}{r-1}$$

$$\alpha = 1 + \frac{3}{r-1} = \frac{r+2}{r-1} = 2 \quad \text{for a strong shock with } r=4 \tag{3.170}$$

## 3.3.2.2 Second-order Fermi Acceleration

- In addition to the very efficient 1st order Fermi acceleration mechanism, there exists also an acceleration mechanism second order in $\beta = \frac{V}{c}$, i.e. much less efficient for non-relativistic flows as it scales as $\frac{V^2}{c^2}$; in Sect. 3.3.1.3, we encountered Alfvénic MHD turbulence which turns out to be negligible for particle acceleration; consider a schematic picture of turbulent eddies in the framework of Alfvénic turbulence (Goldreich & Shridhar '95):




(1) gyroradius of a CR encloses many eddies that are not aligned:

$$L_\perp \ll L_\parallel \sim r_L = \frac{p_\perp c}{Z e B} \tag{3.171}$$

=> this causes a random walk, broadens the gyro resonance and reduces the scattering efficiency

(2) equivalently this argument can be made in k-space where parallel modes decay faster; from eq. (3.158) we get

$$k_\perp \propto L_{MHD}^{1/2} k_\parallel^{3/2}, \quad dk_\perp \propto \frac{3}{2} L_{MHD} k_\parallel^{1/2} dk_\parallel \tag{3.172}$$

perpendicular to $\vec{B}$, we have ordinary Kolmogorov turbulence (3.110) with

$$E(k_\perp) dk_\perp \propto k_\perp^{-5/3} dk_\perp \tag{3.173}$$

but the energy spectrum parallel to $\vec{B}$ is much steeper than Kolmogorov:

$$E(k_\parallel) dk_\parallel \propto k_\parallel^{-5/2} \cdot k_\parallel^{1/2} dk_\parallel \sim k_\parallel^{-2} dk_\parallel \tag{3.174}$$

=> hence there is much less energy available on the resonant scale for scattering and CRs are much weaker coupled to Alfvénic turbulence!

- in addition to Alfénic modes there are compressible fast modes that dominate the CR scattering in spite of damping:

① <u>gyro-resonance</u>:

$$\omega - k_{\parallel} v_{\parallel} = n - \Omega = n \frac{Ze\beta}{\gamma mc} \quad , \; n = \pm 1, \pm 2, \ldots \tag{3.175}$$

which states that the Doppler-shifted MHD frequency is a multiple of the particle's gyro frequency $\Omega$; hence

$$k_{\parallel, \text{resonant}} \sim \frac{\Omega}{v_{\parallel}} = \frac{1}{r_L} \tag{3.176}$$

② <u>non-resonant interactions</u> with "transit time damping":

$$\omega = k_{\parallel} v_{\parallel} \tag{3.177}$$

=> the electron or proton is trapped by a mirror force, surfs the wave and gains energy (head-on collisions are more frequent than tail-on's)

=> stochastic acceleration

- <u>Comparing 1st and 2nd order Fermi acceleration</u>:

1st order: - very efficient at shocks, comparably well understood
- but shocks, in particular strong ones are rare, especially in the cluster volume.

2nd order: - not very efficient
- subsonic turbulence should be ubiquitous in the ICM, which is permanently shaken by minor and (less frequently) major mergers
- difficult plasma physics → no first principle theory suggested

## 3.3.2.3 Cosmic Ray Transport

- consider a spatial random coordinate $x(t)$ of a cosmic ray (CR) particle diffusing in a fluid of bulk velocity $v$; for simplicity, we restrict ourselves to the one-dimensional case; during a time interval, which is much shorter compared to the diffusion time, the particle's position varies by $\Delta x = v\Delta t + \delta x$, where the first contribution is due to the bulk motion of the scattering medium and the second term is due to the random walk diffusion with vanishing mean and the variance $\langle \delta x^2 \rangle = 2D(x,p)\Delta t$ ($D(x,p)$ denotes the diffusion coefficient)

- the distribution of CRs is governed by a competition between injection, escape, energy gain (acceleration), and energy loss (catastrophic and continuous) processes; the transport equation which describes the balance of these processes is a Fokker-Planck type equation that includes the description of fluid motions, radiative losses, and phase space diffusion; it can be obtained by considering the collisionless Boltzmann equation and working out the magneto-hydrodynamic forces acting on a CR particle including the Lorentz force as well as pitch angle scattering on hydro-magnetic waves (details can be found in my *High Energy Astrophysics* lecture; here, I only sketch the picture)

- the transport equation governs the evolution of the isotropic part $f(x,p)$ of the CR distribution function in phase space, assuming weak anisotropy of the CR momentum distribution function:

$$\underbrace{\frac{\partial}{\partial t}f + \frac{\partial}{\partial x}v_t(x,p)f}_{1.)} = \underbrace{-\frac{1}{p^2}\frac{\partial}{\partial p}\left(p^2 A(x,p)f\right)}_{2.)} + \underbrace{\frac{1}{p^2}\frac{\partial}{\partial p}\left(p^2\Gamma(x,p)\frac{\partial}{\partial p}f\right)}_{3.)} + \underbrace{\frac{\partial}{\partial x}\left(D(x,p)\frac{\partial}{\partial x}f\right)}_{4.)} + \underbrace{s(x,p)}_{5.)} \qquad (3.178)$$

the distribution is normalized such that the number density of CRs $n_{\rm CR} = \int f 4\pi p^2 {\rm d}p$; after briefly introducing each term, we discuss them in detail below; the term labeled by 1.) describes the Lagrangian derivative of $f$ as a result of temporal changes and those due to advection with a velocity $v_t = u + v_s$ that is composed of two components: one from the advection with the mean fluid velocity $u$ and one due to the streaming motion of CRs along the magnetic field with velocity $v_s$, 2.) the 'friction' term $A$ describes not only various kinds of (continuous) energy losses but also the energy gain by first order processes in $\beta \equiv u/c$ (adopting relativistic particles), 3.) this term describes diffusion in momentum space, which results in an increasing momentum envelope with time, i.e., energy gain through the second order Fermi process, 4.) this term describes spatial diffusion, and the last term 5.) accounts for sources such as freshly injected CR particles at shocks whose origin can be understood by means of plasma physical calculations and (catastrophic) CR loss processes; we sketch the physical meaning of these processes in the following:

  - **(continuous) synchrotron and inverse Compton losses:** a relativistic charged particle of a Lorentz factor $\gamma = (1-\beta^2)^{-1/2}$ experiences Compton scattering with either real or virtual photons (which represent the magnetic field in the case of synchrotron radiation); this causes the particle to emit photons in the forward direction into a narrow cone of half-angle $\gamma^{-1}$ with respect to its momentum leading to an energy loss which can effectively be described by a friction force in opposite direction to its momentum:

$$A_{\rm rad} \equiv \left.\frac{\langle \Delta p \rangle}{\Delta t}\right|_{\rm rad} = -\frac{4}{3}\sigma_{\rm T}\left(\frac{m_{\rm e}}{m}\right)^2 (\varepsilon_B + \varepsilon_{\rm ph})\gamma^2, \qquad (3.179)$$

    where $\sigma_{\rm T}$ denotes the Thompson cross section, $\varepsilon_B = B^2/(8\pi)$ and $\varepsilon_{\rm ph}$ are the energy densities of the magnetic field (responsible for synchrotron losses) and the low energy photon field (causing the Compton effect in the Thompson regime); the radiative losses of baryons are suppressed by $(m_{\rm e}/m)^2$ such that they can be neglected unless they are ultra-high energetic CRs with energies $\gtrsim 10^{18}$ eV

- **first order Fermi process:** the contribution of the first order Fermi process can be described by a non-inertial entrainment due to the deceleration of the scattering medium: a compressed flow ($\nabla \cdot \boldsymbol{u} < 0$) produces first order acceleration of charged particles. In this situation, the inertial force is $F_j = -p_i(\partial u_j/\partial x_i)$ that gives rise to an accelerating power

$$P_{\rm acc} = -\langle u_j p_i \rangle \frac{\partial u_j}{\partial x_i} = -\frac{pu}{3}\nabla \cdot \boldsymbol{u} \quad \rightarrow \quad A_{\rm acc} = -\frac{p}{3}\frac{\partial u}{\partial x} \qquad (3.180)$$

- **second order Fermi process:** charged particles gyrate around, and travel slowly along magnetic field lines; occasionally, they get scattered on magnetic irregularities and plasma waves (Alfvén and magnetosonic waves); this scattering process can be described by a random walk of the particle's pitch angle with the magnetic field direction, $\theta$, yielding the characteristic variance $\langle \delta\mu^2 \rangle \propto \nu_{\rm s}\Delta t$ where $\nu_{\rm s}$ denotes the average scattering frequency and $\mu = \cos\theta$; because the particle scatters off moving targets, the particle systematically gains energy through random variations of the CR momentum $\delta p = \pm\beta_{\rm A} p \delta\mu$ where $\beta_{\rm A} = v_{\rm A}/c$ is the dimensionless Alfvén velocity in the case of scattering Alfvén waves; the second order Fermi process is thus described by a diffusion process in momentum space with the momentum diffusion coefficient

$$\Gamma \equiv \frac{\langle \delta p^2 \rangle}{2\Delta t} \sim \beta_{\rm A}^2 \nu_{\rm s} p^2 \qquad (3.181)$$

