GALAXY FORMATION

Yannick Bahé (MPA)

Outline of lecture:

1.) Motivation: Why are we interested in Galaxies and their Formation?

2.) Formation of haloes (DM & hot gas)

3.) Gas cooling

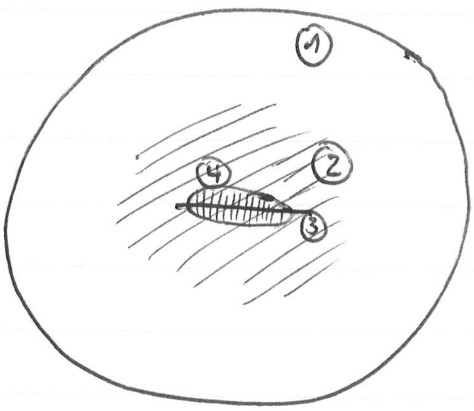
4.) Formation of galactic disks

Text books / Reviews

- Mo, v.d. Bosch & White (2010): "Galaxy Formation & Evolution"
- Binney & Tremaine (2008): "Galactic Dynamics"
- Benson (2010): "Galaxy Formation Theory"

1 MOTIVATION / INTRODUCTION

1.1. What are Galaxies?



↑ Very simple picture of a galaxy

1: Dark Matter halo
2: Extended (hot) gas halo
3: Cool gas disk
4: Stars (disk / spheroid)

more later

→ large range in mass:

$$\sim 10^5 \lesssim \frac{M_*}{M_\odot} \lesssim 10^{13}$$

1.2.: Why do we care about Galaxies?

- They host (almost) all the stellar mass in the Universe
  $\Rightarrow$ visible in optical light

- Building blocks of (observable/optical) Universe on larger scales:
  groups, clusters, superclusters, filaments, ...

- Many galaxies show intricate internal structure – want to understand where this comes from

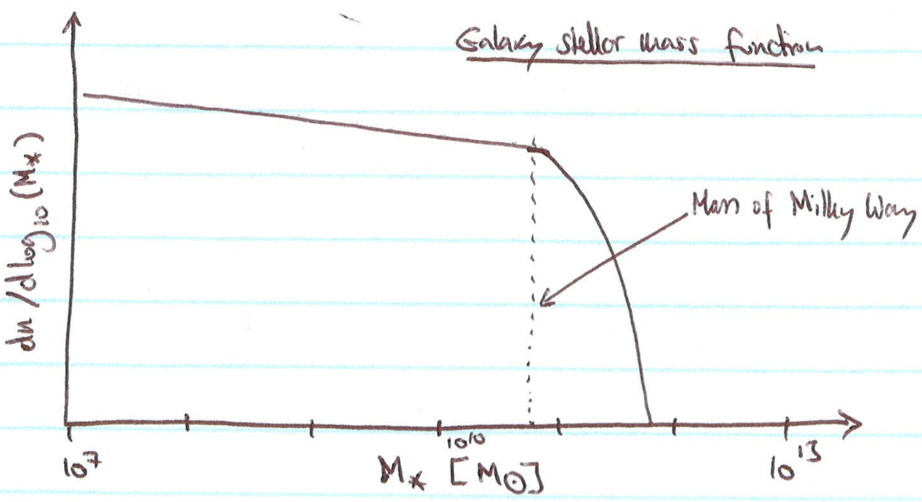
- [Add your own reasons if appropriate]

1.3: Milky Way as a 'typical' galaxy

This is arguably a biased view, but the Milky Way can be considered a 'typical' galaxy.

Stellar mass: $M_* \approx 5 \cdot 10^{10}\ M_\odot$

Total mass: $M_{tot} \approx 10^{12}\ M_\odot$ (Xue et al., 2008)



N.B.: very different from DM halo mass function – see § 3.5

2 FORMATION OF HALOES

- Basic framework of galaxy formation is now well-established – e.g. White & Rees (1978), White & Frenk (1991), Blumenthal et al. (1984)

- Two-stage process:

  (i) Formation of haloes due to gravity ← do this part first...

  (ii) Gas cooling & condensation inside these haloes.

2.1 Linear structure formation

Haloes formed from small density fluctuations in the early Universe, probably seeded / amplified during inflation (e.g. Press & Schechter, 1974). These are usually characterised by their <u>overdensity</u> w.r.t. the background,

$$\delta(\underline{x}) \equiv \frac{\rho(\underline{x})}{\rho_0} - 1$$

(i.e., very small fluctuations have $\delta \sim 0$).

Overdense regions ($\delta > 0$) attract matter through gravity, while underdense regions lose matter, so fluctuations tend to grow with time.

<u>But</u>: Expansion of Universe creates 'drag' which tends to pull matter apart and slow down growth of fluctuations.

⇒ DM fluctuations begin to grow from $z_{\gamma m} \simeq 3100$, (matter-radiation equality) with $\delta \propto t^{2/3}$

→ 'Linear growth' as long as $\delta \ll 1$: Overdensity field retains same shape and is scaled by (growing) factor.

→ Eventually, fluctuations have grown to the point where $\delta \sim 1$, and linear theory no longer applies — non-linear growth.

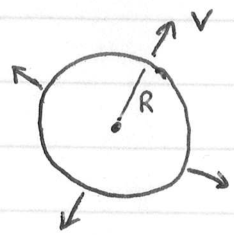
Non-linear growth is challenging to calculate analytically! Two solutions:

(i) Make simplifying assumptions (e.g. spherical symmetry)

(ii) Calculate growth numerically with simulations

2.2 Spherical collapse

Consider the simplified case of a spherically symmetric density perturbation which is affected only by gravity (i.e., no baryonic processes). Due to the expansion of the Universe, the perturbation will (initially) move outwards from its centre.



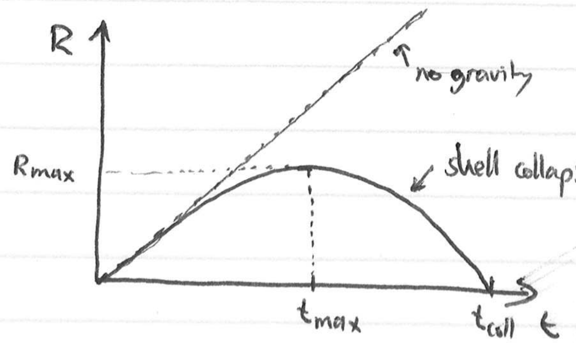
Due to symmetry, we can break this down into a set of (concentric) shells. Each of these moves according to (Newton):

(2.2.1) $$\frac{d^2r}{dt^2} = -\frac{GM(<r)}{r^2}$$

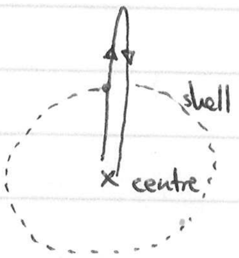
— same as equation for projectile launched vertically from the surface of a spherical body of mass M (Binney & Tremaine 2008)

→ unless velocity is too high, the expansion will eventually stop & turn to collapse

(we will assume that this is the case)



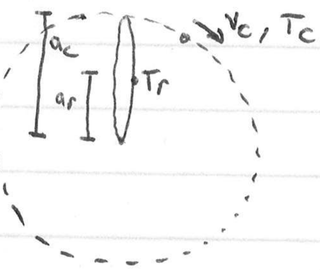
The motion of a point of the shell is (part of) an extreme radial orbit:



$\Rightarrow$ can use solutions for Keple orbits (with eccentricity $e = 1$)

$$\left.\begin{aligned} r &= a(1 - e\cos\eta) \\ t &= \frac{T_r}{2\pi}(\eta - e\sin\eta) \end{aligned}\right\}$$ see Binney & Tremaine

with T_r the (radial) period – find using Kepler III:



$$a_c = 2a_r \Rightarrow T_c^2 = 8T_r^2 = \frac{(2\pi r_{max})^2}{v_c^2}$$

$$\Rightarrow T_r^2 = \frac{4\pi^2 r_{max}^2 \cdot r_{max}}{8GM} = \frac{\pi^2 r_{max}^3}{2GM}$$

$$\Rightarrow \frac{T_r}{2\pi} = \sqrt{\frac{r_{max}^3}{8GM}}$$

$$\Rightarrow r = \frac{r_{max}}{2}(1 - \cos\eta) \quad (2.2.2) \quad ; \quad t = \sqrt{\frac{r_{max}^3}{8GM}}\,(\eta - \sin\eta) \quad (2.2.3) \; *$$

$r = r_{max}$ when $\cos\eta = -1 \;\hat{=}\; \eta = \pi \Rightarrow t_{max} = \sqrt{\dfrac{r_{max}^3 \pi^2}{8GM}}$

- Average density inside sphere: $\rho_s = M \big/ \frac{4}{3}\pi r^3(t)$

  Average background density: $\rho_m = 1 \big/ 6\pi G t^2$ (2.2.4)

$$\delta(t) = \frac{9}{2}\frac{(\eta - \sin\eta)^2}{(1 - \cos\eta)^3} - 1 \quad (2.2.5)$$

$$\Downarrow$$

$$\delta_{max} = \frac{9\pi^2}{16} - 1 \quad (2.2.6)$$

$$\simeq 4.55$$

---

*: This implies that $r = 0$ @ $t = 0$

Relating this to the initial conditions:

If $\delta \ll 1$, then $\eta \ll 1$ (early times) – can expand $\delta(t)$ as a power-law in η:

$$\left.\begin{aligned} \sin\eta &\simeq \eta - \frac{\eta^3}{3!} + \frac{\eta^5}{5!} - \dots \\ \cos\eta &\simeq 1 - \frac{\eta^2}{2!} + \frac{\eta^4}{4!} - \dots \end{aligned}\right\} \quad \delta(t) \simeq \frac{3}{20}\eta^2 \Rightarrow \eta_i^2 = \frac{20}{3}\delta_i$$

Similarly, we have $r_i \simeq \dfrac{r_{max}\,\eta_i^2}{4} \Rightarrow r_{max} = \dfrac{4 r_i \cdot 3}{20\delta_i} = \dfrac{3}{5}\dfrac{r_i}{\delta_i}$

If, at time t_i, the overdensity was δ_i and the radius r_i, the total mass

$$M = \frac{4}{3}\pi(1+\delta_i)\,\rho_m(t_i)\,r_i^3 = \frac{2r_i^3}{9Gt_i^2} \qquad \text{(using 2.2.4)}$$

$$\Rightarrow r_i^3 = \frac{9MGt_i^2}{2}$$

$$r_{max}^3 = \frac{27}{125}\frac{r_i^3}{\delta_i^3}$$

$$r_{max}^3 = \frac{243}{250}\frac{GMt_i^2}{\delta_i^3} \quad \Rightarrow \quad r_{max} = \left(\frac{243}{250}\right)^{1/3}\frac{(GMt_i^2)^{1/3}}{\delta_i}$$

And $t_{max} = \left(\dfrac{\pi^2 r_{max}^3}{8GM}\right)^{1/2} = \pi \cdot \left(\dfrac{243}{2000}\right)^{1/2}\dfrac{t_i}{\delta_i^{3/2}} \simeq 1.095\,\dfrac{t_i}{\delta_i^{3/2}}$ (2.2.7)

Therefore, perturbations with large initial amplitude reach their maximum radius and collapse first.

Result of spherical collapse

According to (2.2.2) & (2.2.3), the perturbation shrinks to a singularity at $t = t_{coll} = 2t_{max}$

→ unrealistic: (even) slight deviations from spherical will lead to an equilibrium halo, supported by velocity dispersion.

- How big is this final halo? → Use Virial Theorem:

$$\underset{\text{pot. energy}}{W} = -2\underset{\text{kinetic energy}}{K} \qquad \Rightarrow K = -\frac{W}{2} \quad \& \quad E = W + K = \frac{W}{2}$$

Now, if the sphere at turnaround had radius r_{max}

$$\Rightarrow E = -\frac{3}{5}\frac{GM^2}{r_{max}} \qquad \text{(pot. energy, kinetic energy zero)}$$

and similarly $W = -\frac{3}{5}\frac{GM^2}{r_{vir}}$ (pot. energy AFTER 'virialisation' of halo)

$$\Rightarrow \underline{r_{max} = 2\, r_{vir}} \qquad (2.2.8)$$

- Density inside virialised halo?

According to (2.2.6), the overdensity at turnaround is $\delta_{max} = \frac{9\pi^2}{16} - 1$

$$\Rightarrow \rho(t_{max}) = \frac{9\pi^2}{16} \cdot \rho_m(t_{max})$$

$$\Rightarrow \rho(t_{coll}) = \left(\frac{r_{max}}{r_{vir}}\right)^3 \cdot \frac{9\pi^2}{16}\rho_m(t_{max})$$

$$= \frac{9}{2}\pi^2 \rho_m(t_{max})$$

And because $\rho_m = \frac{1}{6\pi G t^2}$, $\rho_m(t_{coll}) = \rho_m(t_{max}) \cdot \left(\frac{t_{max}}{t_{coll}}\right)^2$

$$\Rightarrow \frac{\rho(t_{coll})}{\rho_m(t_{coll})} = 18\pi^2 \equiv \Delta_{vir} \qquad (2.2.9)$$

Because $18\pi^2 \simeq 178 \sim 200$, the overdensity of collapsed haloes is often simplified as $\Delta_{vir} = 200$

→ Caveats: (i) All this neglects the influence of Λ - our derivation was based on an Einstein-de Sitter universe.

Bryan & Norman (1998) find that including Λ leads to an overdensity approximately equal to

$$\Delta_{vir} \approx (18\pi^2 + 82x - 39x^2) / \Omega_m(t_{vir}) \qquad (2.2.10)$$

where $x = \Omega_m(t_{vir}) - 1$

(ii) Assumption of spherical symmetry (see below)

(iii) (Strictly) only valid at time of collapse

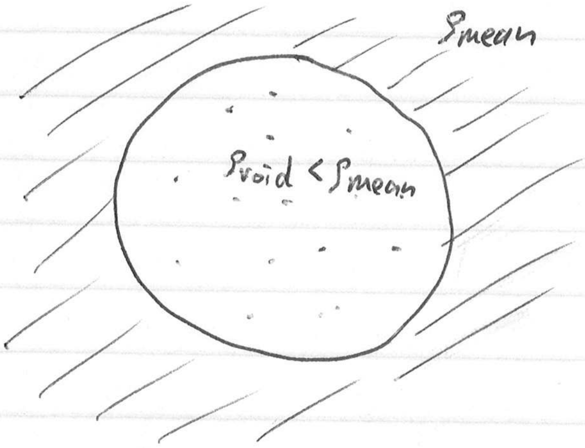
⇒ mass-concentration relation

(iv) Overdensity w.r.t. $\rho_{mean} \neq \rho_{crit.}$ if $\Omega_m \neq 1$

- Mounting evidence that 'haloes' extend significantly beyond 'virial radius', at least on cluster scales

2.3 : Non-spherical collapse

There is no real reason to assume that actual density perturbations grow with spherical symmetry. On the contrary, non-linear density evolution leads to filamentary structures. To see this, consider the evolution of underdense regions (voids):

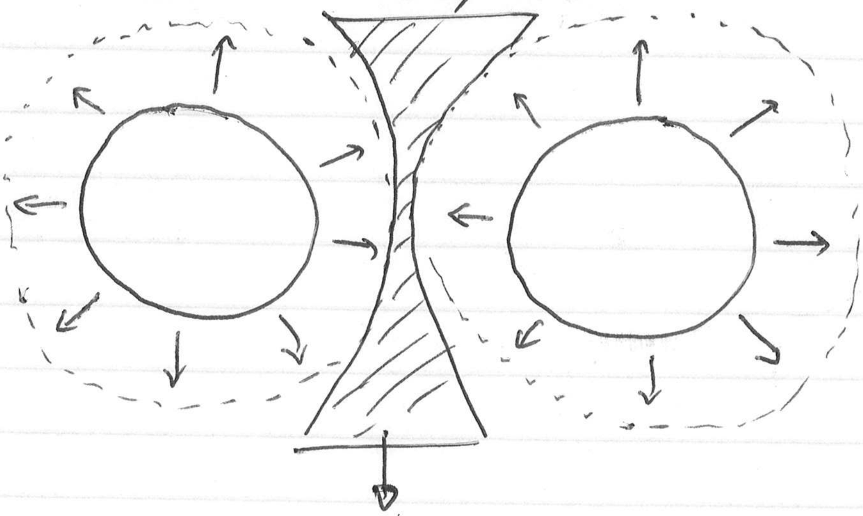


As the universe expands, material in the void is slowed down less than the average by gravity

$\Rightarrow$ void expands in co-moving space

(just like overdense regions shrink)

Now consider two close-by voids:

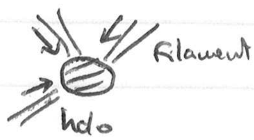


(see Binney & Tremaine 2008, chapter 9.2.2)

flattened 'pancake' formed between expanding voids

→ pancakes themselves are then compressed to filaments, which feed into haloes

"Cosmic Web"



2.4: Hierarchical structure formation

After initial collapse, haloes do not necessarily live in isolation, but can merge to form larger haloes. In fact, mergers are expected to have played a dominant role for the mass growth of all haloes found today.

Recall from (2.2.7) that the first perturbations to collapse are the ones with highest initial overdensity. If the power spectrum follows

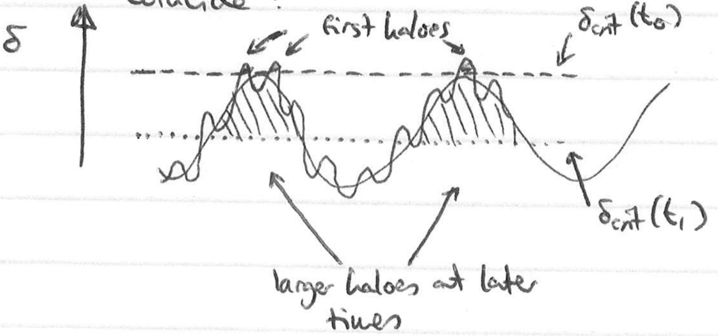
$P(k) \propto k^{-3}$ on small-ish scales ($x \lesssim x_{\gamma m} \approx 100$ Mpc)

(due to suppression of growth on sub-horizon scales in the radiation era)

then the smoothed variance

$$\sigma_k^2 = \frac{1}{2\pi^2} \int_0^k P(k) k^2 \, dk \propto \int_0^k \frac{1}{k} \, dk$$ and each decade contributes equally to the variance

$\Rightarrow$ highest overdensity in regions where crests of different wavelength scales coincide:



$\Rightarrow$ first haloes to form are small-scale haloes

$\Downarrow$

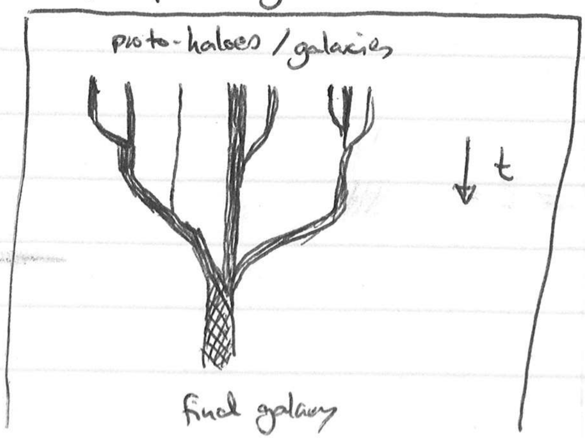
merge to form large structures.

