

ASTR 610

Theory of Galaxy Formation

Lecture 8: Non-Linear Collapse & Virialization

FRANK VAN DEN BOSCH
YALE UNIVERSITY, FALL 2020



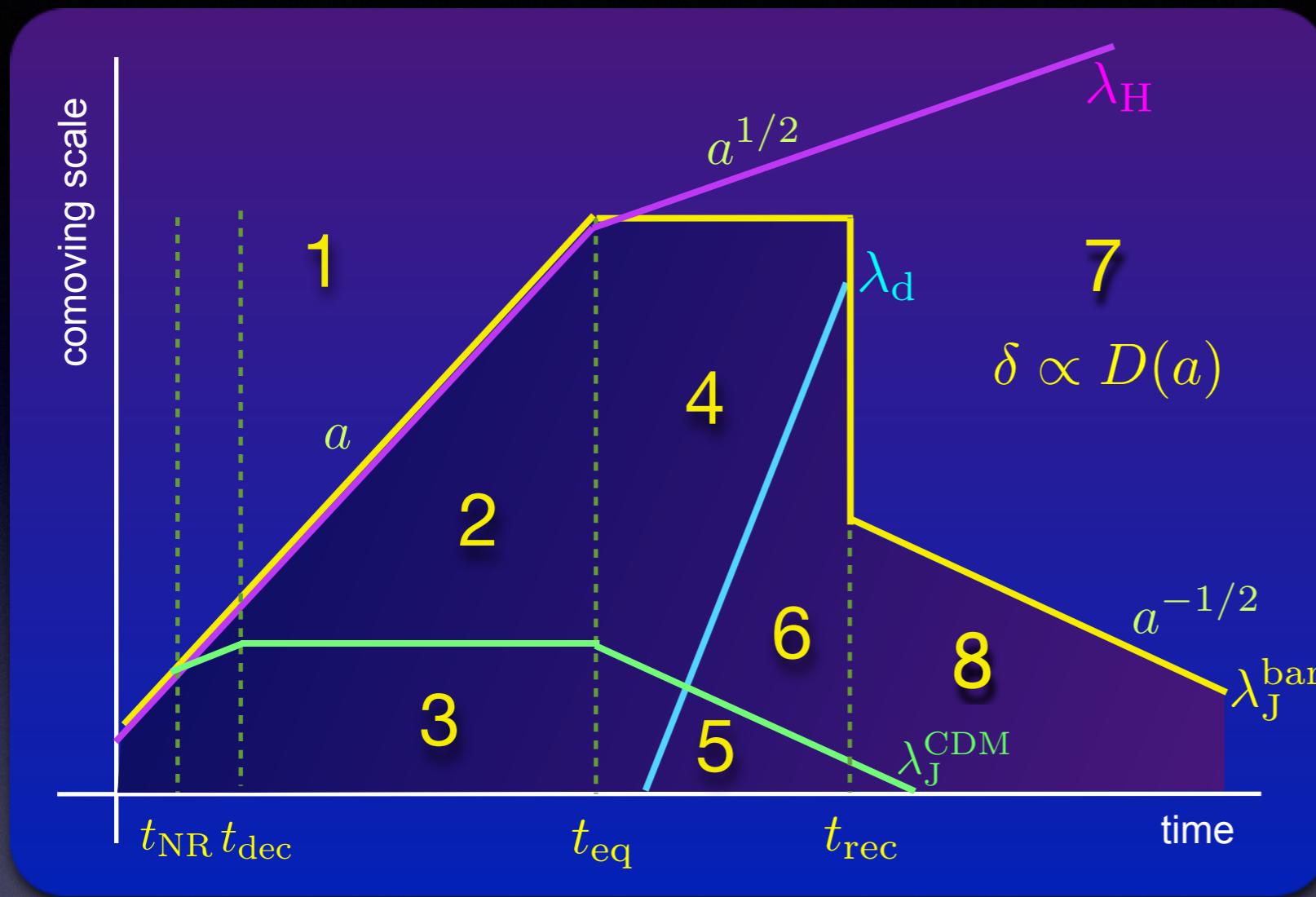
Non-Linear Collapse & Virialization

Having discussed linear theory, we now focus on the non-linear regime. Using simple, analytical models, we will gain insight into how virialized dark matter haloes emerge out of the cosmological density field. We also discuss various relaxation mechanism that bring the collapsed halo into virial equilibrium.