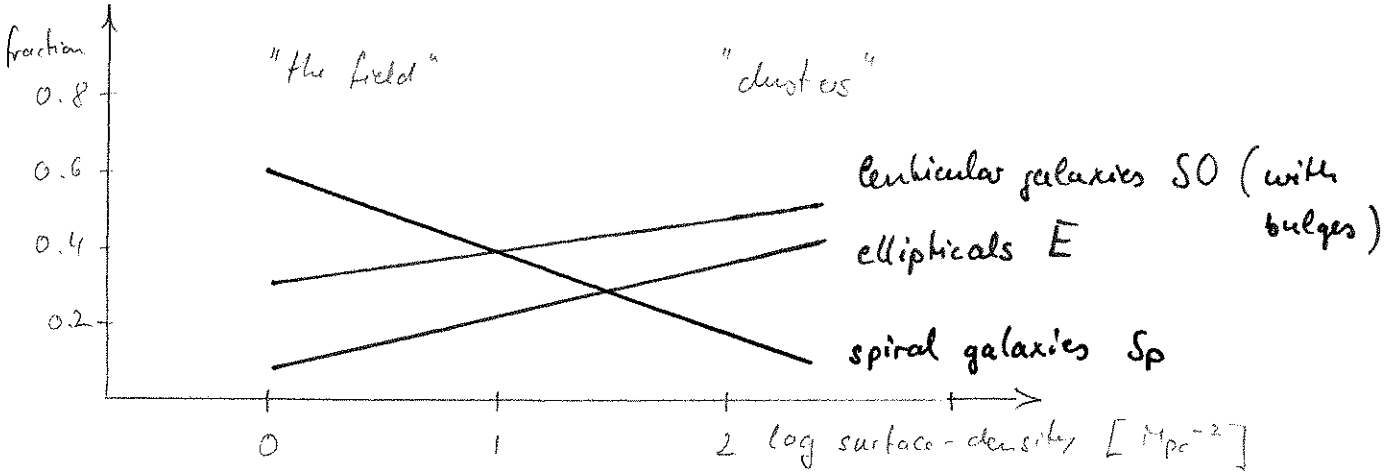
- **spatial diffusive losses and gains:** depending on the gradient of $n_{\rm CR}$ in a particular region, that region will either see a net CR flux outward ($\nabla \cdot n_{\rm CR} > 0$) or inward ($\nabla \cdot n_{\rm CR} < 0$); hence, for a peaked CR distribution in the cluster center, CR diffusion acts to smooth out that peak over a length scale $L$ on the diffusion time scale $\tau = L^2/D$; CRs experience momentum dependent diffusion in a turbulent magnetic field with a Kolmogorov-type spectrum on small scales; this process leads to a loss time scale which is proportional to $p^{-1/3}$; in an equilibrium situation, this results in a steepening of the observed spectrum within the cluster core by $p^{-1/3}$ relative to the injected spectrum
- **Coulomb and ionization losses** are strongest for protons or heavier nuclei, but also relevant for electrons; the ionization process limits the lower energy of the proton spectrum to approximately 50 MeV after traversing a path length through most of the interstellar medium; energetic CRs experience energy losses even within an ionized medium through Coulomb interactions; Coulomb losses efficiently remove the low-energetic part of the injected CR spectrum on a short timescale ($\lesssim$ Gyr) and redistribute these particles and their energy into the thermal pool
- **catastrophic losses:** another loss process is the inelastic reaction of CR nuclei with ions of the intracluster medium; the CR protons interact hadronically with the ambient thermal gas and produce mainly neutral and charged pions, provided their momentum exceeds the kinematic threshold of 0.78 GeV for the reaction; the neutral pions successively decay into $\gamma$-rays while the charged pions decay into secondary neutrinos and electrons that emit radio synchrotron radiation and Compton-upscatter CMB photons to the $\gamma$-ray regime

# 4 Cluster Physics Informed by Different Observables

## 4.1 Optical : Galaxy Properties and Virial Theorem

### 4.1.1 Observational Facts

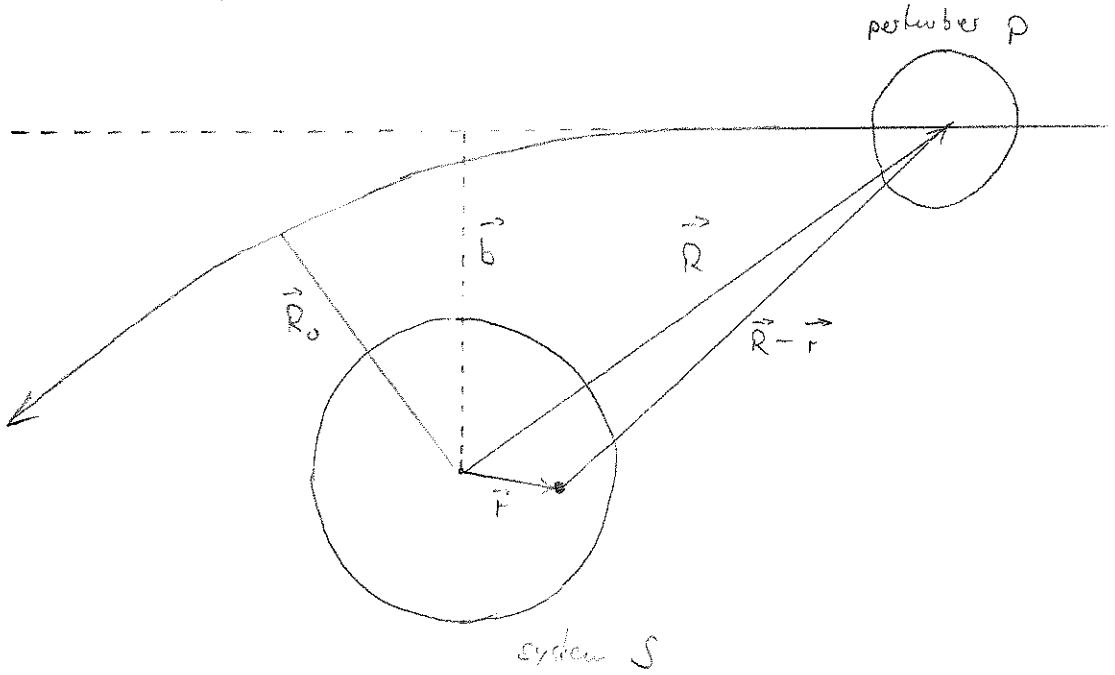
- "morphology - density relation": denser environments host larger fractions of galaxies that are morphologically classified as early types (ellipticals); in addition, galaxies in denser environments are on average redder, are more massive, more concentrated, less gas-rich, and have lower specific star-formation rates




- Butcher-Oemler effect: galaxy populations in clusters at intermediate redshifts ($0.3 \lesssim z \lesssim 0.5$) have a dramatically increased fraction of blue galaxies compared to present-day clusters; in addition, morphological studies have shown that this Butcher-Oemler effect is associated with an increasing spiral fraction as well as a star formation fraction with increasing redshift, and many of these spirals show disturbed morphologies

- although far from unambiguous, these strong environmental and redshift dependencies are often interpreted as indications that galaxies undergo transformations, e.g. late type (spirals) → early type (ellipticals) or star forming → passive; once they became part of a denser environment; clusters as the largest collapsed structures in the Universe are also the environments with the highest number densities of galaxies; hence, galaxy interactions are frequent, making clusters the ideal environments to look for possible transformation processes
- we will now discuss various processes that operate in clusters and may be responsible for transforming star forming disks into passive spheroids:
  (i) tidal interactions with other cluster galaxies or with the cluster potential,
  (ii) dynamical friction, which causes the galaxy to slowly migrate to the cluster center, and
  (iii) interactions with the hot, X-ray emitting ICM

## 4.1.2 Galaxy Interactions and Transformations



- consider a body $S$ which has an encounter with a perturber $P$ with impact parameter $b$; let $q$ be a particle (e.g. a star) in $S$ at a distance $\vec{r}$ from the center; since the gravitational force due to $P$ is not uniform over the body $S$, the particle $q$ experiences a tidal force per unit mass

$$\vec{F}_{tide}(\vec{r}) = -\vec{\nabla}\phi_P(|\vec{R}-\vec{r}|) + \vec{\nabla}\phi_P(\vec{R}) \qquad (4.1)$$

with $\phi_P$ the gravitational potential of $P$; as a result of the encounter, the particle $q$ gains energy at a rate per unit mass

$$\frac{dE_q}{dt} = \vec{v}\cdot\vec{F}_{tide}(\vec{r}) \qquad (4.2)$$

with $\vec{v}$ the velocity of $q$ with respect to the center of $S$;

- similar to the way the Moon induces oceanic tides on the Earth, the gravitational interaction between $S$ and $P$ enhances the gravitational multipole moments of both bodies, which may cause a backreaction on their orbit; let $\tau_{tide}$ be the time for the tide to rise and $\tau_{enc} \approx \frac{R_{max}}{\Delta v_{SP}}$ is the time of the encounter with $\Delta v_{SP}$ the relative velocity between $S$ and $P$ and $R_{max} = max(R_0, R_S, R_P)$ where $R_S$ and $R_P$ are characteristic radii of $S$ and $P$, and $R_0$ is the minimum distance of the encounter

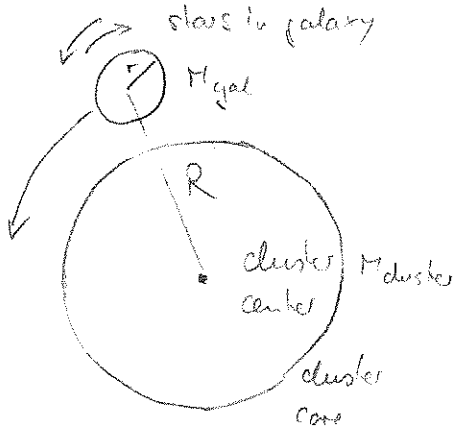
- if $\tau_{enc} \gg \tau_{tide}$, then the internal structure of deformable bodies has enough time to adiabatically adjust to the perturbation in form of tides (due to the change of relative position and orientation of $S$ and $P$); hence the effects of the encounter during approach and departure cancel each other (the deformations are adiabatic) $\rightarrow$ no net transfer of energy

- if $\tau_{enc} \lesssim \tau_{tide}$, the response of the bodies lags behind the instantaneous magnitude and direction of the force $\rightarrow$ backreaction on orbit; the net effect in this case is a transfer of orbital energy to internal energy of the two bodies $\rightarrow$ increase in mutual binding energy

# 4.1.2.1 Tidal Stripping and Shocks

## Tidal radius – Roche problem

- first, consider slow encounters and work out the "tidal radius" outside of which material can get stripped; for simplification, imagine a galaxy on a circular orbit in a cluster (which can be generalized in a straightforward manner); what is the fate of the stars inside this galaxy?