$\rightarrow$ Importance for galaxy formation: Today's galaxies have been assembled from large number of 'proto-galaxies' through mergers

Visualisation of hierarchical structure formation: $\longrightarrow$

"Merge Tree"

(Lacey & Cole, 1993)



2.5: Hot gas haloes

So far, we have only considered the influence of gravity, as appropriate for Dark Matter. But: visible 'galaxies' made up of baryons, which are susceptible to additional effects. In particular:

(i) They are effectively coupled to radiation prior to recombination (at $z_{rec} \approx 1100$, i.e. well after matter-radiation equality)

(ii) They are subject to pressure forces, in addition to gravity.

(iii) Baryons can lose thermal energy through radiation (next section).

Point (ii) implies that baryons can be stabilised against gravitational collapse under appropriate circumstances, i.e. if the perturbation is smaller than the Jeans length $\lambda_J = c_s \sqrt{\pi / G\bar{\rho}}$ (2.5.1)

- Simple derivation: $\lambda_J \approx c_s \cdot t_{\text{free-fall}}$ – maximum distance a sound wave can travel in a free-fall time.

- More rigorous: Perturbing the fluid equations (continuity, Euler, Poisson) leads to the dispersion relation

$$\omega^2 = \frac{k^2 c_s^2}{a^2} - 4\pi G\bar{\rho} \qquad (2.5.2)$$

For perturbations to grow, $\omega^2 < 0 \Rightarrow \frac{2\pi a}{k_J} \equiv \lambda_J = c_s \sqrt{\frac{\pi}{G\bar{\rho}}}$

→ Just before recombination, $c_s \approx c/\sqrt{3}$ and $M_J \sim 10^{17} M_\odot$ ← $\gg M_{gal}$!

due to coupling to photons (see Mo, v.d. Bosch & White)

→ After recombination: $c_s = \left(\frac{5 k_B T}{3 m_p} \right)^{1/2}$ (2.5.3)

(non-rel., monatomic gas)

$$\Rightarrow M_J = \frac{4}{3}\pi \bar{\rho}_{m,0} \left(\frac{\lambda_J}{2}\right)^3 \sim 10^6 \, M_\odot \sim M_{\text{star cluster}} \text{ (glob.)}$$

$\Rightarrow$ Baryons only collapse after recombination on (sub-) galactic scales, and fall into the (already growing) dark matter haloes.

- Dark matter haloes are stabilised by velocity dispersion
  $\rightarrow$ equivalent for gas: <u>thermal motion</u>.

<u>Temperature of gas in collapsed halo?</u>

Start from the Virial Theorem again: $2K = -W$ $\rightarrow$ potential energy:

$$W = -\frac{3}{5}\frac{G M_{gas} M}{r}$$

thermal energy:

$$K = \frac{3}{2} N k_B T = \frac{3 M_{gas} k_B T}{2 \mu m_p}$$

$$\Rightarrow \frac{3 k_B T}{\mu m_p} = \frac{3}{5}\frac{GM}{r} \quad (*)$$

N = number of atoms
k_B = Boltzmann's constant
T = temperature
M_{gas} = gas mass in halo
M = total mass in halo
G - Newton's constant
r = halo radius

Now,

$$M = \frac{4}{3}\pi \rho r^3$$

$$\Rightarrow r = \left(\frac{3M}{4\pi\rho}\right)^{1/3}$$

$$\Rightarrow T = \frac{G \cdot M^{2/3} \cdot \mu \cdot m_p \cdot (4\pi)^{1/3}}{5 \cdot 3^{1/3} \cdot k_B} \cdot \rho^{1/3}$$

and $\rho \equiv (1+\delta)\bar{\rho} = (1+\delta)(\Omega_{m,0} h^2)(1+z)^3 \frac{3\mathcal{H}_0^2}{8\pi G}$

where δ = overdensity

$\Omega_{m,0}$ = present-day matter density parameter

$\mathcal{H}_0 \equiv 100 \text{ km}/\text{s Mpc}$

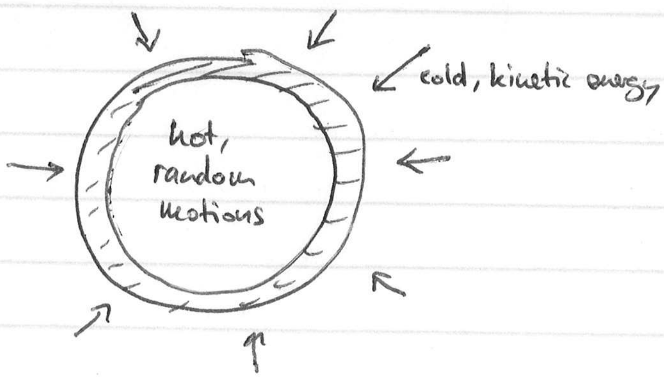
$$\Rightarrow T = \frac{1}{2^{1/3}\cdot 5} \cdot \frac{(G H_0)^{2/3} \cdot \mu m_p}{k_B} \cdot (1+\delta)^{1/3} \cdot (\Omega_{m,0} \cdot h^2)^{1/3} \cdot (1+z) \cdot M^{2/3}$$

$$\simeq 6 \cdot 10^4 \, K \cdot (1+\delta)^{1/3} (\Omega_{m,0} h^2)^{1/3} (1+z) \cdot M_{12}^{2/3} \qquad (2.5.4)$$

with $M_{12} \equiv \dfrac{M}{10^{12} M_\odot}$

For example, a halo that has just collapsed ($\delta \approx 200$) at $z = 3$ in the ΛCDM cosmology ($\Omega_{m,0} h^2 \approx 0.15$) with $M_{12} = 1$ has a gas temperature of $\underline{T \sim 8 \cdot 10^5 \, K}$.

Aside: Physical mechanism that yields these temperatures is <u>shock heating</u> at the halo boundary



→ only works if galaxy can support a stable accretion shock (see next section).

3 GAS COOLING

Gas observed near the halo centre (in the actual 'galaxies') is not hot ($T \sim 10^6$ K). Several reasons for this, including

- Observed 21-cm signal in radio waves from atomic hydrogen shows the gas in the galaxies is mostly neutral, but at $T \sim 10^6$ K almost all gas is ionized.

- Jeans mass in 10^6 K gas is much too large for star formation: $M_J > M_{halo}$ by definition (gas stabilised by thermal motion)

- Disk shape and rotation support of gas (see next section)

$\Rightarrow$ Galaxy formation requires gas halo to cool!

3.1: Physical processes responsible for gas cooling

- Gas acquired its high temperature through shock heating
  - $\rightarrow$ increase in entropy, irreversible!
  - $\rightarrow$ Need to get rid of thermal energy by other means: interaction with photons (= emission of photons)

  Three main mechanisms relevant for galaxy formation:

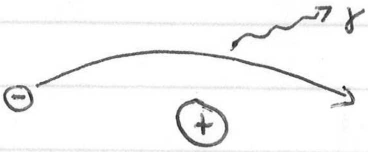
  (i) Bremsstrahlung

  (ii) Collisional ionisation + recombination

  (iii) Collisional excitation + radiative de-excitation

  $\rightarrow$ N.B.: Assume optically thin gas, i.e. radiation not re-absorbed.

(i) Bremsstrahlung ('free-free')



Basic principle: electrodynamics requires radiation emission whenever a charged particle is accelerated.

→ requires gas to be ionised: most effective at $T \gtrsim 10^6\,K$.

- Cooling rate $\mathcal{C}_{ff} \equiv \frac{dE}{dt} = \int \varepsilon_{ff}(\nu)\, d\nu$ ← "Free-free emissivity"

  → N.B.: per unit volume

with $\varepsilon_{ff}(\nu) = n_i n_e \int P(v,\nu)\, f(v)\, dv$

n_i: ion density; n_e: electron density

$P(v,\nu)$: power radiated at frequency ν for electron at speed v → electrodynamics

$f(v)$: Electron velocity (speed) distribution – Maxwellian

For a (fully) ionised plasma, $n_i = n_e$ and $\varepsilon_{ff}(\nu) \propto n_e^2\, T^{-1/2}\, e^{-\frac{h\nu}{kT}}$

$$\Rightarrow \mathcal{C}_{ff} \approx 1.4 \cdot 10^{-23} \left(\frac{T}{10^8 K}\right)^{1/2} \left(\frac{n_e}{cm^{-3}}\right)^2 \text{erg } s^{-1}\, cm^{-3}$$

(for more details, see e.g. Sutherland & Dopita, 1993 or Mo, v.d. Bosch & White, 2010)

→ convenient to separate density and temperature-dependent parts:

$\mathcal{C}_{ff} \equiv n_H^2\, \Lambda(T)$ ← "Cooling function", depends only on T

Hydrogen atom number density

From above $\underline{\Lambda_{ff} \propto T^{1/2}}$.