Topics that will be covered include:

- Top-Hat Spherical Collapse
- Secondary Infall Model
- Zel'dovich approximation
- Ellipsoidal Collapse
- Violent Relaxation
- Phase-Mixing

Recap: Linear Perturbation Growth



	Baryons	CDM
1	growth	growth
2	oscillations	stagnation
3	oscillations	free-streaming
4	oscillations	growth
5	Silk-damping	free-streaming
6	Silk-damping	growth
7	growth	growth
8	oscillations	growth

As shown in Lecture 6, after **recombination** the growth of linear perturbations on our scales of interest ($10^6 M_\odot < M < M^{15} M_\odot$) is governed by the **linear growth rate**; $D(a)$

In this linear regime, all modes k evolve similarly and independently: $\delta_{\vec{k}} \propto D(a)$



$$P(k, t) = P_i(k) T^2(k) D^2(t)$$

Once perturbations become of order unity, structure formation becomes **non-linear**....

Non-Linear Evolution

In the **linear** regime ($\delta \ll 1$) we can calculate the evolution of a density field of arbitrary form using **linear** perturbation theory.

In the **non-linear** regime ($\delta > 1$) perturbation theory is no longer valid. Modes start to couple to each other, and one can no longer describe the evolution of the density field with a simple growth rate: in general, no analytic solutions exist...

Because of this mode-coupling, the density field loses its Gaussian properties, i.e., in the **non-linear** regime, we no longer have a Gaussian random field.

Hence, higher-order moments are required to completely specify density field.

How to proceed?

- Oversimplified, but insightful, analytical model (this lecture)
- Higher-order perturbation theory (not covered, see MBW §4.1.7)
- Numerical simulations (lecture 18)
- The Halo Model (lecture 13)

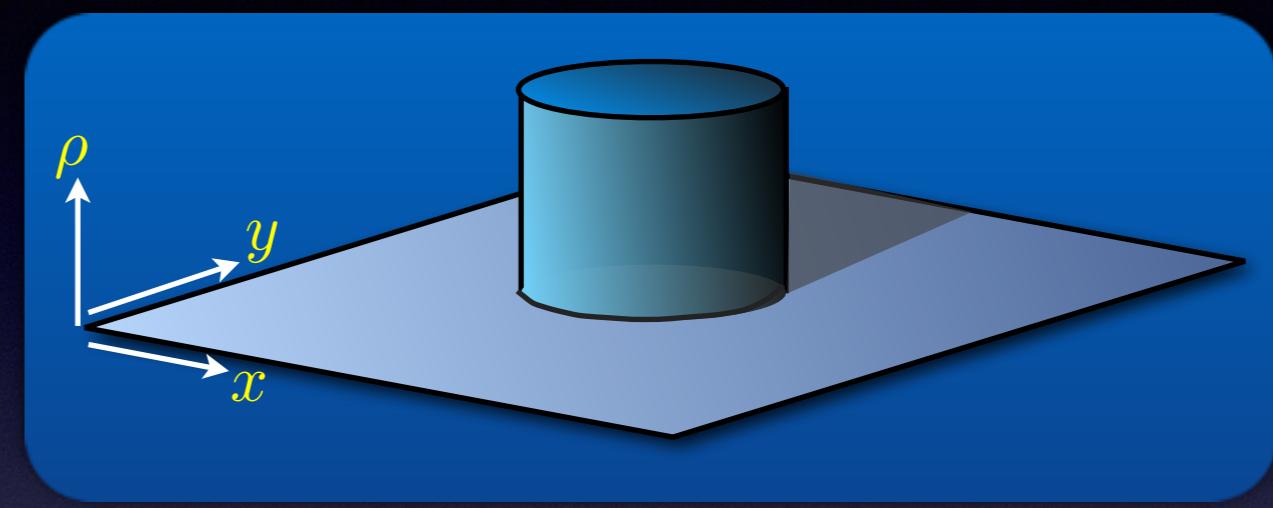
Top-Hat Spherical Collapse

In order to gain insight into the non-linear evolution of density perturbations, we now consider the highly idealized case of Top-Hat Spherical Collapse.

- Universe is homogeneous, except for a single, top-hat, spherical perturbation.
- Universe is in matter-dominated phase, after recombination...
- Collisionless fluid → treatment is only valid for collisionless Dark Matter.
- Einstein-de Sitter (EdS) cosmology



$$\begin{aligned}\Omega(t) &= 1.0 & H(t) \cdot t &= \frac{2}{3} \\ \bar{\rho} &= \frac{1}{6\pi G t^2} & D(a) &\propto t^{2/3}\end{aligned}$$



NOTE:

Although the following treatment is only valid for an EdS cosmology, similar models can be constructed for other cosmologies as well, including Λ CDM (see MBW §5.1.1 + 5.1.2)

Furthermore, since all cosmologies behave similar to EdS at early times, this treatment is always good approximation at high z

Illustration of Spherical Collapse



Onion Model

you can think of overdensity
as consisting of many
individual, thin mass shells

Because of collisionless nature, the shell
crosses itself and starts to oscillate



the evolution of a single mass shell
consisting of **collisionless dark matter**
in a homogeneous universe