- the differential tidal force per unit mass between star and galaxy center is

$$F_{tid} = r \frac{d}{dR}\left(\frac{G M_{cluster}(R)}{R^2}\right) \qquad (4.3)$$

the restoring force from the galaxy is

$$F_{gal} = -\frac{G M_{gal}(r)}{r^2} \qquad (4.4)$$

- the tidal radius is defined by $F_{tid} = F_{gal}$

$$\frac{M_{gal}(r_{tid})}{r_{tid}^3} = -\frac{d}{dR}\left(M_{cluster}(R)\,R^{-2}\right) = \left(2 - \frac{d\ln M_{cluster}}{d\ln R}\right)\frac{M_{cluster}}{R^3}$$

$$\rightarrow \quad r_{tid} = \left[\frac{M_{gal}}{M_{cluster}(R)}\,\frac{1}{\left(2 - \frac{d\ln M_{gal}}{d\ln R}\right)}\right]^{1/3} R \qquad (4.5)$$

$$\text{or} \quad \bar{S}_{gal}(r_{tid}) = \left(2 - \frac{d\ln M_{cluster}}{d\ln R}\right)\bar{S}_{cluster}(R) \qquad (4.6)$$

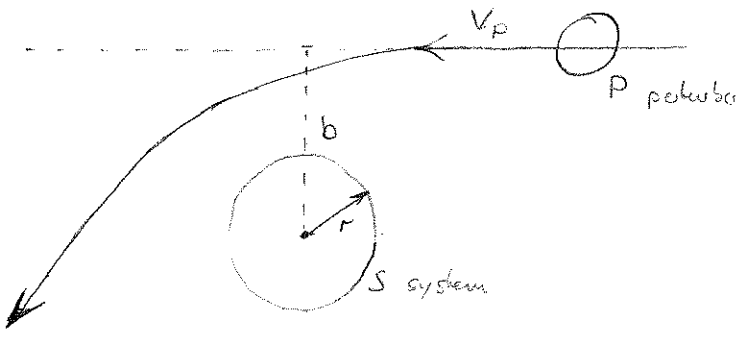
this is the "Roche criterion" for slow encounters so that the internal (stellar) distribution can adjust adiabatically to the perturbation!

- outside $r_{tid}$ , stars are only loosely bound so that they can be "stripped off" by tidal forces exerted by the cluster core on the stars in the galaxy ; this implies a resonance condition when the period of the stellar orbit in the galaxy at $r_t$ matches the period of the galaxy orbit inside the cluster

$$P_* (r_t) \approx P_{gal} (R) \qquad (4.7)$$

- complications :
  (i) orbits are not circular ~ apply formula at pericenter
  (ii) effects depend on phase and inclination of stellar orbits within the galaxy

## Tidal shocks ( galaxy harassment )




- now : treat rapid encounters in the impulse approximation and assume time of passing $\tau_{enc} \ll$ internal dynamical time $\tau_{tid}$
- since the tidal force per unit mass exerted by P on a star in S is $F_{tid} = \frac{dv_*}{dt}$ , we obtain an impulsive change in star velocity relative to the center of mass in S over the interaction time $\frac{b}{V_p}$ )

$$\Delta V_* \sim r \frac{G M_p}{b^3} \cdot \frac{b}{V_p} \sim \frac{G M_p r}{b^2 V_p} \qquad (4.8)$$

- the change in the total energy per unit mass of a star in S is given by

$$\Delta E = \frac{1}{2} (\vec{v}_* + \Delta \vec{v}_*)^2 - \frac{1}{2} \vec{v}_*^2 = \vec{v}_* \cdot \Delta \vec{v}_* + \frac{1}{2} |\Delta \vec{v}_*|^2 \qquad (4.9)$$

- we want to compute $\Delta E_S$ by integrating $\Delta \bar{E}$ over the entire system S; because of symmetry, the integral of the first term on the right-hand-side of equ. (4.9) is identical to zero, so that

$$\Delta E_S = \frac{1}{2} \int |\Delta \vec{v}_*(\vec{r})|^2 \, \rho(\vec{r}) \, d^3 r \qquad (4.10)$$

- inserting (4.8), we get a change in the energy per unit mass (Spitzer 1958)

$$\Delta E_S = \frac{4}{3} \frac{G^2 M_p^2 r^2}{b^4 v_p^2} \qquad (4.11)$$

where the coefficient (4/3) averages over all stars in a shell at radius r

- What is the relative change in energy due to an impulsive encounter? using the approximation $E_S(r) \approx -\frac{G M_S(r)}{r}$, we obtain

$$\frac{\Delta E_S}{E_S} = -\frac{4}{3} \frac{G M_p^2 r^3}{M_S(r) v_p^2 b^4} = -\frac{4}{3} \frac{\bar{\rho}_p(b)}{\bar{\rho}_S(r)} \frac{v_{parabolic}^2}{v_p^2} \qquad (4.12)$$

where we used

$$\bar{\rho}_p(b) = \frac{M_p(b)}{\frac{4}{3}\pi b^3}, \quad \bar{\rho}_S(r) = \frac{M_S(r)}{\frac{4}{3}\pi r^3}, \quad v_{parabolic}^2 = \frac{G M_p}{b}$$

- notes:

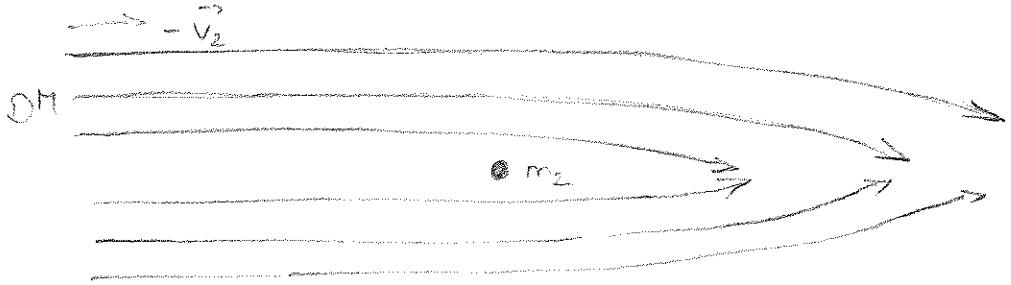
(i) $v_p \sim v_{parabolic} \Rightarrow \frac{\Delta E}{E} \sim 1$ for $\rho_p \sim \rho_S$, i.e. we recover the Roche criterion

(ii) formalism only applicable to fast encounters, i.e. down to $r_{crit}$, below which adiabatic rules applies

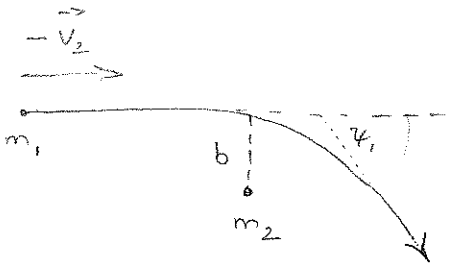
(iii) cumulative effect of many encounters can strip stars to $r \ll r_{Roche}$

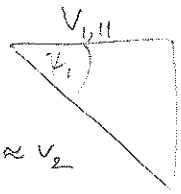
## 4.1.2.2 Dynamical Friction

- when a heavy object of mass $m_2$ (e.g. a galaxy) moves through a large collisionless system that is constituted of particles of mass $m_1 \ll m_2$ (e.g. DM particles), it experiences a drag force that is called dynamical friction, which transfers energy and momentum from the galaxy to the DM particles; the reason for this is that a system evolves toward thermodynamic equilibrium through energy exchange by means of two-body encounters; if particles have different masses, thermodynamic equilibrium implies $m_1 \langle v_1^2 \rangle = m_2 \langle v_2^2 \rangle$; since $m_2 \gg m_1$ and DM particles at the same radius have similar orbital velocities, the galaxy has usually larger kinetic energy than the DM particles it encounters $\rightarrow$ galaxy looses net energy and momentum
- an alternative way to understand dynamical friction is to move in the rest system of the galaxy; DM particles are deflected through the gravity of the galaxy to form an overdensity of DM in the wake (a so-called "trailing enhancement"); the gravitational pull of this wake on the galaxy slows it down




- in real systems with a gravitational potential, dynamical friction (DF) produces mass segregation and not equipartition; e.g. a massive galaxy on circular orbit within a cluster experiences DF $\rightsquigarrow$ tighter bound orbit $\rightsquigarrow$ galaxy moves faster!




$$v_{1,\perp} = v_2 \cos\psi$$

$$\Delta v_{1,\perp} \approx v_2 - v_{1,\perp} = v_2 (1 - \cos\psi_1)$$

- in a single encounter, the deflection due to gravity is (in the rest system of the galaxy with $m_2$)

$$\psi_1 = \frac{\Delta v_\perp}{v_2} \approx \underbrace{\frac{G m_2}{b^2}}_{|\vec{F}_\perp|} \cdot \underbrace{\frac{2b}{v_2}}_{t} \cdot \frac{1}{v_2} = \frac{2 G m_2}{b v_2^2} \qquad (4.13)$$

- momentum balance yields (with $\frac{m_1}{m_2} \ll 1$):

$$\Delta v_{2,\perp} = -\frac{m_1}{m_2} \Delta v_{1,\perp} = -\frac{m_1}{m_2} v_2 (1 - \cos\psi_1) \approx -\frac{2 G^2 m_1 m_2}{b^2 v_2^3} \qquad (4.14)$$

($1 - \cos\psi_1 \approx 1 - \frac{\psi_1^2}{2}$)

- Since the rate of this encounter is given by

$$\Gamma = \int n v \, d\sigma$$

we obtain for the rate of change for the galaxy velocity

$$\frac{d\vec{v}_2}{dt} = \int_{b_{min}}^{b_{max}} \Delta v_{2,\perp} \, n_1 \, v_2 \, 2\pi b \, db \, \frac{\vec{v}_2}{|\vec{v}_2|}$$

$$= -4\pi G^2 m_2 \, \rho_1(v_2) \ln\Lambda \, \frac{\vec{v}_2}{|\vec{v}_2|^3} \quad \text{with} \quad \Lambda = \frac{b_{max}}{b_{min}}$$