→ Bremsstrahlung most effective at high temperatures

$\hat{=}$ highest halo masses (galaxy clusters), but also relevant to galaxy scales.

(ii) Collisional ionisation + radiative recombination

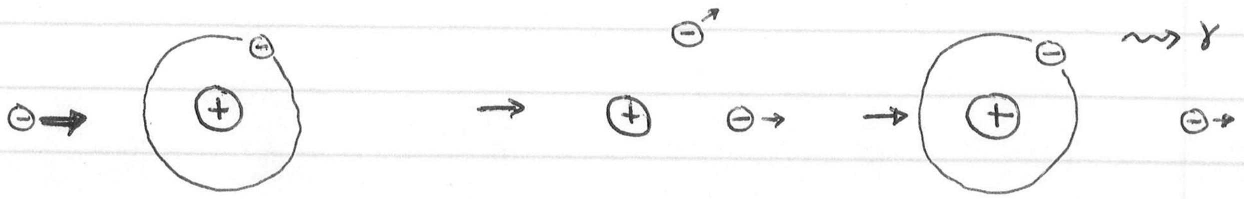
Bremsstrahlung is only effective in (mostly) ionised gases.
Estimate minimum temperature:

$$E = \frac{3}{2} kT \quad \text{for one electron} \rightarrow \text{must exceed } E_I$$

(= 13.6 ~~k~~eV)
↑ for H

$$\Rightarrow T_{min} \sim \frac{E_I}{k} \simeq \underline{1.5 \cdot 10^5 \text{ K}}$$

→ expect significant neutral gas below this temperature.



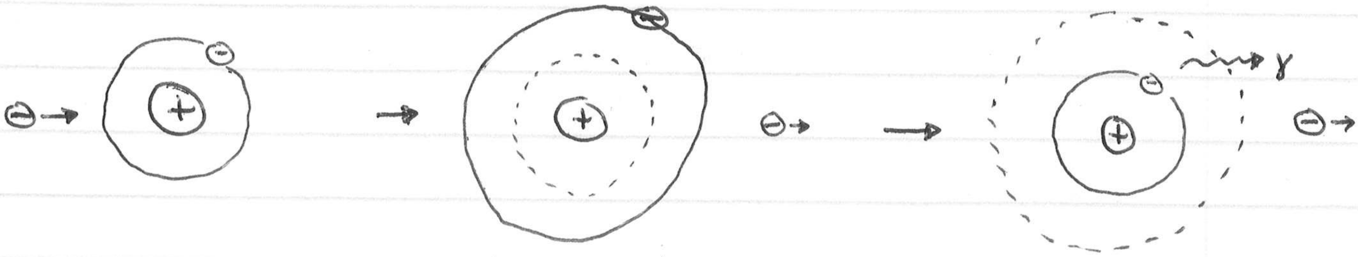
Electron ($E > 13.6$ eV) collides with atom

Ionised atom

Recombination: $E_{thermal} \rightarrow E_\gamma$

(iii) Collisional excitation

We had argued above that at $T \lesssim 2 \cdot 10^5$ K most electrons are <u>not</u> energetic enough to ionise H atoms ⇒ more relevant: <u>excitation</u>.



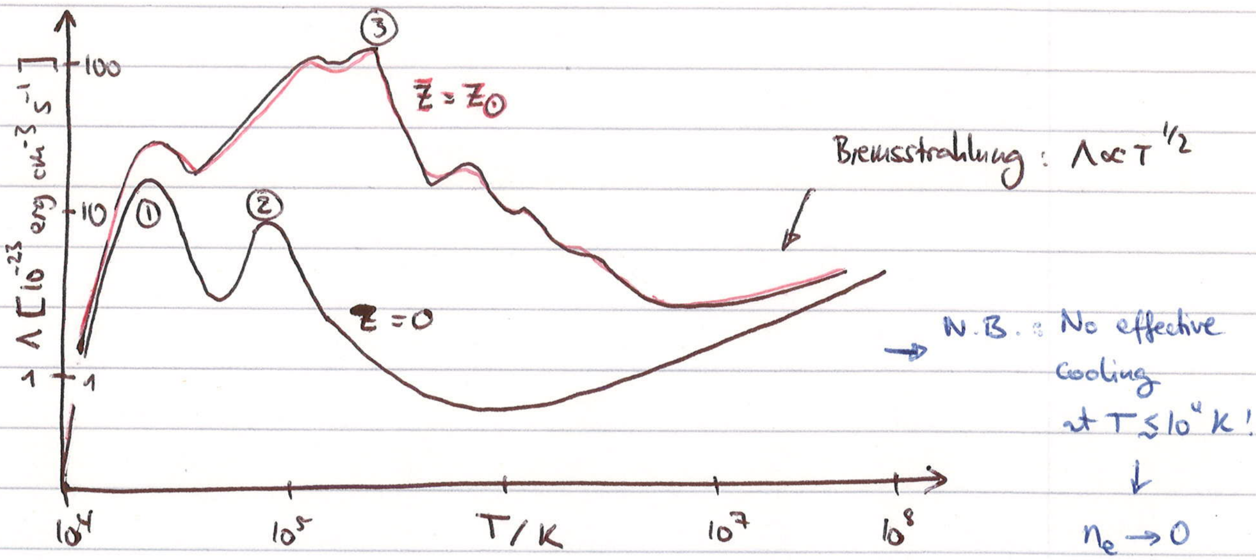
electron excited to higher energy level

→ lower radiation energy per occasion, but much more common ⇒ <u>dominant</u>.

→ resulting cooling functions can be calculated from Quantum mechanics (see Sutherland & Dopita 1993)

→ If gas is non-primordial (metals), there will be more excitations possible and therefore the cooling function will be higher.

→ Overall cooling function looks like this:



Some features: ①: Excitations in H^0 at $T \sim 15\,000$ K

②: Excitations in He^+ at $T \sim 10^5$ K

③: Metal line excitations (esp. C, O, Ne, Fe)

⇒ Presence / absence of metals can have significant impact on cooling rate!

3.2 Gas cooling times

We can estimate how long it takes for the hot gas to radiate away all its thermal energy – the 'cooling time' t_{cool}.

$$t_{cool} \equiv \frac{\rho \cdot \varepsilon}{\mathcal{C}}$$

with ε = specific internal energy (per unit mass, hence factor ρ)

$$\Rightarrow t_{cool} = \frac{3nk_BT}{2n_H^2\Lambda(T)} \simeq 3.3\cdot10^9\,\text{yr}\left(\frac{T}{10^6\,\text{K}}\right)\left(\frac{n}{10^{-3}\,\text{cm}^{-3}}\right)^{-1}\left(\frac{\Lambda(T)}{10^{-23}\,\text{erg s}^{-1}\,\text{cm}^{-3}}\right)^{-1}$$

→ Values for 'typical' galaxy? Example before:

$T \approx 8\cdot10^5\,\text{K}$ $\qquad \Lambda \approx 0.5\cdot10^{-23}\,\text{erg s}^{-1}\,\text{cm}^{-3}$

To find particle number density n:

$$n = \frac{27}{12}n_H \quad \text{(assuming fully ionised)}$$

$$= \frac{27}{12}\cdot\frac{\rho X}{m_p} \quad (\text{with } X = \text{hydrogen mass fraction} \approx 0.752)$$

$$= \frac{27}{12}\frac{X}{m_p}(1+\delta)(\Omega_{m,0}h^2)(1+z)^3\cdot\frac{3H_0^2}{8\pi G} \quad \text{(see section 2.5)}$$

$$\simeq 5.5\cdot10^{-3}\,\text{cm}^{-3}\cdot\left(\frac{1+z}{4}\right)^3$$

$\Rightarrow \underline{t_{cool} \approx 10^9\,\text{yr}}$ ← less than age of Universe - good!!

restores equilibrium ($\rho\uparrow \Rightarrow T\uparrow$)

BUT: As the gas cools, its Jeans mass reduces and it collapses

$\Rightarrow$ cooling can only be successful if $t_{cool} \lesssim t_{free\text{-}fall}$!

For example galaxy: $t_{ff} = \sqrt{\frac{3\pi}{32G\rho}} \approx \underline{4\cdot10^8\,\text{yr}}$

$\Rightarrow$ Do NOT expect effective cooling!