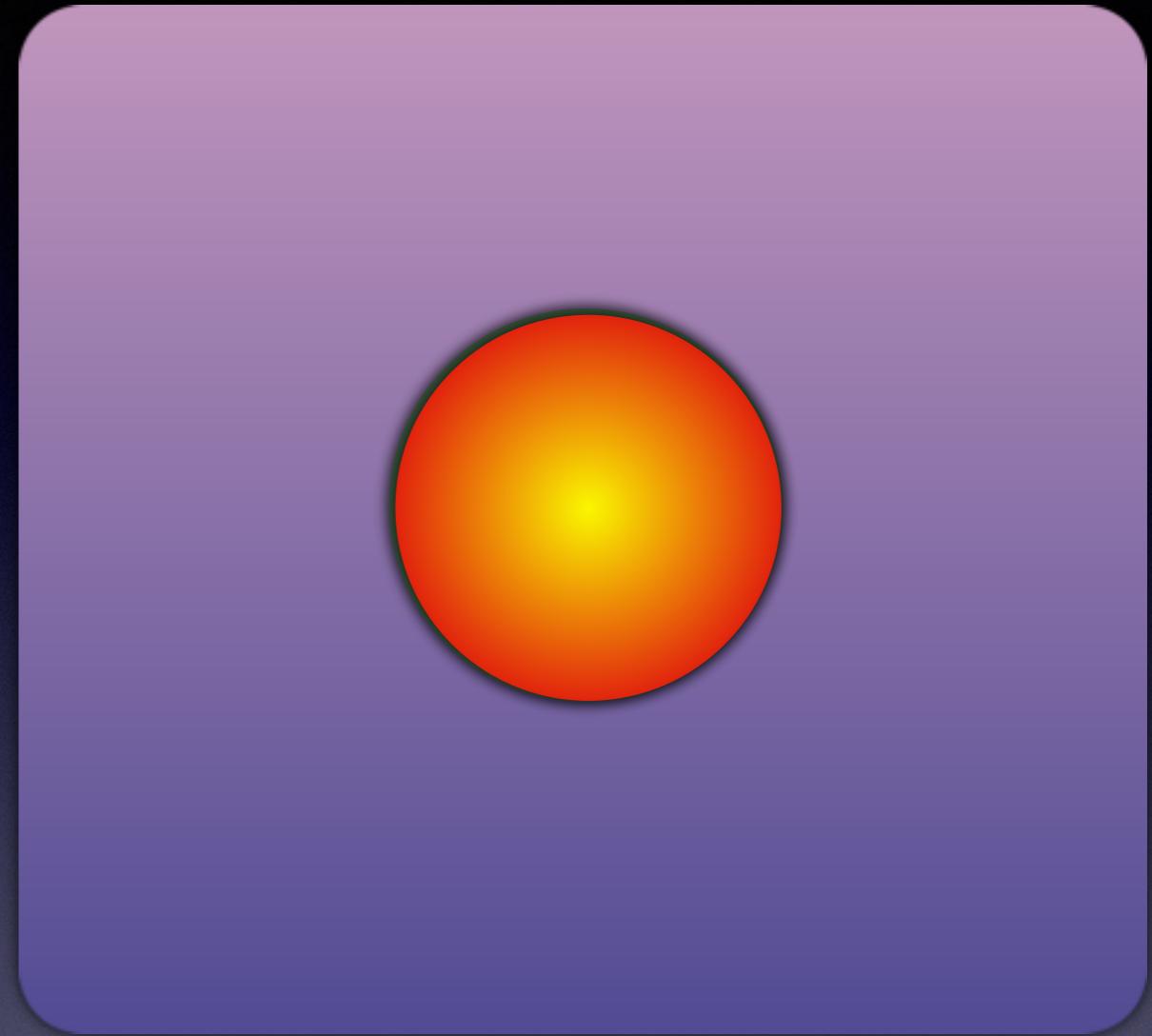
Illustration of Spherical Collapse



Onion Model

you can think of overdensity
as consisting of many
individual, thin mass shells

Because of pressure a shock develops,
which heats the gas and makes it expand



the evolution of a single mass shell
consisting of **baryonic** matter
in a homogeneous universe