→ this is the Chandrasekhar formula, where $\rho_1(v_2)$ is the mass density of light particles moving slower than $v_2$ (and assuming $f(v_1)$ to be isotropic).

notes:

(i) $\dot{v}_{gal}$ is independent of the DM mass (whether it is composed of axions or black holes), $\dot{v}_2$ depends only on $\Omega$

(ii) $\dot{v}_{gal} \propto m_{galaxy}$: heavier galaxies experience a larger drag which moves them faster to the bottom of the cluster potential

(iii) assume Boltzmann distribution for DM and $v_{gal} \ll \sigma_{DM}$:

$$\left.\begin{aligned} f &\propto e^{-\frac{v^2}{2\sigma^2}} \to 1, \quad v_{gal} \ll \sigma_{DM} \\ &\sim g(<v_{gal}) \propto \frac{v_{gal}^3}{\sigma_{DM}^3} \quad \text{for } v_{gal} \ll \sigma_{DM} \end{aligned}\right\} \quad \dot{v}_{gal} \propto v_{gal} \text{ for } v_{gal} \ll \sigma_{DM} \quad (4.13)$$

(Stokes friction)

(iv) in the opposite limit $v_{gal} \gg \sigma_{DM}$:

$$\dot{v}_{gal} \propto v_{gal}^{-2} \quad (4.14)$$

Orbit decay through dynamical friction

- consider a simple SIS model with circular galaxy orbits:

$$f(E) \propto e^{-\frac{E}{\sigma_1^2}} \; ; \; \rho(r) = \frac{\sigma_1^2}{2\pi G r^2} \; , \quad M(r) = \frac{2\sigma_1^2}{G} r \quad (4.15)$$

→ a galaxy on a circular orbit has

$$\frac{v_2^2}{r} = \frac{G M(r)}{r^2} \quad \Rightarrow \quad v_2 = \sqrt{2}\,\sigma_1 \quad \text{for all radii } r \quad (4.16)$$

→ its binding energy per unit mass is

$$\frac{\phi(r)}{M(r)} = \frac{G m_{shell}}{r} = 4\pi G \int_{r_0}^{r} \frac{\tilde{r}^2 \rho \, d\tilde{r}}{\tilde{r}} = 2\sigma_1^2 \int_{r_0}^{r} \frac{d\tilde{r}}{\tilde{r}} = 2\sigma_1 \ln \frac{r}{r_0} \quad (4.17)$$

→ this evolves according to dynamical friction

$$\dot{E} = \vec{v}_2 \cdot \vec{F} = \vec{v}_2 \cdot \frac{d\vec{v}_2}{dt} = -4\pi G^2 m_2 \ln\Lambda \, \frac{0.43\sigma_1^2}{2\pi G r^2} \, \frac{1}{\sqrt{2}\sigma_1}$$

$$= -0.61 \, \frac{G m_2 \sigma_1 \ln\Lambda}{r^2} \quad (4.18)$$

→ but we have also

$$\dot{E} = \frac{2\sigma_1^2}{r} \frac{dr}{dt} \quad (4.19)$$

→ equating these yields

$$\frac{2\sigma_1^2}{r}\frac{dr}{dt} = -0.61\,\frac{G m_2 \sigma_1 \ln\Lambda}{r^2}$$

$$\rightsquigarrow \frac{dr^2}{dt} = 2r\frac{dr}{dt} = -0.61\,\frac{G m_2 \ln\Lambda}{\sigma_1} = -\frac{G m_2 \ln\Lambda}{1.64\,\sigma_1 r_1^2}\,r_1^2 \equiv -\frac{r_1^2}{T}$$

→ integration gives

$$r^2 = r_1^2\left(1 - \frac{t}{T}\right) \quad \text{where} \quad T = 1.64\,\frac{r_1^2\sigma_1}{G m_2 \ln\Lambda} = \frac{0.185}{\ln\Lambda}\,\frac{M_1(r_1)}{m_2}\,t_{\text{orbit, initial}} \qquad (4.20)$$

$$\text{and} \quad t_{\text{orbit, initial}} = 2\pi\frac{r_1}{V_c} = \sqrt{2}\,\pi\,\frac{r_1}{\sigma_1}$$

⇒ in this simple model, $T$ is the orbit decay time due to dynamical friction; as expected $T \propto m_2^{-1}$ and large galaxies migrate faster to the center; to get a numerical value, we set the minimum impact parameter, $b_{max}$, equal to the size of the host system, $r_{200}$, and we use that $\bar{v}_{rel} \simeq V_c = \sqrt{\frac{G M_{200}}{r_{200}}}$; the Coulomb logarithm becomes

$$\ln\Lambda \simeq \ln\frac{M_{200}}{m_{gal}} \qquad (m_2 = m_{gal}) \qquad (4.21)$$

→ we use

$$\frac{r_{200}}{V_c} = \left(\frac{r_{200}^3}{G M_{200}}\right)^{1/2} = \frac{1}{10\,H(z)} \qquad (4.22)$$

with $\dfrac{M_{200}}{\frac{4}{3}\pi r_{200}^3} = 200\,\dfrac{3H^2(z)}{8\pi G} \quad \rightsquigarrow \quad \dfrac{G M_{200}}{r_{200}^3} = 100\,H^2(z)$

$\rightsquigarrow$ hence we can rewrite (4.20)

$$T \simeq \frac{1.17}{\ln\left(\frac{M_{200}}{m_{gal}}\right)}\,\frac{M_{200}}{m_{gal}}\,\frac{1}{10\,H(z)} \qquad (4.23)$$

→ thus, the dynamical friction decay time from the edge of a halo to the center is longer than the age of the universe for $M_{200}/m_{gal} \gtrsim 15$; only the most massive subhalos and satellite galaxies are expected to be substantially disrupted by mass

- if a galaxy sinks to the center due to dynamical friction (e.g. because of its large mass or highly eccentric orbit), it will merge with the central galaxy that resides there, a process called "galactic cannibalism"

### 4.1.2.3 Ram Pressure Stripping

- consider a disk of radius $r_d$ moving through an ICM of density $\rho_{ICM}$ with velocity $\vec{v}$; for simplicity, we consider a face-on moving disk (i.e. $\vec{v}$ is parallel to disk normal); the amount of ICM material swept by the disk is

$$\dot{M}_{ICM} = \sigma_d \rho_{ICM} v = \pi r_d^2 \rho_{ICM} v \qquad (4.24)$$

- if we assume, that the interstellar medium (ISM) stopps the wind, then the momentum transferred to the disk per unit time is

$$\dot{p} = v \dot{M}_{ICM} = \pi r_d^2 \rho_{ICM} v^2 = \pi r_d^2 P_{ram} \qquad (4.25)$$

with the ram pressure $P_{ram} = \rho_{ICM} v^2$. (4.26)

- if this pressure exceeds the force per unit area that binds the ISM to the disk, then the gas gets stripped from the disk: to estimate the binding force, we assume that the mean surface density of interstellar gas is $\Sigma_{ISM}$ and that the mean mass density (usually dominated by stars) is $\Sigma_*$; the gravitational field of the disk (in the disk) is

$$\phi_{grav} \approx 2\pi G \Sigma_* \qquad (4.27)$$

hence, the gravitational force per unit area on the interstellar gas is

$$\frac{F_{grav}}{A} \approx 2\pi G \Sigma_* \Sigma_{ISM}$$

- stripping occurs if $P_{ram} > \frac{F_{grav}}{A^2}$, or

$$\rho_{ICM} > \frac{2\pi G \Sigma_* \Sigma_{ISM}}{v^2} \qquad (4.28)$$

- consider a Milky Way - like disk with

$M_* \sim 5 \cdot 10^{10} M_\odot$, $M_{ISM} \sim 5 \cdot 10^9 M_\odot$, $r = 10$ kpc,

which is moving with $v = 1000 \frac{km}{s}$; (4.28) gives $\rho_{ICM} > 4.6 \cdot 10^{-27} \frac{g}{cm^3}$ for stripping to occur, or $n \geq 4.6 \cdot 10^{-3}\ cm^{-3}$ which compares well to ICM densities toward the cluster core;

- in general, $P_{ram}$ will change as a function of time due to the eccentric orbit of a galaxy (causing $v$ and $\rho_{ICM}$ to change); hence (4.28) depends on time and eccentricity; since $\Sigma_*$ and $\Sigma_{ISM}$ decrease to the disk periphery, there will be a critical radius beyond which ram pressure stripping will be effective

- when spirals loose their gas → potential for future star formation is reduced → $P_{ram}$ - stripping is considered important to explain the quenching of star formation in clusters

- if only parts of the outer disk gets stripped, star formation can continue until it runs out of its fuel after a few Gyrs; this is called "strangulation" of a galaxy as supposed for the abrupt quenching due to ram pressure stripping

# 4.1.3 Virial Theorem

## 4.1.3.1 Derivation

- start with Newtonian N-body equations

$$m_i \ddot{\vec{r}}_i = -\vec{\nabla} \sum_{j \neq i} \phi_{ij} \quad , \quad i = 1 \dots N \tag{4.29}$$

take $\phi_{ij}$ to be Newtonian : $\phi_{ij} = -\frac{G m_i m_j}{r_{ij}}$ , $r_{ij} = |\vec{r}_i - \vec{r}_j|$

- derivation :

$$\sum_i \vec{r}_i \cdot \left| \quad m_i \ddot{\vec{r}}_i = -\vec{\nabla} \sum_{j \neq i} \phi_{ij} \right.$$

$$\sum_i m_i \vec{r}_i \cdot \ddot{\vec{r}}_i = \sum_i \sum_{j \neq i} \frac{G m_i m_j \vec{r}_i \cdot (\vec{r}_j - \vec{r}_i)}{|\vec{r}_i - \vec{r}_j|^3} \tag{4.30}$$

→ switching i and j, we have

$$\sum_j m_j \ddot{\vec{r}}_j \cdot \vec{r}_j = \sum_j \sum_{j \neq i} \frac{G m_j m_i \vec{r}_j \cdot (\vec{r}_i - \vec{r}_j)}{|\vec{r}_j - \vec{r}_i|^3} \tag{4.31}$$