→ if $t_{cool} > t_{free\text{-}fall}$: 'Quasi-hydrostatic'

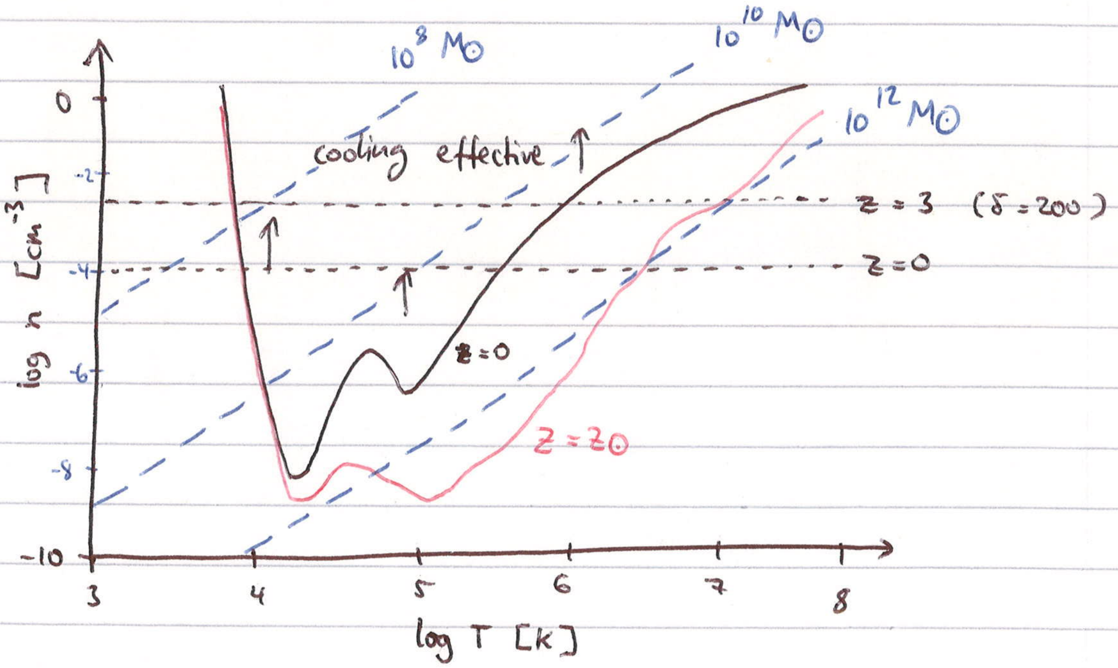
We can estimate in which cases cooling is effective (and galaxies form!) by considering

$$t_{cool} = t_{ff}$$

Writing $\rho = \frac{n \mu m_p}{f_{gas}}$, we have $t_{ff} = \sqrt{\frac{3\pi f_{gas}}{32 G n \mu m_p}}$

$$\simeq 2.1 \cdot 10^{9}\,\mathrm{yr}\ f_{gas}^{-1/2} \cdot \left(\frac{n}{10^{-3}\,\mathrm{cm}^{-3}}\right)^{-1/2}$$

Equating this to t_{cool} and solving for $n(T)$ gives the following (left as exercise for the reader, or see Mo, v.d. Bosch & White):



From section 2.5: $T \propto M^{2/3} n^{1/3} \Rightarrow \log n = 3\left(c + \log T - \frac{2}{3}\log M\right)$ proportionality constant

→ haloes of given mass lie on straight line with slope $= 3$, intercept proportional to $\log M$

→ dashed lines (blue)

$\Rightarrow$ expect cooling to work in haloes with $10^{7} \lesssim \frac{M_{gas}}{M_\odot} \lesssim 10^{12}$

This agrees quite well with the mass range of actual galaxies (considering not all gas necessarily turns into stars, see below).

CAVEATS:

(i) Mergers can lead to build-up of larger galaxies than 'allowed' by direct cooling.

(ii) Most galaxies formed from mergers of small haloes at high redshift → cooling expected to be very effective for these!

Observations indicate that only a small fraction of baryons are actually turned into stars

⇒ suggests that cooling is offset by heating to large extent.

- Photoheating from UV / X-ray background (see e.g. Wiersma et al., 2009)
- Energy injection from supernova explosions

(iii) The usual warnings about oversimplifications: spherical symmetry, ionisation equilibrium, uniform distribution, ...

Nevertheless, provides very useful insight.

3.3: Cooling radius

Let's relax the assumption of uniform gas distribution, and consider haloes with power-law (adiabatic) gas profiles:

$$\rho = \rho_0 \left(\frac{r}{r_0} \right)^{-\alpha} \quad ; \quad P = P_0 \left(\frac{r}{r_0} \right)^{-\beta}$$

$\Rightarrow$ (ideal gas law) : $P = n k_B T \quad : \quad T = \dfrac{P \mu m_p}{\rho k_B}$

$$\Rightarrow T = T_0 \left(\frac{r}{r_0} \right)^{\alpha - \beta} \qquad \text{with } T_0 = \frac{P_0}{\rho_0} \frac{\mu m_p}{k_B}$$

If we also approximate the cooling function as a (piecewise) power-law, we have

$$\Lambda(T) = \Lambda_0 \left(\frac{T}{T_0} \right)^{\nu}$$

So that the cooling time is given by

$$t_{cool} = \frac{3 n k_B T}{2 n_H^2 \Lambda(T)} = \frac{3 k_B}{2} \cdot \left(\frac{n^2}{n_H^2} \right) \cdot \frac{\mu m_p}{\rho} \cdot \frac{T}{\Lambda(T)}$$

$$= \frac{3 k_B \mu m_p}{2} \left(\frac{n}{n_H} \right)^2 \cdot T_0 \left(\frac{r}{r_0} \right)^{\alpha - \beta} \rho_0^{-1} \left(\frac{r}{r_0} \right)^{\alpha} \Lambda_0^{-1} T^{-\nu} T_0^{\nu}$$

$$= \frac{3 k_B \mu m_p}{2} \left(\frac{n}{n_H} \right)^2 \cdot \frac{T_0}{\rho_0 \Lambda_0} \cdot T_0^{-\nu} \left(\frac{r}{r_0} \right)^{-\nu(\alpha - \beta)} T_0^{\nu} \left(\frac{r}{r_0} \right)^{2\alpha - \beta}$$

$$= \underbrace{\frac{3 k_B \mu m_p}{2} \left(\frac{n}{n_H} \right)^2 \cdot \frac{T_0}{\rho_0 \Lambda_0}}_{t_0} \left(\frac{r}{r_0} \right)^{\overbrace{2\alpha - \beta - \nu\alpha + \nu\beta}^{1/\tau}}$$

$$\equiv t_0 \left(\frac{r}{r_0} \right)^{1/\tau}$$

For an isothermal halo, $\alpha = \beta = 2 \Rightarrow \tau = 1/2$ (indep. of ν !)

$\Rightarrow$ Cooling time is shortest near the centre (highest density), and becomes longe at large radii

We can also estimate the cooling radius r_{cool}, s.t. $t = t_{cool}$:

$$t = t_0 \left(\frac{r_{cool}}{r_0} \right)^{1/\tau} \Rightarrow \underline{r_{cool} = r_0 \left(\frac{t}{t_0} \right)^{\tau}}$$

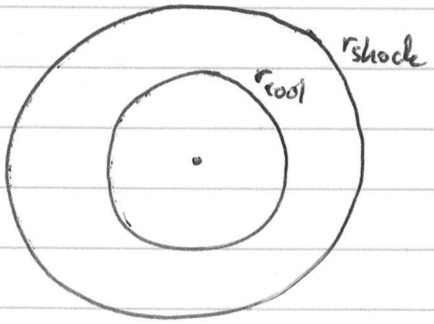
$\Rightarrow$ cooling radius moves outward with time - 'cooling wave'.

3.4 Hot and cold accretion

We have so far assumed that the galaxy is surrounded by / formed from gas which was shock-heated to high temperatures, but this is only the case if there is actually a (stable) accretion shock to heat the infalling gas.

$\rightarrow$ Paper by Birnboim & Dekel (2003): When is this the case?

Consider a system with spherical symmetry:



If $r_{shock} \gg r_{cool}$: Gas is first shock-heated & then cooled

But if $r_{shock} \ll r_{cool}$, no shock-heating happens:

"Cold flow" directly to centre

We have seen above that $r_{cool} \propto \Lambda(t)^{1/2}\, t^{1/2}$

and $r_{shock} \propto r_{vir} \propto V_c \cdot t$ (see White & Frenk, 1991)

$\Rightarrow$ shock grows faster than cooling radius

$\Rightarrow$ expect cold accretion to be more important at high redshift

$\rightarrow$ also: expect shocks to be stable in massive haloes at any given epoch (V_c dependence of r_{shock})

$\Rightarrow$ "Critical mass" above which hot mode dominates (and vice versa).

Birnboim & Dekel (2003) show that shock stability requires

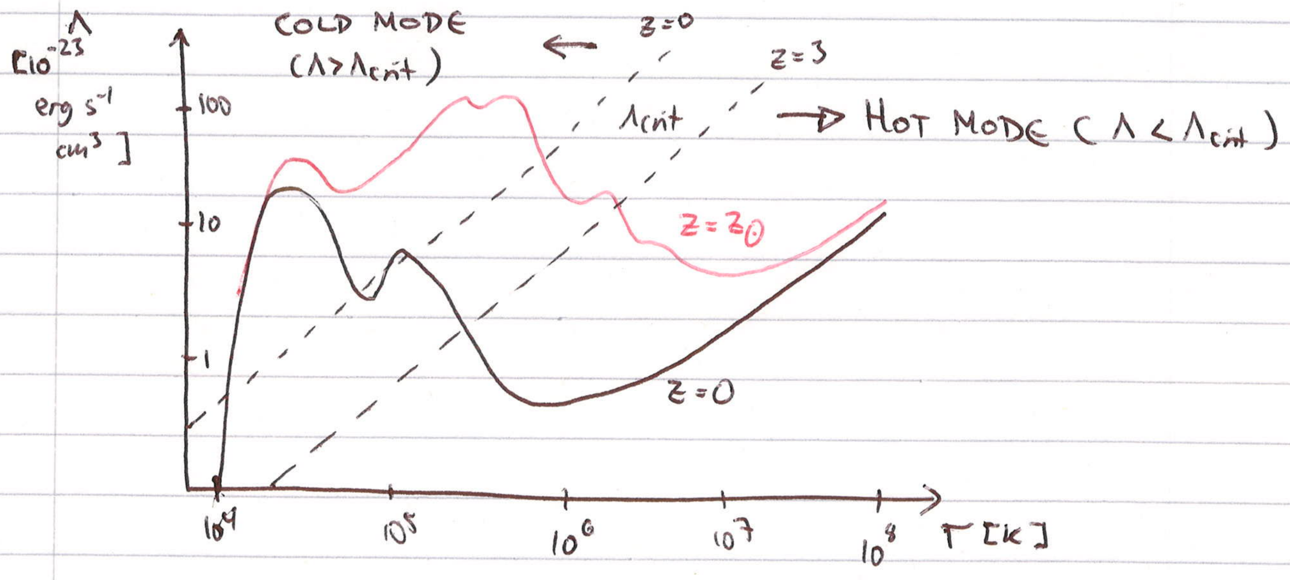
$$\gamma_{eff} \equiv \frac{d\ln P}{d\ln \rho} > \frac{2\gamma}{\gamma + 2/3}.$$