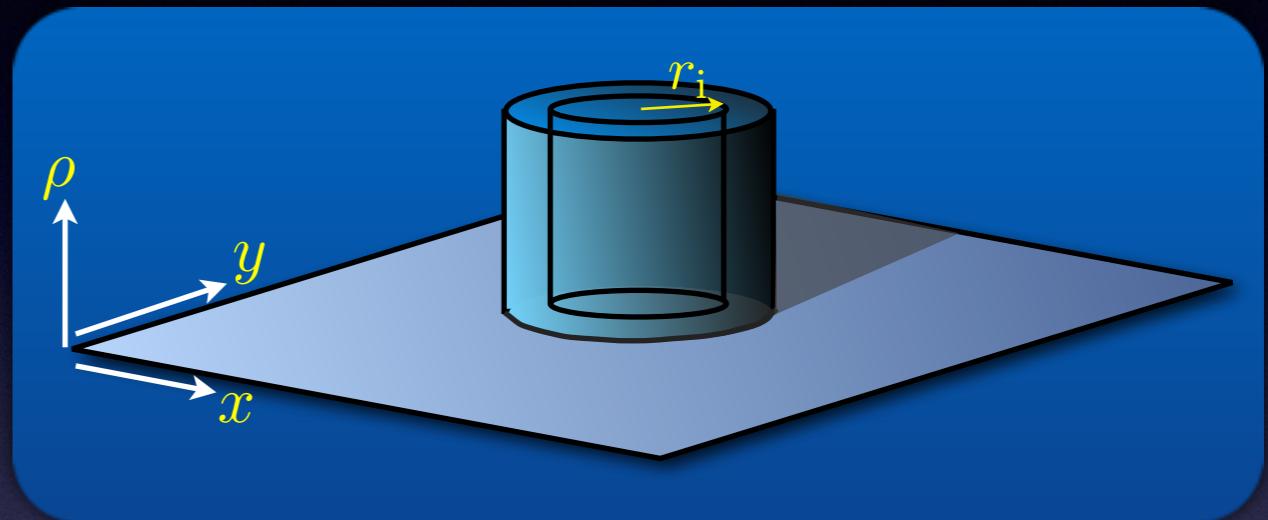
In what follows we focus exclusively on dark matter. We will discuss the evolution
of the baryonic component only at a much later stage (lecture 14+)...

Top-Hat Spherical Collapse

Consider our **spherical top-hat** perturbation: Let r_i denote the radius of some mass shell inside the top-hat at some initial time, t_i , and let δ_i and $\bar{\rho}_i$ denote the top-hat overdensity and the back-ground density at that same time.

The mass enclosed by the shell is

$$\begin{aligned} M(< r) &= \frac{4}{3}\pi r_i^3 \bar{\rho}_i [1 + \delta_i] \\ &= \frac{4}{3}\pi r^3(t) \bar{\rho}(t) [1 + \delta(t)] \end{aligned}$$



where the second equality expresses mass conservation: because of spherical symmetry, the mass inside the shell is conserved, but only up to shell crossing !!!

NOTE: r is the physical radius, not the comoving radius.!!

Newton's first Theorem:

a spherically symmetric matter distribution outside a sphere exerts no force on that sphere



Equation of motion

$$\frac{d^2r}{dt^2} = -\frac{GM}{r^2}$$

the GR equivalent of this is known as Birkhoff's theorem

Top-Hat Spherical Collapse

Integrating the equation of motion once yields

$$\frac{1}{2} \left(\frac{dr}{dt} \right)^2 - \frac{GM}{r} = E$$

where the integration constant E is clearly the specific energy of our shell.

Recall Classical Mechanics: $E < 0$ corresponds to the gravitationally bound case, which for our mass shell implies 'collapse'

- For $E = 0$ the solution to the above equation is simple: $r(t) = \left(\frac{9GM}{2} \right)^{1/3} t^{2/3}$ which shows that $r \propto a$; the mass shell (and hence the top-hat) grows at the same rate as Universe $\rightarrow \delta(t) = \delta_i$ (no growth)
- For $E < 0$ the solution can be written in parametric form:

$$r = A (1 - \cos \theta) \quad \theta \in [0, 2\pi]$$
$$t = B (\theta - \sin \theta)$$

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

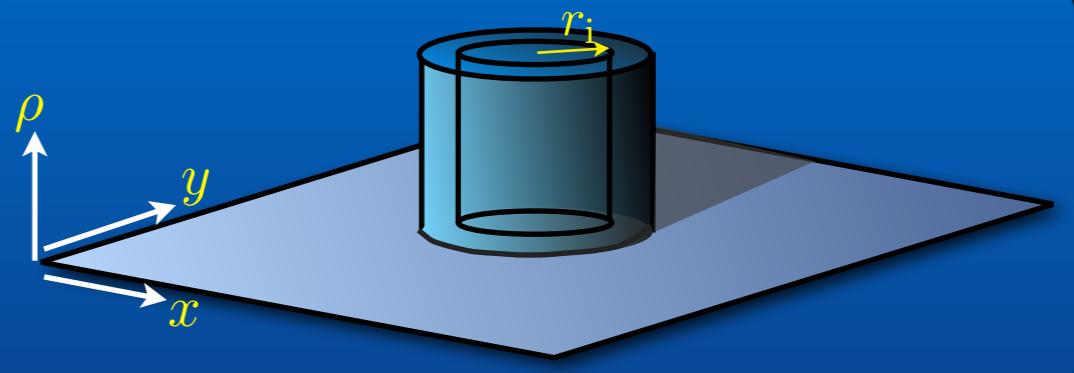
Top-Hat Spherical Collapse

$$r = A(1 - \cos\theta)$$

$$t = B(\theta - \sin\theta)$$

$$A = \frac{GM}{2|E|}$$

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



This solution implies the following evolution for our mass shell:

- shell expands from $r = 0$ at $\theta = 0$ ($t = 0$)
- shell reaches a maximum radius r_{\max} at $\theta = \pi$ ($t = t_{\max} = \pi B$)
- shell collapses back to $r = 0$ at $\theta = 2\pi$ ($t = t_{\text{coll}} = 2t_{\max}$)

The time of maximum size is often called the turn-around time, $t_{\text{ta}} = t_{\max}$, while the time of collapse is also called the virialization time $t_{\text{vir}} = t_{\text{coll}} = 2t_{\text{ta}}$

We will use both notations intermittently...

Top-Hat Spherical Collapse

To get some further insight into the non-linear evolution of (spherical) overdensities, we can use the concept of energy conservation:

Under the approximation that the initial velocity of our mass shell is simply the Hubble flow, we have $v_i = dr_i/dt = d(ax_i)/dt = \dot{a}x_i + a\dot{x}_i \simeq \dot{a}x_i = H_i r_i$ where x is the radius of the mass shell in comoving units.

The initial, specific energy of the mass shell is $E_i = K_i + W_i = \frac{1}{2}H_i^2 r_i^2 - \frac{GM}{r_i}$

Using that $M = \frac{4}{3}\pi r_i^3 \bar{\rho}_i (1 + \delta_i) = \frac{H_i^2 r_i^3}{2G} (1 + \delta_i)$ we obtain $E_i = K_i - K_i(1 + \delta_i)$

→ Collapse requires $E_i < 0$ which thus translates into $\delta_i > 0$

Hence, in an EdS cosmology, all overdensities will (ultimately) collapse.

In a non-EdS cosmology, it is straightforward to show that the above criterion becomes $\delta_i > \Omega_i^{-1} - 1$; overdensities can be prevented from collapse if background density is sufficiently low....Note though, that all cosmologies behave as EdS at early times...

Top-Hat Spherical Collapse

At turn-around ($\theta = \pi, t = t_{\max}$), the shell has zero kinetic energy; $K_{ta} = 0$



$$E_{ta} = W_{ta} = -\frac{GM}{r_{\max}} = -\frac{H_i^2 r_i^3}{2r_{\max}} (1 + \delta_i)$$

We can compare this to our initial energy: $E_i = -K_i \delta_i = -\frac{H_i^2 r_i^2}{2} \delta_i$

Energy conservation then implies that

$$\frac{E_{ta}}{E_i} = 1 = \frac{r_i}{r_{\max}} \frac{1 + \delta_i}{\delta_i}$$



$$\frac{r_{\max}}{r_i} = \frac{1 + \delta_i}{\delta_i} \simeq \delta_i^{-1}$$

Hence, the turn-around radius depends only on the initial overdensity (not on the actual mass enclosed by the shell). Note also that smaller perturbations have larger radii at turn-around (which implies that they turnaround/collapse later...).

Top-Hat Spherical Collapse

Now let us focus on the evolution of the actual overdensity:

The mean density of the top-hat is

$$\rho = \frac{3M}{4\pi r^3} = \frac{3M}{4\pi A^3} (1 - \cos \theta)^{-3}$$

The mean density of the background is $\bar{\rho} = \frac{1}{6\pi G t^2} = \frac{1}{6\pi G B^2} (\theta - \sin \theta)^{-2}$

Hence, the actual overdensity of our spherical top-hat region, according to the spherical collapse (SC) model, which in general will be non-linear, is

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

where we have used that $A^3 = GM B^2$.

Before we examine this SC model in some detail, we first compare it to predictions from linear theory....

Top-Hat Spherical Collapse

For a number of reasons (in particular for use in EPS theory), it is also useful to compare this SC overdensity model to what linear theory predicts for $\delta(t)$.

According to linear theory, perturbation in EdS cosmology evolve as

$$\delta_{\text{lin}} \propto D(a) \propto a \propto t^{2/3}$$

In order to use the correct initial conditions (ICs), we have to use our parametric solution of $r(t)$ in the limit $\theta \ll 1$. Using a Taylor series expansion of $\sin \theta$ and $\cos \theta$ one can show that (see problem set 3):

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

NOTE: this implies that since $\delta(r) = \text{constant}$ inside the top-hat, each mass shell that is part of the top-hat will turn-around (reach maximum expansion) at the same time....

Top-Hat Spherical Collapse

For a number of reasons (in particular for use in EPS theory), it is also useful to compare this SC overdensity model to what linear theory predicts for $\delta(t)$.