→ adding (4.30) and (4.31), we get

$$2 \sum_i m_i \vec{r}_i \cdot \ddot{\vec{r}}_i = \sum_i \sum_{j \neq i} \frac{G m_i m_j}{|\vec{r}_i - \vec{r}_j|^3} \left[ \vec{r}_i \cdot (\vec{r}_j - \vec{r}_i) + \vec{r}_j \cdot (\vec{r}_i - \vec{r}_j) \right] \tag{4.32}$$

$$- (\vec{r}_i - \vec{r}_j)^2$$

→ dividing by 2 yields

$$\sum_i m_i \vec{r}_i \cdot \ddot{\vec{r}}_i = -\frac{1}{2} \sum_{\substack{i,j \\ j \neq i}} \frac{G m_i m_j}{|\vec{r}_i - \vec{r}_j|} = -\sum_{\substack{i,j \\ i > j}} \frac{G m_i m_j}{|\vec{r}_i - \vec{r}_j|} \tag{4.33}$$

→ hence we have

$$\frac{1}{2} \frac{d^2}{dt^2} \left( \sum_i m_i |\vec{r}_i|^2 \right) - \sum_i m_i |\dot{\vec{r}}_i|^2 = -\sum_{\substack{i,j \\ i>j}} \frac{G m_i m_j}{|\vec{r}_i - \vec{r}_j|} \tag{4.34}$$

$$\rightsquigarrow \quad \ddot{I} = 2T + W = 0 \quad \text{in equilibrium} \tag{4.35}$$

$I = \sum_i m_i \vec{r}_i^{\,2}$ moment of inertia

$T = \sum_i \frac{m_i}{2} |\dot{\vec{r}}_i|^2$ total kinetic energy

$W = \sum_{\substack{i,j \\ i>j}} \frac{G m_i m_j}{|\vec{r}_i - \vec{r}_j|}$ total gravitational potential energy

$M = \sum_i m_i$ is total mass

$V^2 = \frac{2T}{M}$ is 3-dimensional velocity dispersion

$R_g \equiv -\frac{GM^2}{2W}$ is the gravitational radius

$E = T + W$ is the total energy

$t_{dyn}$ or $t_{cross} \equiv \frac{2R_g}{V}$ is the dynamical or crossing time

$\rightsquigarrow$ a system in equilibrium is characterized by $2T + W = 0$ and determined by E and M only !

## 4.1.3.2 Weighting a Cluster with Galaxies

- how can we weight a galaxy cluster through observations of galaxy spectra ? we emphasize the underlying assumptions that (if not fully realized) will bias the mass

- assumption (1) : system is close to equilibrium - not just forming

$$\rightsquigarrow \sum_i m_i v_i^2 - \sum_{\substack{i,j \\ i>j}} \frac{G m_i m_j}{r_{ij}} = 0$$

- usually, we only have radial velocities $v_{los}$ + projected separations between galaxy i and j, $r_{\perp,ij} = D\,\theta_{ij}$, where D is the angular diameter distance

assumption (2): System is viewed from random direction (test particles representative) → not biased by selection

$v_{los}$:


$$v_{los,i} = v_i \cos\theta_i$$

$$\sum_i m_i v_i^2 \cos^2\theta_i \equiv \sum_i m_i v_{los,i}^2 \tag{4.36}$$

taking the averages on both sides, we obtain

$$\left\langle \sum_i m_i v_i^2 \right\rangle \left\langle \cos^2\theta_i \right\rangle = \left\langle \sum_i m_i v_{los,i}^2 \right\rangle \tag{4.37}$$

making use of the fact that $m_i$ and $v_i$ are not correlated with $\theta_i$

$$\langle v_{los,i}^2 \rangle = \langle v_i^2 \rangle \langle \cos^2\theta_i \rangle =$$

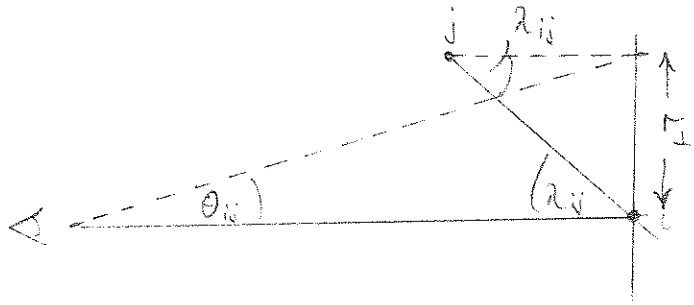
$$= \langle v_i^2 \rangle \frac{\int_0^{2\pi}\int_0^{\pi/2} \cos^2\theta \sin\theta \, d\theta \, d\varphi}{\int_0^{2\pi}\int_0^{\pi/2} \sin\theta \, d\theta \, d\varphi} = \langle v_i^2 \rangle \left. \frac{x^3}{3} \right|_0^1 = \left\langle \frac{v_i^2}{3} \right\rangle \tag{4.38}$$

↑ sub: $x = \cos\theta$

⇒ unbiased velocity estimator:

$$v_i^2 = 3 v_{los,i}^2 \quad \text{with} \quad 2T = 3 \left\langle \sum_i m_i v_{los,i}^2 \right\rangle \tag{4.39}$$

$r_{\perp,ij}$:




$$r_{\perp,ij} = r_{ij} \sin\lambda_{ij} = D \tan\theta_{ij} \approx D\theta_{ij}$$

$$\sum_{\substack{i,j \\ i>j}} \frac{G m_i m_j}{r_{ij} \sin\lambda_{ij}} = \sum_{\substack{i,j \\ i>j}} \frac{G m_i m_j}{r_{\perp,ij}} \tag{4.40}$$

- taking again the averages with respect to the random sample yields

$$-\omega \left\langle \frac{1}{\sin \lambda_{ij}} \right\rangle = \left\langle \sum_{\substack{i,j \\ i>j}} \frac{G\, m_i m_j}{r_{\perp,ij}} \right\rangle \qquad (4.111)$$

$$\left\langle \frac{1}{\sin \lambda_{ij}} \right\rangle = \frac{\int_0^{2\pi} \int_0^{\pi/2} \frac{1}{\sin \lambda_{ij}} \sin \lambda_{ij}\, d\lambda_{ij}\, d\varphi_{ij}}{\int_0^{2\pi} \int_0^{\pi/2} \sin \lambda_{ij}\, d\lambda_{ij}\, d\varphi_{ij}} = \frac{\pi}{2} \qquad (4.112)$$

$$\Rightarrow -\omega = \frac{2}{\pi} \sum_{\substack{i,j \\ i>j}} G \frac{m_i m_j}{r_{\perp,ij}} \simeq \frac{2}{\pi} \sum_{\substack{i,j \\ i>j}} G \frac{m_i m_j}{D\theta_{ij}} \qquad (4.113)$$

$\Rightarrow$ <u>unbiased estimator for the angular separation:</u>

$$\frac{1}{r_{ij}} = \left\langle \frac{2}{\pi} \frac{1}{D\theta_{ij}} \right\rangle \qquad (4.114)$$

- what about fluctuations around this mean, i.e. $\left\langle \frac{1}{\theta_{ij}^2} \right\rangle$ ?

$$\left\langle \frac{1}{\sin^2 \lambda_{ij}} \right\rangle = \frac{\int_0^{\pi/2} \frac{1}{\sin^2 \lambda_{ij}} \sin \lambda_{ij}\, d\lambda_{ij}}{\int_0^{\pi/2} \sin \lambda_{ij}\, d\lambda_{ij}} = \log \tan \frac{\lambda_{ij}}{2} \Big|_0^{\pi/2} \to \infty \qquad (4.115)$$

$\Rightarrow$ Virial Theorem mass estimates are subject to large fluctuations!

<u>assumption (3)</u>: all particles are counted

$\rightarrow$ all mass is in observed objects with $m_i = \left(\frac{M}{L}\right) L_i$

(i) no objects excluded by selection

(ii) all mass in galaxies

for a system of equal mass particles we obtain from (4.34), (4.39), and (4.44)

$$m = \frac{3 \sum_i v_{\mathrm{pec},i}^2}{\frac{2}{\pi} \frac{G}{D} \sum_{\substack{i,j \\ i>j}} \theta_{ij}^{-1}} \qquad (4.116)$$

hence,

$$\frac{M}{L} = \frac{3\pi D \sum_i L_i\, v_{\mathrm{pec},i}^2}{2G \sum_{\substack{i,j \\ i>j}} L_i L_j\, \theta_{ij}^{-1}} \qquad (4.117)$$

problems : - $L$ - weighting is irrelevant
- $\frac{M}{L}$ may not be constant
- not all mass may be attached to galaxies

assumption (4) : positions and velocities of objects are representative :

$$M = \frac{3\pi D \sum_i v_{los,i}^2}{2G \sum_{i \neq j} \theta_{ij}^{-1}} \qquad (4.48)$$

problem : → bad statistical behaviour of $\theta_{ij}^{-1}$

solution : obtain angular size by fitting a model

- weight a cluster : $R_{cl} \sim 1\,Mpc \quad v_{cl} \sim 1000 \frac{km}{s}$

$N_{gal} \sim 200$ with $R_{gal} \sim 3\,kpc$, $v_{gal} \sim 150 \frac{km}{s}$

thus $\frac{\text{Mass in visible galaxies}}{\text{total mass}} = \frac{N_{gal} M_{gal}}{M_{cl}} \sim \frac{N_{gal} R_{gal} v_{gal}^2}{R_{cl} v_{cl}^2}$

$$\sim N_{gal} \left( \frac{R_{gal}}{R_{cl}} \right) \left( \frac{v_{gal}}{v_{cl}} \right)^2 \sim 0.0135 \qquad (4.49)$$

→ to have $\sum_i m_{gal} \sim M_{tot}$, you need $R_{gal} \sim 0.2\, R_{cl}$

Either : (i) galaxies must more extended than light
(ii) most mass not attached to galaxies
(iii) clusters very far from equilibrium
(iv) gravity not Newtonian

## 4.2 Gravitational Lensing

The theory presented in this section is based on two main assumptions: (i) the Newtonian limit of a slowly varying gravitational field is taken from Einstein's field equations, namely $|\Phi| \ll c^2$ and $|v_{\rm lens}| \ll c$, in order to characterize the properties of lenses, and (ii) the lensing objects are considered to be thin, i.e. the deflecting mass is isolated and concentrated within a region $L$ much smaller than the distances between source and deflector and deflector and observer, $L \ll cH_0^{-1}$. This approximation holds remarkably well in the astrophysical cases of galaxies or clusters of galaxies.