Combining this with the Rankine-Hugoniot jump conditions at the shock interface, one finds that a stable shock requires

$$\underset{\substack{\uparrow \\ \text{post-shock} \\ \text{Temp.}}}{\Lambda(T_2)} < \Lambda_{crit} = 0.022 \frac{m_p^2 |v_1|^3 \leftarrow \text{accretion velocity}}{\underset{\substack{\uparrow \\ \text{pre-shock density}}}{\rho_1} r_{shock}}$$

Assuming that $\rho_1 = \bar{\rho}_{baryon}$, $v_1 = V_{vir}$ and a dark matter overdensity in the halo of δ_{vir},

$$\Lambda_{crit} \approx 6 \cdot 10^{-22} \text{ erg cm}^3 \text{ s}^{-1} \left(\frac{T_{vir}}{10^6 K}\right)\left(\frac{\delta_{vir}}{100}\right)^{1/2}\left(\frac{\Omega_{b,0} h^2}{0.024}\right)^{-1}\left(\frac{h}{0.7}\right)(1+z)^{-3}$$



- Depending on metallicity and redshift, critical temperature $T_{crit} \approx 10^5 - 10^6$ K $\hat{=}$ M_{crit} of $\sim 10^{11} - 10^{12}\ M_\odot$

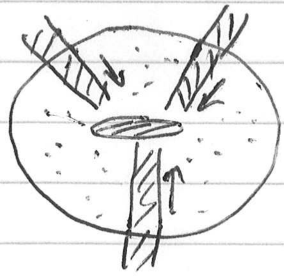
⇒ "Typical" galaxies like Milky Way are expected to lie in the HOT mode part.

CAVEATS:

(i) X-ray emission from these hot, cooling haloes is still not fully convincingly detected, and certainly not at level expected from the simple picture outlined here

→ Crain et al., 2010: Recent simulations agree with this, due to entropy increase at $z \sim 2$ (due to supernova feedback)

(ii) Simple 'either-or' picture is likely oversimplified: Plausible to have both hot halo and cold streams:

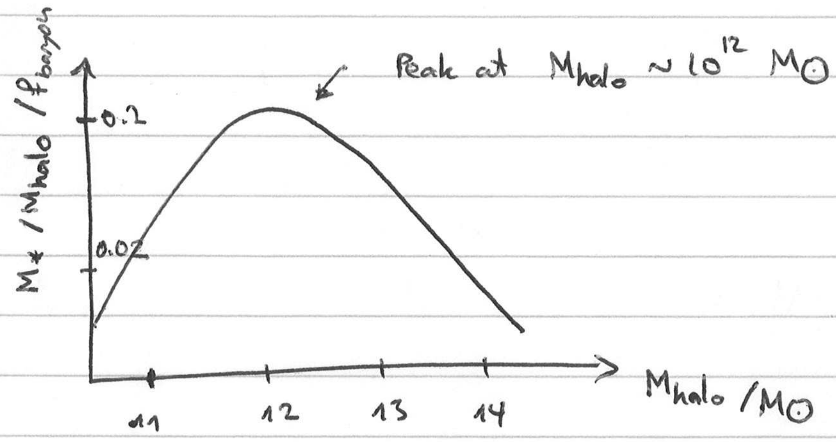


→ still subject of active research.

3.5 Mass-to-light ratio of galaxies

The 'critical' mass for support of a hot halo is suspiciously close to the 'knee' in the galaxy stellar mass function (section 1.3), for a Milky-Way M_* / M_{total} ratio. Look at this in slightly more detail, by comparing the stellar and halo masses of galaxies

(see e.g. Moster et al., 2010)



Galaxies at $M_{halo} \sim 10^{12} M_\odot$ are most effective at converting their gas into stars.

→ Still far from 100% conversion ⇒ effective feedback

→ Star formation appears much less effective at the low- and high mass end. Possible explanations:

(i) Low-M: higher effectivity of supernova feedback

$E_{SN} \propto M$; $W_{pot} \propto -M^2$

⇒ supernovae can remove less gas from a Milky-Way-like galaxy than from a much smaller dwarf.

(ii) High-M: Influence of accreting super-massive black holes ('AGN feedback')

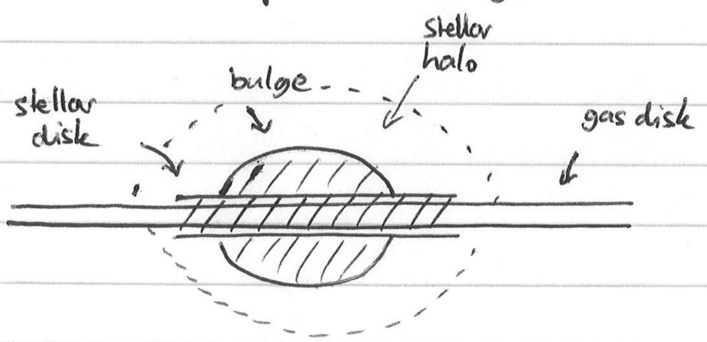
4 FORMATION OF GALACTIC DISKS

We have so far considered the central halo regions mostly as 'dump' of cooled gas, without looking at their detailed structure. But these central, baryon-rich regions are actually the 'galaxies' that are observed in optical / IR / UV / radio / ...

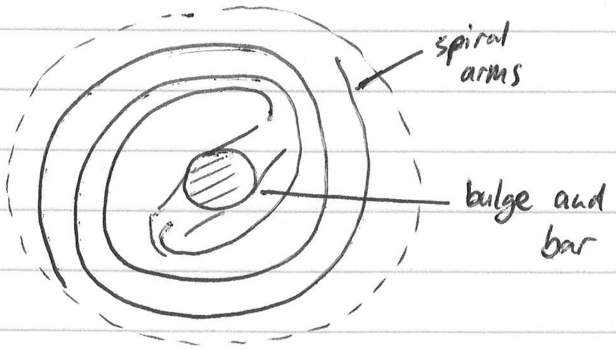
→ look at them in more detail now.

4.1 Structure of (actual) galaxies

If we again take the Milky Way as a 'typical' example galaxy, we have most components that galaxies can feature:



Side-on view



face-on

→ gas forms a highly flattened, rotating disk

→ some stars (but not necessarily all) also lie in a disk, but typically less extended than the gas disk

→ Both gas and stars often show spiral structure, and exponential surface density profiles

Bulges presumably form through galaxy mergers, which disrupt the disk structure

$\Rightarrow$ disk as 'undisturbed' state, result of gas cooling

4.2 Formation of gas disks

As the gas cools, it loses energy and sinks deeper into the potential well of the halo.

__But__: Gas cannot (easily) lose its angular momentum

$\rightarrow$ radiation emitted isotropically

$\Rightarrow$ gas spins up and becomes flattened (centrifugal force), until it is supported fully by rotation and

$$\frac{GM(<r)}{r^2} = \frac{v(r)^2}{r}$$

$\rightarrow$ Initial angular momentum provided by tidal torques from surrounding large-scale structure.

__Sounds good – does this actually work?__

$\rightarrow$ Consider time required to reach rotationally supported configuration:

(following Efstathiou & Silk, 1983)

Spin parameter $\lambda \equiv \dfrac{J|E|^{1/2}}{GM^{5/2}}$ – dimensionless parameter, higher for faster rotation

For an isolated exponential disk:

$$\Sigma = \Sigma_0 \, e^{-R/R_d}$$

which (for a thin disk) leads to a rotation curve

$$V_c^2(R) = -4\pi G \Sigma_0 R_d y^2 \left[I_0(y) K_0(y) - I_1(y) K_1(y) \right]$$

where $y = R/2R_d$ and I_n, K_n are modified Bessel functions of the first and second kinds

The disk's total angular momentum is therefore

$$J_D = 2\pi \int_0^\infty V_c(R)\, \Sigma(R)\, R^2\, dR \approx 1.11\, G^{1/2} M_d^{3/2} R_d^{1/2}$$

where M_d is the total mass of the disk:

$$M_d = 2\pi \int_0^\infty \Sigma(R)\, R\, dR = 2\pi \Sigma_0 R_d^2$$

The disk's energy can be calculated from the virial theorem:

$$W = -2K \Rightarrow E = -K = -\int \frac{1}{2} v^2\, dM$$

$$= -\pi \int_0^\infty V_c^2(R)\, \Sigma(R)\, R\, dR \approx -0.147\, G M_d^2 R_d^{-1}$$

Combining all these results, we obtain a spin parameter of

$$\lambda_D = \underbrace{1.11\, G^{1/2} M_d^{3/2} R_d^{1/2}}_{J_d} \underbrace{(0.147)^{1/2} G^{1/2} M_d R_d^{-1/2}}_{|E|^{1/2}} G^{-1} M_d^{-5/2}$$

$$\approx 1.11 \cdot \sqrt{0.147} = 0.425$$

To put this into perspective, tidal torques are expected to produce spin parameters in the range $0.01 \lesssim \lambda < 0.1$ with median

$\tilde{\lambda} \approx 0.05$ (see e.g. Bullock et al., 2001)

During collapse, J & M are conserved, and the binding energy decreases:

$E \propto R^{-1}$

$\Rightarrow$ spin parameter evolves as $\lambda = \lambda_i \left(\frac{R}{R_i} \right)^{-1/2}$

$$\Rightarrow \frac{R}{R_i} = \left(\frac{\lambda}{\lambda_i} \right)^2 \approx \left(\frac{0.425}{0.05} \right)^2 = 72$$

$\Rightarrow$ Disk must have shrunk by factor ~ 70 before reaching rotational equilibrium.