According to linear theory, perturbation in EdS cosmology evolve as

$$\delta_{\text{lin}} \propto D(a) \propto a \propto t^{2/3}$$

In order to use the correct initial conditions (ICs), we have to use our parametric solution of $r(t)$ in the limit $\theta \ll 1$. Using a Taylor series expansion of $\sin \theta$ and $\cos \theta$ one can show that (see problem set 3):

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

Combining the above, we have that, according to linear theory:

$$\delta_{\text{lin}} = \delta_i \left(\frac{t}{t_i} \right)^{2/3} = \frac{3}{20} (6\pi)^{2/3} \left(\frac{t}{t_{\max}} \right)^{2/3}$$

Turn Around & Collapse

Spherical Collapse (SC) model:

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

Linear Theory

$$\delta_{\text{lin}} = \delta_i \left(\frac{t}{t_i} \right)^{2/3} = \frac{3}{20} (6\pi)^{2/3} \left(\frac{t}{t_{\max}} \right)^{2/3}$$

Turn-Around: ($t_{\text{ta}} = t_{\max}; \theta = \pi$)

SC model: $1 + \delta(t_{\text{ta}}) = \frac{9\pi^2}{16} \simeq 5.55$

linear theory: $\delta_{\text{lin}}(t_{\text{ta}}) = \frac{3}{20} (6\pi)^{2/3} \simeq 1.062$

Collapse (shell crossing) ($t_{\text{coll}} = 2t_{\text{ta}}$)

SC model: $\delta(t_{\text{coll}}) = \infty$



linear theory: $\delta(t_{\text{coll}}) = \frac{3}{20} (12\pi)^{2/3} = \frac{3}{5} \left(\frac{3\pi}{2} \right)^{2/3} \simeq 1.686$

Critical Overdensity for Collapse

According to linear theory, regions in the linearly extrapolated density field with $\delta_{\text{lin}} \geq 1.686$ should have collapsed. This is often called the “critical overdensity for collapse”, and denoted by the symbol δ_c . As we will see, δ_c plays an important role in (extended) Press-Schechter theory!

Note, though, that our derivation is only valid for an EdS cosmology. Fortunately, similar calculations also exist for non-EdS cosmologies (see MBW §5.1.1 & 5.1.2)

The non-EdS results are well approximated by:

$$\delta_c = \delta_{\text{lin}}(t_{\text{coll}}) = \frac{3}{5} \left(\frac{3\pi}{2} \right)^{2/3} [\Omega_m(t_{\text{coll}})]^{0.0185} \simeq 1.686 [\Omega_m(t_{\text{coll}})]^{0.0185} \quad (\Omega_\Lambda = 0.0)$$

$$\delta_c = \delta_{\text{lin}}(t_{\text{coll}}) = \frac{3}{5} \left(\frac{3\pi}{2} \right)^{2/3} [\Omega_m(t_{\text{coll}})]^{0.0055} \simeq 1.686 [\Omega_m(t_{\text{coll}})]^{0.0055} \quad (\Omega_\Lambda \neq 0.0)$$