### 4.2.1 Deflection Angle $\alpha$

Linearizing the gravitational field equations and taking non–relativistic sources results in the "post–Minkowskian" metric to first order, neglecting the gravitational vector potential,

$$\mathrm{d}s^2 = \left(1 + \frac{2\Phi}{c^2}\right) c^2 \mathrm{d}t^2 - \left(1 - \frac{2\Phi}{c^2}\right) \mathrm{d}\boldsymbol{r}^2 , \tag{4.50}$$

where $\Phi$ represents the Newtonian potential and $\mathrm{d}\boldsymbol{r}$ characterizes the spatial part of the Minkowski metric. Using the fact that light propagates on null geodesics, namely $\mathrm{d}s^2 = 0$, yields an effective velocity of light $c'$ in the presence of a weak gravitational field,

$$c' = \frac{|\mathrm{d}\boldsymbol{r}|}{\mathrm{d}t} \simeq c\left(1 + \frac{2\Phi}{c^2}\right) \equiv \frac{c}{n} , \tag{4.51}$$

$$n = \left(1 - \frac{2\Phi}{c^2}\right) \geq 1 , \tag{4.52}$$

which defines an *effective index of refraction* $n$ of the gravitational field in analogy to geometrical optics in dense media. Note that the gravitational potential $\Phi$ is by definition negative as it represents an attractive gravitational force. Applying Fermat's principle leads to an equation for the spatial light paths by using the Euler–Lagrange equations for carrying out the variation

$$\delta \int_A^B n \, \mathrm{d}l = \delta \int_A^B n(\boldsymbol{r}) \sqrt{|\dot{\boldsymbol{r}}|^2} \, \mathrm{d}\lambda \stackrel{!}{=} 0 \tag{4.53}$$

$$\text{or} \qquad \ddot{\boldsymbol{r}} = -\frac{2}{c^2} \nabla_\perp \Phi(\boldsymbol{r}), \tag{4.54}$$

where the different curves are parametrized by the affine curve parameter $\lambda$, the dot denotes a derivative with respect to $\lambda$ and $\nabla_\perp \Phi(\boldsymbol{r})$ is the gradient of the potential perpendicular to the perturbed light ray. The total deflection is therefore the integral along the light path of the differential displacements,

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{r}) = -\int \nabla_\perp n(\boldsymbol{r}) \, \mathrm{d}l = \frac{2}{c^2} \int \nabla_\perp \Phi(\boldsymbol{r}) \, \mathrm{d}l . \tag{4.55}$$

Because in nearly all cases of astrophysical interest the deflection angle is small, $\hat{\boldsymbol{\alpha}} \ll 1$, one usually applies the "Born approximation" and evaluates the integral along the unperturbed ray, i.e. along a straight line. Since the non–relativistic matter is characterized by its density perturbations only, the gravitational potential which gives rise to light deflections (4.55) neither depends on the actual nature of matter nor its composition or physical state. Therefore gravitational light deflection probes the total matter density of gravitationally interacting particles irrespective of baryonic and dark matter.

### 4.2.2 Lens Equation

The lensing equation relates the intrinsic angular source position of an astrophysical object to its observable image position on the sky which was possibly changed in the presence of gravitational light deflection along the line of sight. In order to derive this equation in the thin screen approximation, it is useful first to consider lensing by a point mass. The Newtonian potential as well as its perpendicular gradient can be written as

$$\Phi(\boldsymbol{\xi}, z) = -\frac{GM}{\sqrt{\xi^2 + z^2}} \tag{4.56}$$

$$\text{and} \qquad \nabla_\perp \Phi(\boldsymbol{\xi}, z) = \frac{GM\boldsymbol{\xi}}{(\xi^2 + z^2)^{3/2}}\,, \tag{4.57}$$

where the tree dimensional vector $\boldsymbol{r}$ is decomposed into the $z$–coordinate along the unperturbed ray and the two dimensional impact parameter $\boldsymbol{\xi}$ orthogonal to the unperturbed ray pointing towards the point mass. Equation (4.55) leads to the deflection angle

$$\hat{\alpha}(\boldsymbol{\xi}) = \frac{2}{c^2}\int_{-\infty}^{\infty} \frac{GM\boldsymbol{\xi}}{(\xi^2 + z^2)^{3/2}}\,\mathrm{d}z = \frac{4GM}{c^2\xi}\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|} = \frac{2R_S}{\xi}\frac{\boldsymbol{\xi}}{|\boldsymbol{\xi}|}\,, \tag{4.58}$$

with $R_S$ being the Schwarzschild radius of the point mass. The Born approximation in this context makes sure that the integral is evaluated along the straight coordinate line $z$.

If we now consider extended objects acting as lenses, but still located within a small region compared to the total distance between lens and observer, the mass distribution of the lensing object can be projected along the line of sight. The smooth three–dimensional distribution can then be replaced by a mass layer perpendicular to the line of sight, which is called *lens plane*. The surface mass density on the lens plane is given by

$$\Sigma(\boldsymbol{\xi}) = \int \rho(\boldsymbol{\xi}, z)\,\mathrm{d}z\,, \tag{4.59}$$

and the deflection angle at position $\boldsymbol{\xi}$ is the overall deflection effect due to a superposition of "point–mass" elements in the plane because of linearity of the system:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\xi}) = \frac{4G}{c^2}\int \frac{\Sigma(\boldsymbol{\xi}')(\boldsymbol{\xi} - \boldsymbol{\xi}')}{|\boldsymbol{\xi} - \boldsymbol{\xi}'|^2}\,\mathrm{d}^2\xi'\,. \tag{4.60}$$

This equation holds in the lens plane with the impact parameter measured in physical units. Assuming the small angle approximation, the *lens equation* relates the position of the source to the observable image position on the sky. The geometry of a typical gravitational lens system is shown in figure (1).

The true position of the source with respect to some arbitrarily chosen optical axis is denoted by $\boldsymbol{\beta}$ and the angular image position on the sky as viewed by an observer is given by $\boldsymbol{\theta}$. All distances along the line of sight are angular diameter distances, where $D_{ls}$ denotes the distance between lens and source, $D_l$ the distance between lens and observer and $D_s$ the distance between source and observer. Using the relation $\boldsymbol{\xi} \simeq D_l\boldsymbol{\theta}$ and introducing the reduced deflection angle,

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{D_{ls}}{D_s}\hat{\boldsymbol{\alpha}}(\boldsymbol{\theta}), \tag{4.61}$$

equation (4.60) can be written as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{4G}{c^2}\frac{D_l D_{ls}}{D_s}\int \frac{\Sigma(\boldsymbol{\theta}')(\boldsymbol{\theta} - \boldsymbol{\theta}')}{|\boldsymbol{\theta} - \boldsymbol{\theta}'|^2}\,\mathrm{d}^2\theta'\,. \tag{4.62}$$

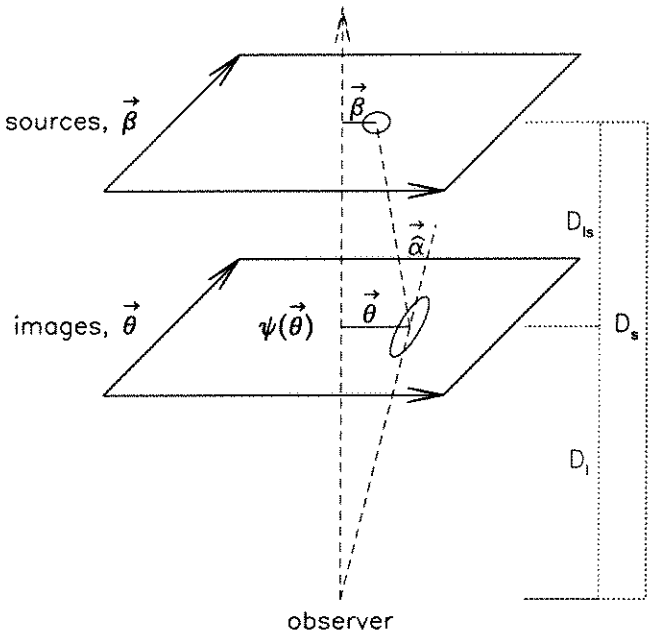



Figure 1: Illustration of a typical gravitational lens system. The angles are exaggerated for visualization purposes.

The critical surface mass density $\Sigma_{\rm cr}$ and the convergence $\kappa$ are defined by

$$\Sigma_{\rm cr} \equiv \left(\frac{4\pi G}{c^2}\frac{D_l D_{ls}}{D_s}\right)^{-1} \qquad \text{and} \qquad \kappa \equiv \frac{\Sigma}{\Sigma_{\rm cr}}\,. \tag{4.63}$$

It is important to note that the distance combination appearing in equation (4.62), $\frac{D_l D_{ls}}{D_s}$, acts as a lensing efficiency function. It approaches zero at both the source and the observer and has a maximum in between. Using definitions (4.63), the deflection angle as a function of the image position $\boldsymbol{\theta}$ reduces to

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{1}{\pi}\int \kappa(\boldsymbol{\theta}')\frac{\boldsymbol{\theta}-\boldsymbol{\theta}'}{|\boldsymbol{\theta}-\boldsymbol{\theta}'|^2}\,\mathrm{d}^2\theta'\,. \tag{4.64}$$

This equation shows that only the ration of $\Sigma$ and $\Sigma_{\rm cr}$ is measurable, or in other words, using gravitational lensing on its own, one is not able to determine both the mass of a lensing object and the involved distances independently. From figure (1) we can read off $\boldsymbol{\theta}D_s - \hat{\boldsymbol{\alpha}}D_{ls} = \boldsymbol{\beta}D_s$, assuming the small angle approximation and using the theorem on intersecting lines. Using the expression for the reduced deflection angle, this establishes the *lens equation* in its simplest form

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta})\,. \tag{4.65}$$

In general, this equation is nonlinear and can thus yield multiple images on the sky for a single source position $\boldsymbol{\beta}$. Moreover, the shape and the size of the images will differ from the original source because light bundles are deflected differentially.