For a disk like the Milky Way ($M \sim 5 \cdot 10^{10}\, M_\odot$, $R \sim 10$ kpc) this gives an initial size of ~ 700 kpc and density

$$\rho_i \sim \frac{3M}{4\pi R_i^3} \approx 2.4 \cdot 10^{-27}\ \mathrm{kg/m^3}$$

which gives a free-fall time of

$$t_{ff} = \sqrt{\frac{3\pi}{32 G\rho}} \approx 1.4 \cdot 10^{18}\ \mathrm{s} \mathrel{\hat{=}} 4.3 \cdot 10^{10}\ \mathrm{yr} \quad (!)$$

($\gg t_{universe}$)

$\Rightarrow$ no way for <u>isolated</u> gas cloud to reach rotational equilibrium!

- But if the disk is embedded in a dark matter halo, such that

$$V_c(R) = \text{const.} = V_c$$

and $\rho_H = v_c^2 / 4\pi G r^2$ (s.t. $M_H = \int_0^{r_H} 4\pi r^2 \rho_H(r)\, dr$

$= \frac{V_c^2}{G} \cdot r_H \Rightarrow V_c^2 = \frac{G M_H}{r_H}$)

things look different:

Assume gas conserves its specific angular momentum during contraction:

$$J = R \cdot V_c = R_i \, V_{rot,i} \quad \Rightarrow R_i = R \cdot \frac{V_c}{V_{rot,i}}$$

initial rotation velocity at initial radius (R_i)

From numerical ~~sim~~ simulations, Bullock et al. (2001) find that

$$J(r) \propto r^{1.1 \pm 0.3}$$

, so we can (for simplicity) assume

$$J(r) \propto r \Rightarrow V_{rot,i} = \text{const.}$$

$$\Rightarrow V_{rot,i} = \eta V_c \quad \text{with } \eta = \frac{R}{R_i} \text{ is the contraction factor}$$

→ find $V_{rot,i}$ from λ_i (omit subscript 'i' for simplicity)

$$J_H = \int_0^{r_H} 4\pi r^2 \rho_H(r) \, V_{rot} \, r \, dr \qquad (r_H = \text{halo truncation radius})$$

$$= \frac{V_c^2}{G} V_{rot} \int_0^{r_H} r \, dr = \underline{\frac{V_c^2 V_{rot} r_H^2}{2G}} \qquad (*)$$

$$E = -K = -\frac{1}{2} M_h V_c^2$$

because V_c is the same throughout the halo (N.B.: not V_{rot}!)

$$\Rightarrow \lambda = \frac{J \, |E|^{1/2}}{G M^{5/2}} = \frac{V_c^2 V_{rot} r_H^2 M_H^{1/2} V_c}{2G \cdot 2^{1/2} \, G M_H^{5/2}}$$

$$= \frac{V_c^4 \cdot \eta \cdot r_H^2}{2^{3/2} G^2 M_H^2} \qquad \text{because } V_{rot} = \eta V_c$$

$$= 2^{-3/2} \eta \qquad \text{because } V_c^2 = \frac{G M_H}{r_H}$$

$\Rightarrow$ Contraction factor $\eta = \lambda_i \, 2^{3/2} \approx 0.14$

$\Rightarrow \dfrac{R_i}{R} \approx 7$ (order of magnitude lower!)

and because $t_{ff} \propto \rho^{-1/2}$ and $\rho_i \propto R_i^{-3}$ $\Rightarrow t_{ff} \propto R_i^{3/2}$

$\Rightarrow$ revised free-fall time ~ 30 times lower ~ 1.4 Gyr $\ll t_{univ.}$

Therefore, the existence of rotationally supported disks is only possible with dark matter haloes - irrespective of rotation curve shape!

4.3 Reason for exponential surface density profiles

Why do galactic disks follow an exponential surface density profile? (At least) two possible explanations:

(i) This is a consequence of the initial conditions (distribution in dark matter haloes of angular momentum)

(ii) It reflects some process happening in the disk, such as angular momentum transport through viscosity.

→ Angular momentum in DM haloes

From their simulations, Bullock et al. (2001) find the following specific angular momentum distribution in DM haloes:

$$P(J) = \frac{\mu J_0}{(J_0 + J)^2}$$

(μ = free parameter)
(J_0 related to J_{tot})

Assume for the moment that each individual gas element conserves its specific angular momentum ('strong conservation')

Then, if the gas has the same (initial) J distibution as the DM, we have

$$\frac{M_d(<R)}{M_d} = \frac{M_h(<J)}{M_{vir}} \equiv \int_0^J P(J')\,dJ'$$

(material further out has higher angular momentum, $V_c = \text{const}$)

$$\Rightarrow \Sigma(R) = \frac{1}{2\pi R} \frac{dM_d(<R)}{dR}$$

$= V_c$ because $J = V_c R$

$$= \frac{1}{2\pi R} \cdot M_d \cdot P(J) \cdot \frac{dJ}{dR}$$

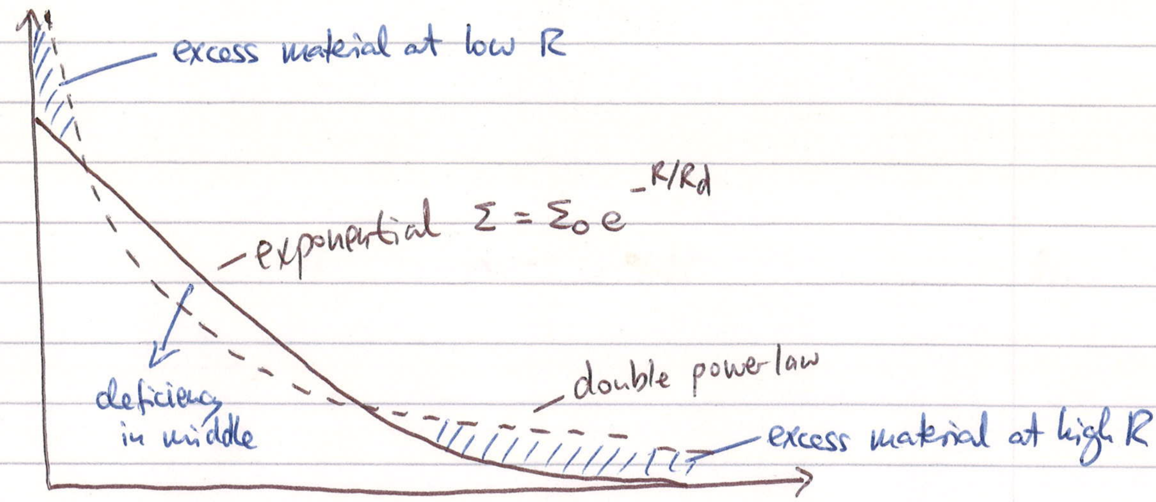
$$= \frac{M_d}{2\pi R} \frac{\mu J_0}{(J_0 + J)^2} V_c$$

Now substitute $R_s \equiv J_0 / V_c$ and use $J = V_c R$:

$$\Sigma(R) = \frac{M_d \mu R_s V_c^2}{2\pi R (R_s V_c + R V_c)^2} = \frac{M_d \mu R_s V_c^2}{2\pi R \cdot V_c^2 R_s^2 (1 + R/R_s)^2}$$

$$= \frac{M_d \mu}{2\pi R_s^2} \left(\frac{R}{R_s}\right)^{-1} \left(1 + \frac{R}{R_s}\right)^{-2}$$

Note that this is not exactly an exponential profile! It is a (double) power-law, which rises faster at low R and declines slower at large R than an exponential profile:



$\Rightarrow$ Simple 'relic angular momentum' model does not explain exponential surface density profiles!

Possible solutions:

(i) Transform excess material near centre (low J) into bulge (v.d. Bosch, 2001) - but not all disks have bulges

(ii) Disks form from preferentially high-J material
$\rightarrow$ e.g. feedback removal of low-J gas (Dutton & v.d. Bosch, 2009)

(iii) Strong J-conservation does not apply: Angular momentum redistributed in disk (e.g. through gravitational torques or viscosity)
$\rightarrow$ discussed below

Aside: Another angular momentum problem has been seen in hydrodynamical simulations of galaxy formation - gas transfers most of its J to the dark matter and does not form realistic discs (i.e. violation of weak J-conservation). Reason: gas forms stars too quickly and does not form disk.

Now plausibly solved by implementation of strong feedback. (e.g. Crain et al. 2009, McCarthy et al. 2012)

Exponential (stellar) disks from viscosity

Lin & Pringle (1987) have shown that viscosity leads to the formation of exponential stellar disks, provided the viscous timescale t_ν and the star-formation time scale t_* are approximately equal.