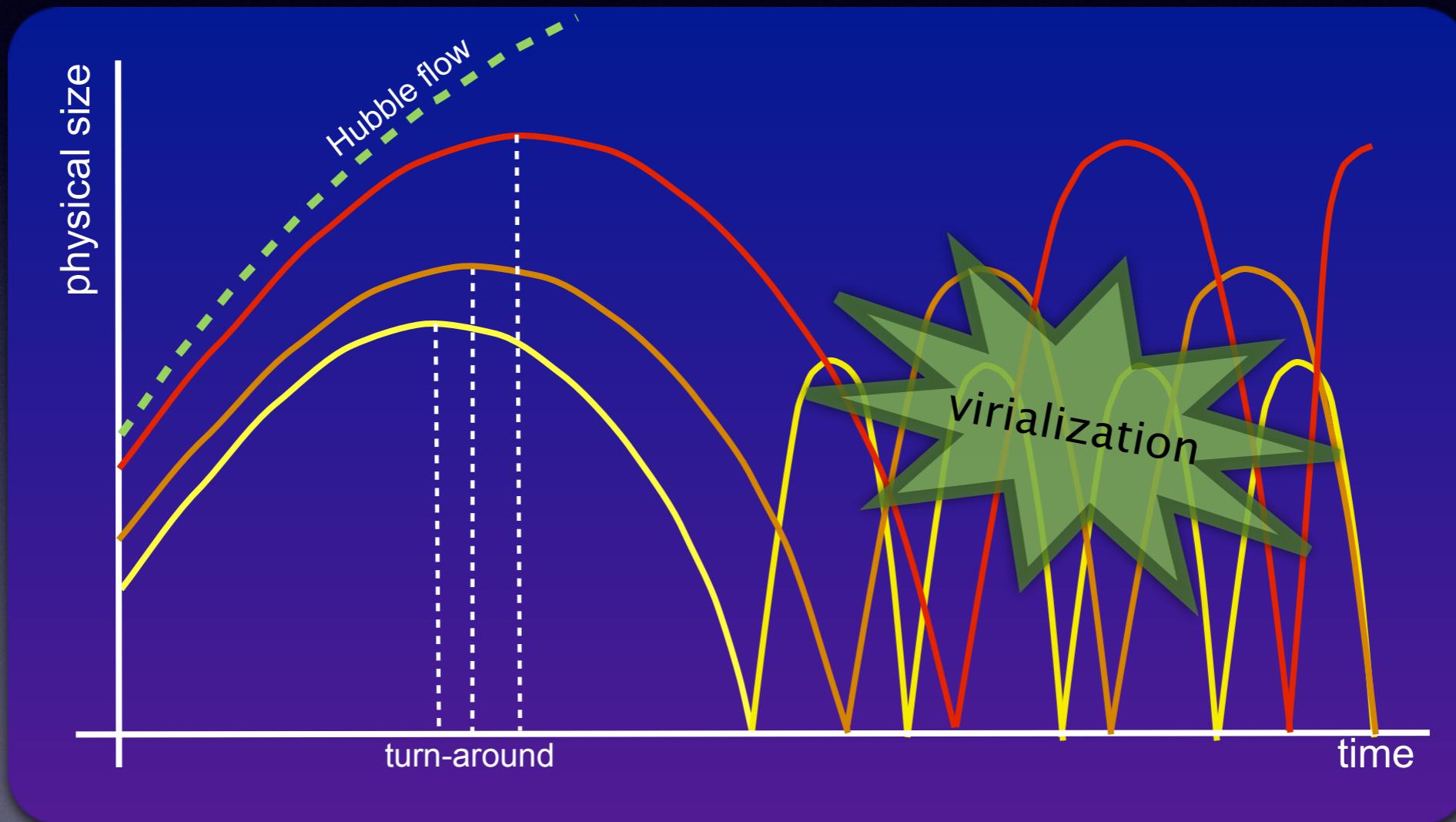
Both approximations are accurate to better than 1%, and show that δ_c has only a very weak dependence on cosmology: to good approximation one can simply adopt $\delta_c \simeq 1.686$

Shell Crossing & Virialization

The SC model discussed above is only valid up to the point of shell crossing.

Afterall, after shell crossing $M(r)$ is no longer a conserved quantity!

According to the SC model, $\delta(t_{\text{coll}}) = \infty$, which would result in the formation of a black hole. However, in reality, the collapse is never perfectly spherical.



Individual oscillating shells interact gravitationally, exchanging energy (virializing). This process, to be described in more detail below, results in a virialized dark matter halo

Density of a Collapsed Dark Matter Halo

Virialization means that the system relaxes towards virial equilibrium:

We can use the virial theorem to make a simple estimate of the final density of our collapsed & virialized dark matter halo:

Virial Equilibrium: $2 K_f + W_f = 0$

Energy conservation: $E_f = K_f + W_f = E_i = E_{ta}$

$$\left. \begin{aligned} E_{ta} &= W_{ta} = -\frac{GM}{r_{ta}} \\ E_f &= W_f/2 = -\frac{GM}{2r_{vir}} \end{aligned} \right\} \rightarrow r_{vir} = r_{ta}/2$$



A mass shell is expected to virialize at half its turn-around radius.

Hence, after virialization, the average density of the material enclosed by the mass shell is 8 times denser than at turn-around....

Density of a Collapsed Dark Matter Halo

We now compute the average overdensity of a virialized dark matter halo:

$$1 + \Delta_{\text{vir}} \equiv 1 + \delta(t_{\text{coll}}) = \frac{\rho(t_{\text{coll}})}{\bar{\rho}(t_{\text{coll}})}$$

NOTE: for consistency with many textbooks and journal articles, we use the symbol Δ_{vir} , rather than δ_{vir} to indicate the virialized overdensity....

Using that $\bar{\rho} \propto a^{-3} \propto t^{-2}$ (EdS), and that $t_{\text{coll}} = 2t_{\text{ta}}$ we have that



$$1 + \Delta_{\text{vir}} = \frac{8 \rho_{\text{ta}}}{\bar{\rho}(t_{\text{ta}})/4} = 32 (1 + \delta_{\text{ta}}) = 18\pi^2 \simeq 178$$