### 4.2.3 Circular Symmetric Lenses – Einstein Radius

Consider a circularly symmetric lens with an arbitrary mass profile. Due to the high degree of symmetry, we can place the coordinate origin at the center of symmetry and reduce light deflection to a one-dimensional problem. The deflection angle always points toward the center of symmetry with a modulus

$$\hat{\alpha}(\xi) = \frac{4GM(\xi)}{c^2\xi}, \tag{4.66}$$

where $\xi = D_l\theta$ is the distance from the lens center and $M(\xi)$ is the enclosed mass within $\xi$,

$$M(\xi) = 2\pi \int_0^\xi \Sigma(\xi')\xi' \mathrm{d}\xi'. \tag{4.67}$$

Combining equations (4.61) and (4.66) enables us to rewrite the lensing equation (4.68), yielding

$$\beta(\theta) = \theta - \frac{D_{ls}}{D_l D_s} \frac{4GM(\theta)}{c^2\theta} \tag{4.68}$$

Owing to the rotational symmetry of the lens, a source, which lies exactly on the optical axis ($\beta = 0$) is imaged as a ring if the lens is supercritical ($\Sigma > \Sigma_{\rm cr}$). Setting $\beta = 0$ in equation (4.68) gives the radius of the ring, the so-called *Einstein radius*,

$$\theta_{\rm E} = \sqrt{\frac{4GM(\theta)}{c^2} \frac{D_{ls}}{D_l D_s}}. \tag{4.69}$$

It is not only a property of the lens, but also depends on the distance efficiency function. It provides the natural angular scale to describe the lensing geometry for the following reasons: (i) in the case of multiple imaging, the angular separation of images is of order $2\theta_{\rm E}$, (ii) sources that lie closer than approximately $\theta_{\rm E}$ to the optical axis experience strong lensing yielding to strong magnification and sheared images whereas sources at much larger distances are only very little magnified, and (iii) in many lens models the Einstein radius roughly represents the boundary of source positions that are either multiply-imaged if they lie inside $\theta_{\rm E}$ or singly-imaged. Comparing equations (4.63) and (4.69) reveals that the surface mass density inside $\theta_{\rm E}$ is exactly the critical surface density $\Sigma_{\rm cr}$. For a point mass, we can obtain the Einstein radius

$$\theta_{\rm E} = \sqrt{\frac{4GM}{c^2} \frac{D_{ls}}{D_l D_s}} \approx 30'' \left(\frac{M}{10^{14} M_\odot}\right)^{1/2} \left(\frac{D}{\rm Gpc}\right)^{-1/2}, \tag{4.70}$$

where we defined the lensing efficiency distance $D = D_l D_s / D_{ls}$ and inserted typical values for clusters to highlight the relevant angular scales for giant (tangential) arcs in clusters. In the case of clusters, detailed mass modeling is necessary since only a fraction of the cluster mass resides within the Einstein radius.

For a point mass lens, we can use the Einstein radius of equation (4.69) to rewrite the lens equation in the form

$$\beta = \theta - \frac{\theta_{\rm E}^2}{\theta}. \tag{4.71}$$

This equation has two solutions

$$\theta_\pm = \frac{1}{2}\left(\beta \pm \sqrt{\beta^2 + 4\theta_{\rm E}^2}\right). \tag{4.72}$$

Any source inside $\theta_{\rm E}$ is imaged twice by a point mass lens. The two images are on either side of the source with one image inside the Einstein ring and the other one outside. As the source moves away from the center of the lens (i.e., with increasing $\beta$), one of the images approaches the lens and becomes very faint, while the other image approaches the true position of the source and tends toward a magnification of order unity.

### 4.2.4 The Lensing Potential $\Psi$ and Local Lens Properties

It is convenient to define the lensing potential $\Psi(\boldsymbol{\theta})$ which is the scaled and projected Newtonian potential of the lens,

$$\Psi(\boldsymbol{\theta}) = \frac{D_{ls}}{D_l D_s}\frac{2}{c^2}\int \Phi(D_l\boldsymbol{\theta}, z)\mathrm{d}z\,. \tag{4.73}$$

The lensing potential has the nice property that its gradient with respect to $\boldsymbol{\theta}$ is the deflection angle

$$\nabla_\theta \Psi(\boldsymbol{\theta}) = \frac{D_{ls}}{D_s}\frac{2}{c^2}\int \nabla_\perp \Phi(\boldsymbol{\xi}, z)\mathrm{d}z = \boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{4.74}$$

where the perpendicular gradient is now acting on the physical impact parameter having used the small angle approximation $\boldsymbol{\xi} \simeq D_l\boldsymbol{\theta}$. Assuming further that the changes of the Newtonian potential along the line of sight average out, which is true for instance, as long as the lensing object is only slowly varying and does not undergo a rapid collapse. More precisely, the time–scale on which light travels across the lensing object, has to be much smaller than the collapse time–scale of the light deflecting object. Then the two–dimensional Laplacian can be replaced by its three–dimensional analogue,

$$\Delta^{(2)}\Phi(\boldsymbol{r}) = \sum_{i=1}^{2}\frac{\partial^2\Phi(\boldsymbol{r})}{\partial\xi_i^2} \simeq \sum_{i=1}^{3}\frac{\partial^2\Phi(\boldsymbol{r})}{\partial r_i^2} = \Delta^{(3)}\Phi(\boldsymbol{r})\,. \tag{4.75}$$

Therefore, the Laplacian of the lensing potential acting on its angular coordinate $\boldsymbol{\theta}$ equals twice the surface mass density scaled with its critical value, i.e. the convergence $\kappa$,

$$\Delta_\theta^{(2)}\Psi(\boldsymbol{\theta}) = \frac{2}{c^2}\frac{D_l D_{ls}}{D_s}\int \Delta^{(3)}\Phi(\boldsymbol{\xi}, z)\mathrm{d}z = 2\frac{4\pi G}{c^2}\frac{D_l D_{ls}}{D_s}\int \rho(\boldsymbol{\xi}, z)\mathrm{d}z = 2\kappa(\boldsymbol{\theta})\,, \tag{4.76}$$

where Poisson's equation has been used in the second step. Since $\Psi$ satisfies the two–dimensional Poisson's equation, its Green's function has to be considered, namely

$$\Delta^{(2)}G(\boldsymbol{\theta}, \boldsymbol{\theta}') = 2\pi\delta_D(\boldsymbol{\theta}, \boldsymbol{\theta}') \quad \Longrightarrow \quad G(\boldsymbol{\theta}, \boldsymbol{\theta}') = \ln|\boldsymbol{\theta} - \boldsymbol{\theta}'|\,. \tag{4.77}$$

Therefore the lensing potential $\Psi(\boldsymbol{\theta})$ is given by the convolution integral of the source function $\kappa(\boldsymbol{\theta})$ and the Green's function in two dimensions,

$$\Psi(\boldsymbol{\theta}) = \frac{1}{\pi}\int \kappa(\boldsymbol{\theta}')\ln|\boldsymbol{\theta} - \boldsymbol{\theta}'|\mathrm{d}^2\theta'\,. \tag{4.78}$$

Liouville's theorem and the conservation of the physical number density of photons during the process of gravitational light bending imply that lensing conserves surface brightness or specific intensity. Assuming that the angular scale on which the lens properties change is much larger than the extent of the source, the lens equation can locally be linearized yielding

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \boldsymbol{\alpha}(\boldsymbol{\theta}) \simeq \boldsymbol{\beta}_0 + \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}}(\boldsymbol{\theta} - \boldsymbol{\theta}_0)\,. \tag{4.79}$$

The local lens properties of the lens mapping are described by its Jacobian matrix $\mathcal{A}$

$$\mathcal{A} \equiv \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}} = \left(\delta_{ij} - \frac{\partial\alpha_i(\boldsymbol{\theta})}{\partial\theta_j}\right) = \left(\delta_{ij} - \frac{\partial^2\Psi(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) \equiv \left(\delta_{ij} - \Psi_{,ij}\left(\boldsymbol{\theta}\right)\right) = \mathcal{M}^{-1}\,, \tag{4.80}$$

where an abbreviation for partial derivatives has been introduced and $\mathcal{A}$ is the inverse of the magnification tensor $\mathcal{M}$. This is justified, because a solid–angle element $\delta\beta^2$ of the source is

mapped onto the solid–angle element $\delta\theta^2$ on the image, and thus the magnification due to the mapping is given by

$$\frac{\delta\theta^2}{\delta\beta^2} = \det \mathcal{M} = \frac{1}{\det \mathcal{A}}\,. \tag{4.81}$$

The trace of the Jacobian $\mathcal{A}$ describes the isotropic magnification of the source,

$$\mathrm{tr}(\mathcal{A}) = (1 - \Psi_{,11}) + (1 - \Psi_{,22}) = 2(1 - \kappa)\,. \tag{4.82}$$

This also intuitively explains the meaning of the convergence $\kappa$, which is a measure for how much the lens focuses light rays isotropically. Subtracting the trace from $\mathcal{A}$ leads to an expression for anisotropic distortion (astigmatism) of the image,

$$\mathcal{A}_{ij} - \frac{1}{2}\delta_{ij}\,\mathrm{tr}(\mathcal{A}) = \delta_{ij} - \Psi_{,ij} - \delta_{ij}(1 - \kappa) = -\Psi_{,ij} + \kappa\delta_{ij} \equiv \Gamma\,, \tag{4.83}$$

where the *shear tensor* $\Gamma$ has been defined in the last step. This distortion is due to the tidal gravitational field. Particularly, it decomposes in

$$\Gamma = \begin{pmatrix} \gamma_1 & \gamma_2 \\ \gamma_2 & -\gamma_1 \end{pmatrix} \tag{4.84}$$

and

$$\gamma_1 = \frac{1}{2}(\Psi_{,11} - \Psi_{,22}) \equiv \gamma(\boldsymbol{\theta})\cos\left(2\phi(\boldsymbol{\theta})\right) \tag{4.85}$$

$$\gamma_2 = \Psi_{,12} = \Psi_{,21} \equiv \gamma(\boldsymbol{\theta})\sin\left(2\phi(\boldsymbol{\theta})\right)\,. \tag{4.86}$$

Here $\gamma = \sqrt{\gamma_1^2 + \gamma_2^2}$ describes the magnitude of the shear and $\phi$ its orientation, whereas the factor 2 shows that $\gamma$ is not a vector, but a $2 \times 2$–tensor.