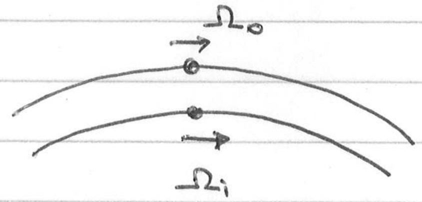
Consider the extreme cases:

$t_\nu \gg t_*$: Viscosity is ineffective, and star formation consumes the gas without significant re-distribution

$t_\nu \ll t_*$: Star formation is ineffective, and viscosity can redistribute gas unhindered.

$\Downarrow$

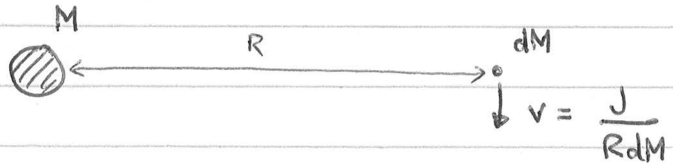
moves mass inward and J outward (Lynden-Bell & Pringle, 1974)



two nearby orbits

$\Omega_i > \Omega_o \Rightarrow$ inner particle slows down, outer particle speeds up

Minimum energy configuration
has (almost) all mass in centre and a (tiny) fraction at large radius



$\Rightarrow$ neither produces an exponential distribution!

But (from numerical calculations), if $t_\nu \approx t_*$, an exponential stellar disk can be produced

$\rightarrow$ BUT: still does not solve central excess...

4.4 Star formation and (local) disk stability

Cooling through electronic excitation / de-excitation becomes ineffective at $T \lesssim 10^4$ K because the gas is almost fully neutral.

What is the Jeans mass for this gas?

$$M_J = \frac{4}{3}\pi \rho \left(\lambda/2\right)^3$$

$$= \frac{\pi}{6}\rho\, c_s^3 \left(\frac{\pi}{G\rho}\right)^{3/2}$$

$$= \frac{\pi^{5/2}}{6}\rho^{-1/2}\left(\frac{5 k_B T}{3 m_p G}\right)^{3/2}$$

$$= \left(\frac{\pi^5 \cdot 125}{36 \cdot 27}\right)^{1/2} \cdot \left(\frac{k_B}{m_p G}\right)^{3/2} \cdot T^{3/2} \cdot \rho^{-1/2}$$

$$\simeq 4.3 \cdot 10^{-9}\, M_\odot \left(\frac{T}{K}\right)^{3/2}\left(\frac{\rho}{kg\, m^{-3}}\right)^{-1/2}$$

[N.B.: lowest for highest density]

For $T = 10^4$ K and $n_H \approx 100\ cm^{-3}$ ($\hat{=}\ \rho \approx 10^{-19}\ kg/m^3$)

we therefore find that $M_J \approx 10^7\, M_\odot \gg M_*$

⇒ star formation requires further cooling and condensation.

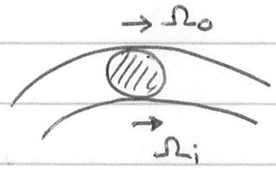
The dominant coolant to reach these low temperatures is molecular hydrogen (see e.g. Abel et al., 1997), which can be excited through vibration and rotation due to its non-spherical shape.

→ Supported by observations: Star formation in the Milky Way and other nearby galaxies restricted to 'Giant Molecular Clouds' (GMCs).

Observations also show that GMCs have higher internal pressure than the surrounding ISM, which indicates that they are gravitationally self-bound (e.g. Larson, 1981).

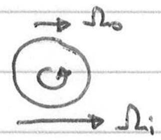
$\Rightarrow$ GMC formation requires gravitational collapse

- Added complication: Disk rotation $\rightarrow$ can stabilise perturbations in addition to pressure



Unless $\Omega_i = \Omega_o$ (i.e., $V_c \propto R^{-1}$), the perturbation will rotate around its own centre

$\Rightarrow$ centrifugal & Coriolis forces which oppose gravity!



- Criterion for gravitational collapse in rotating disk:

  Toomre parameter $Q \equiv \frac{c_s \kappa}{\pi G \Sigma} < 1$ (see Toomre, 1964)

Simplified derivation following Schaye (2004):

Consider the three relevant timescales

t_{ff} - free-fall time-scale

$$\approx \sqrt{\frac{\lambda}{G\Sigma}}$$

t_{sc} - sound crossing

$$\approx c_s^{-1} \lambda$$

t_{rot} - rotation time-scale

$$\approx \frac{2\pi}{\kappa} \quad \text{where } \kappa^2 = 2\left(\frac{V_c^2}{R^2} + \frac{V_c}{R}\frac{dV_c}{dR}\right)^{1/2}$$

(epicyclic frequency, see Binney & Tremaine)

Collapse requires $t_{ff} < t_{sc}, t_{rot}$

- The first inequality equates to

$$t_{ff} < t_{sc} \Rightarrow t_{ff}^2 < t_{sc}^2 \Rightarrow \frac{\lambda}{G\Sigma} < \frac{\lambda^2}{c_s^2} \Rightarrow \underline{\lambda > \frac{c_s^2}{G\Sigma}}$$

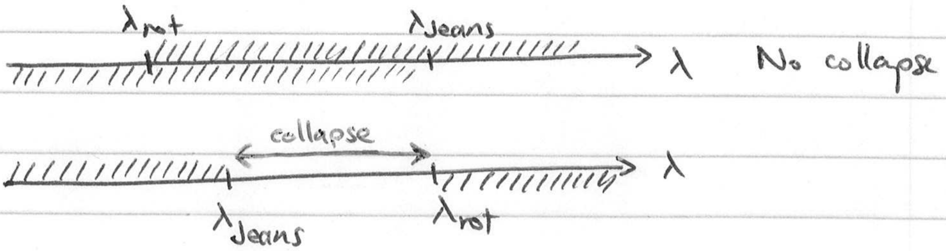
(analogous to Jeans criterion)

Note: pressure prevents collapse for small λ, rotation for large λ!

- The second inequality implies

$$t_{ff}^2 < t_{rot}^2 \Rightarrow \underline{\frac{\lambda}{G\Sigma} < \frac{4\pi^2}{\kappa^2}}$$

$\Rightarrow$ Collapse is only possible if $\lambda_{Jeans} = \frac{c_s^2}{G\Sigma} < \lambda_{rot} = \frac{4\pi^2 G\Sigma}{\kappa^2}$



$\Rightarrow$ Limit when $\lambda_{rot} = \lambda_{Jeans} \Rightarrow c_s^2 \kappa^2 = (2\pi G\Sigma)^2$

$$\Rightarrow Q = \frac{c_s \kappa}{2\pi G\Sigma} = 1 \text{ at limit}$$

off by factor 2 ...

<u>Aside</u>: More rigorous derivation uses perturbation of fluid equations in cylindrical polars to first order, and assuming that spiral arms are tightly wound $\rightarrow$ axial symmetry. See Mo, v.d. Bosch & White (Chapter 11.5.1) for full details.

Resulting dispersion relation:

$$\omega^2 = \kappa^2 - 2\pi G \Sigma_0 |k| + k^2 c_s^2 \qquad (k \equiv \tfrac{2\pi}{\lambda})$$

"unperturbed"

As before, instability requires $\omega^2 < 0$

$\Rightarrow$ At limit $\omega^2 = 0$ and therefore

$$k^2 - \frac{2\pi G \Sigma_0 k}{c_s^2} + \frac{\kappa^2}{c_s^2} = 0$$

$$\Rightarrow k = \frac{\pi G \Sigma_0}{c_s^2} \pm \underbrace{\sqrt{\frac{(\pi G \Sigma_0)^2}{c_s^4} - \frac{\kappa^2}{c_s^2}}}_{=0 \text{ at limit}}$$

$=0$ at limit – instability only for one k!

$$\Rightarrow \lambda_{crit} = \frac{2\pi}{\pi G \Sigma_0} \cdot c_s^2$$

and $\pi G \Sigma_0 = \kappa c_s$ (from radicand condition)

$$\Rightarrow Q \equiv \frac{c_s \kappa}{\pi G \Sigma_0} = 1$$

$$\Rightarrow \lambda_{crit} = \frac{2}{G\Sigma_0} \cdot \frac{\pi^2 G^2 \Sigma_0^2}{\kappa^2} = \frac{2\pi^2 G \Sigma_0}{\kappa^2}$$

What is λ_{crit} for gas with $T \sim 10^4$ K in galactic disk?

$\Sigma \simeq 10\ M_\odot/pc^2$; $\kappa \sim \Omega \sim 10^{-8}\ yr^{-1}$

$\Rightarrow \lambda_{crit} \sim 3 \cdot 10^{18}$ m $\hat{=}$ 1 kpc

$\Rightarrow M_{crit} = \pi \left(\lambda_{crit}/2\right)^2 \Sigma \approx 10^7\ M_\odot \gg M_{GMC}$

and $Q \approx 0.8$, so only perturbations with $\lambda \sim \lambda_{crit}$ are unstable.

But if gas cools first (due to e.g. H_2 formation), c_s is lower by a factor of $\left(T_{warm}/T_{cold}\right)^{1/2} \sim \sqrt{1000} \sim 30$

⇒ Q is very low and perturbations on scales of GMCs can collapse

⇒ suggests that thermal instability (cooling to low T) triggers gravitational collapse and hence star formation.

(Note: Toomre parameter only valid for thin disk – not really true for the warm component, which can therefore be stable. Also note that the short-wavelength limit to instability is set by the Jeans length, and not rotation. See Schaye (2004) for further discussion).

Caveat: GMCs also involve turbulence and magnetic fields
⇒ complicated, and probably reason for low star formation efficiency.