### 4.2.5 Strong and Weak Cluster Lensing

The name of the game in cluster lensing research consists in reconstructing the mass distribution. Depending on the type of lensing – strong or weak lensing – there are different algorithms used.

**4.2.5.1 Strong Cluster Lensing** Using our simplified mass density model of a singular isothermal sphere (SIS, see Section 2.4.1), we can readily work out the relevant strong lensing properties for this model. Recall that the mass density and rotational velocity in this model was given by

$$\rho(r) = \frac{\sigma_v^2}{2\pi G}\frac{1}{r^2} \quad \text{and} \quad v_{\mathrm{rot}}^2 = \frac{GM(r)}{r} = 2\sigma_v^2 = \text{const.} \tag{4.87}$$

Upon projection the density along the line-of-sight, we obtain the surface mass density

$$\Sigma(\xi) = \int_0^s \rho(\boldsymbol{\xi}, z)\mathrm{d}z = 2\int_\xi^\infty \frac{\rho(r)r\mathrm{d}r}{\sqrt{r^2 - \xi^2}} = \frac{\sigma_v^2}{2G}\frac{1}{\xi}, \tag{4.88}$$

where $\xi$ is the distance from the center of the two-dimensional profile. Using equation (4.66), we obtain the deflection angle

$$\hat{\alpha} = 4\pi\frac{\sigma_v^2}{c^2} \tag{4.89}$$

which is independent of $\xi$ and points to the center of the lens. The Einstein radius of the SIS is given by equation (4.69),

$$\theta_{\mathrm{E}} = 4\pi\frac{\sigma_v^2}{c^2}\frac{D_{ls}}{D_s} = \hat{\alpha}\frac{D_{ls}}{D_s} = \alpha. \tag{4.90}$$

The symmetry of the problem reduces the dimensionality of the problem to become one-dimensional. Multiple images are only obtained if the source lies inside the Einstein ring, i.e., if $\beta < \theta_{\rm E}$. If this condition is satisfied, we obtain the following two solutions,

$$\theta_{\pm} = \beta \pm \theta_{\rm E}. \tag{4.91}$$

The images at $\theta_{\pm}$, the source, and the lens all lie on a straight line. (The third image with zero flux lies at $\theta = 0$ and only acquires a non-zero flux if the singularity of the lens is replaced by a core of finite density).

Rich concentrated clusters can produce giant arcs when a background galaxy is aligned with one of the cluster caustics. Typically, a parametrized lens model (such as the SIS above or a more complicated functional form) is optimized so as to obtain a good fit to the observed image. If there are many constraints from a number of strongly lensed galaxies such as their position and detailed properties of their distortion (magnitude distribution across their arcs), ray tracing through an adaptive grid is possible. This can even constrain the detailed mass distribution within the cluster including their substructure mass distribution.

**4.2.5.2 Weak Cluster Lensing – The Kaiser & Squires Algorithm** Every cluster weakly distorts images of a large number of background galaxies giving rise to so-called arclets – this phenomenon is referred to as weak lensing. With the development of the Kaiser & Squires (1993) algorithm, weak lensing is being used increasingly to derive parameter-free two-dimensional mass maps of galaxy clusters.

Kaiser & Squires Algorithm – topic of mini-symposium.

# 4.3 X-rays : Astrophysics at High Resolution

## 4.3.1 Hydrostatic Equilibrium Masses and Biases

- unless the intracluster gas gets continuously disturbed, we expect it to relax on roughly a sound crossing time of the cluster

$$t_s \equiv \frac{D}{c_s} \approx 6.6 \cdot 10^8 \text{ yr} \left(\frac{T}{10^8 \text{K}}\right)^{-1/2} \left(\frac{D}{1 \text{ Mpc}}\right) \qquad (4.92)$$

since $t_s$ is shorter than the age of a typical cluster, which is a fraction of the Hubble time, the gas should be close to hydrostatic equilibrium; this, however, should not be applicable to clusters that experienced a recent merger or to regions where an AGN has injected energy recently

- in hydrostatic equilibrium, the pressure force balances gravity

$$\vec{\nabla} P = -\rho \vec{\nabla} \phi \quad \rightarrow \quad \frac{1}{\rho} \frac{dP}{dr} = -\frac{GM(r)}{r^2} \quad \text{for spherical symmetry} \qquad (4.93)$$

where $M(r)$ is the total enclosed mass at radius $r$; since this is a single equation for $\rho$ and $P$, we must specify an equation of state to solve this; we assume an ideal gas, $P = \frac{\rho k T}{m}$, where $m$ is the particle mass (e.g., $m = \mu m_p$ for the intracluster plasma)

$$\frac{kT}{m} \frac{d\rho}{dr} + \frac{\rho k}{m} \frac{dT}{dr} = -\frac{GM\rho}{r^2} \qquad (4.94)$$

- considers galaxy motions within the cluster's DM dominated potential as the motion of a gas with temperature

$$\frac{3}{2} kT = \frac{m}{2} \left(3 \sigma_v^2\right) \quad \Rightarrow \quad kT = m \sigma_v^2 \qquad (4.95)$$

where $\sigma_v$ is the velocity dispersion along one coordinate direction

- hence, equation (4.94) becomes

$$M(r) = -\frac{r\sigma_v^2}{G}\left(\frac{d\ln \rho_{gal}}{d\ln r} + \frac{d\ln \sigma_v^2}{d\ln r}\right) \qquad (4.96)$$

where $\rho_{gal}$ is now the (number) density of galaxies

- equivalently, we find for the intracluster plasma with mass density $\rho_{gas}$

$$M(r) = -\frac{rkT}{Gm}\left(\frac{d\ln \rho_{gas}}{d\ln r} + \frac{d\ln T}{d\ln r}\right) \qquad (4.97)$$

- combining both mass estimates yields

$$\sigma_v^2\left(\frac{d\ln \rho_{gal}}{d\ln r} + \frac{d\ln \sigma_v^2}{d\ln r}\right) = \frac{kT}{m}\left(\frac{d\ln \rho_{gas}}{d\ln r} + \frac{d\ln T}{d\ln r}\right) \qquad (4.98)$$

- introducing the ratio of specific energies

$$\beta \equiv \frac{m\sigma_v^2}{kT} \Rightarrow d\ln\beta = d\ln\sigma_v^2 - d\ln T \qquad (4.99)$$

leads to the expression

$$d\ln \rho_{gas} = \beta\left(d\ln \rho_{gal} + d\ln \sigma_v^2\right) - d\ln T \qquad (4.100)$$

- combining (4.99) and (4.100) to eliminate $\sigma_v$ yields

$$d\ln \rho_{gas} = \beta\, d\ln \rho_{gal} + (\beta - 1)\, d\ln T - d\beta \qquad (4.101)$$

or equivalently

$$\rho_{gas} \propto \rho_{gal}^{\beta}\, T^{\beta-1} \quad \text{for } \beta = \text{const.} \qquad (4.102)$$

- since the galaxy distribution follows a King profile

$$\rho_{gal} = \rho_0\left[1 + \left(\frac{r}{r_c}\right)^2\right]^{-\frac{3}{2}} \qquad (4.103)$$

and for an isothermal gas distribution, we obtain a $\beta$-profile

$$\rho_{gas} = \rho_0\left[1 + \left(\frac{r}{r_c}\right)^2\right]^{-\frac{3\beta}{2}} \qquad (4.105)$$

- Since the X-ray emissivity $\varepsilon_X \propto \rho^2$, we obtain

$$\varepsilon_X \propto \left(1 + \frac{r^2}{r_c^2}\right)^{-3\beta} \tag{4.106}$$

which yields upon a line-of-sight integration

$$S_X = \int_0^{\text{cluster}} \varepsilon_X \, dz = S_{X0} \left(1 + \frac{r_\perp^2}{r_c^2}\right)^{-3\beta + \frac{1}{2}} \tag{4.107}$$

this functional form provides excellent fits to the X-ray surface brightness of observed clusters with $\beta \approx \frac{2}{3}$

- to obtain $M(r)$ we calculate

$$\begin{aligned} \frac{d \ln S_{gas}}{d \ln r} &= -\frac{3\beta}{2} r \frac{d}{dr} \ln\left[S_0 \left(1 + \frac{r^2}{r_c^2}\right)\right] \\ &= -3\beta \frac{\frac{r^2}{r_c^2}}{1 + \frac{r^2}{r_c^2}} \end{aligned} \tag{4.108}$$

for an isothermal cluster, we obtain from (4.97)

$$M(r) = \frac{3\beta r kT}{G m} \frac{\frac{r^2}{r_c^2}}{1 + \frac{r^2}{r_c^2}} \xrightarrow{r \gg r_c} \frac{3\beta kT}{G m} \cdot r \tag{4.109}$$

- assuming NFW DM density profiles, gas in hydrostatic equilibrium, spherical symmetry, and isothermal gas yields X-ray surface brightness profiles that are excellently fit with an $\beta$ profile, but the resulting mass profiles are wrong:
  - → many of the simplifying assumptions break down, especially in the cluster outskirts
  - → X-ray observations are mostly sensitive to regions around the core radius where the DM density profile (that imprints its structure onto the gas) scales as $\rho_{DM} \propto r^{-2}$; coincidentally this is the asymptotic scaling of the gas density, $\rho_{gas} \propto r^{-3\beta} \sim r^{-2}$ for $\beta \approx \frac{2}{3}$!