For non-EdS cosmologies, the virial overdensities are well approximated by

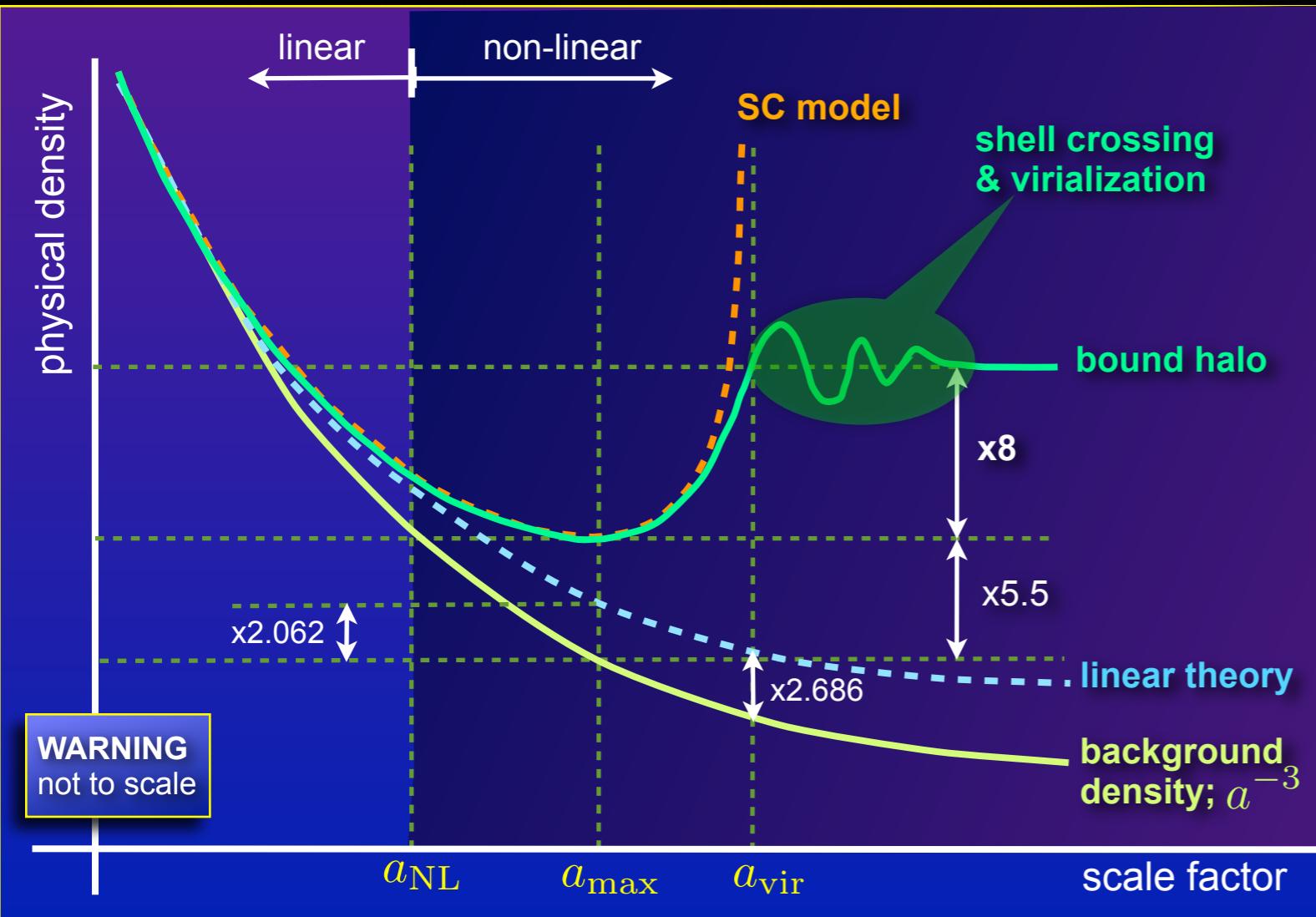
$$\Delta_{\text{vir}} \approx (18\pi^2 + 60x - 32x^2)/\Omega_m(t_{\text{vir}}) \quad (\Omega_\Lambda = 0)$$

$$\Delta_{\text{vir}} \approx (18\pi^2 + 82x - 39x^2)/\Omega_m(t_{\text{vir}}) \quad (\Omega_\Lambda \neq 0)$$

(Bryan & Norman 1998)

Here $x = \Omega_m(t_{\text{vir}}) - 1$. These equations are often used to 'define' dark matter haloes in N-body simulations or in analytical models....

The Spherical Collapse (SC) Model



$\delta = \rho/\bar{\rho} - 1$	turn-around	collapse
SC model	4.55	∞
linear model	1.062	1.686

Although the SC model becomes inaccurate (brakes down) shortly after turn-around it is still a useful model to identify important epochs in the linearly evolved density field...

The linearly extrapolated density field collapses when $\delta_{lin} = \delta_c \simeq 1.686$

Virialized dark matter haloes have an average overdensity of $\Delta_{vir} \simeq 178$

$$\Delta_{vir} \approx (18\pi^2 + 60x - 32x^2)/\Omega_m(t_{vir}) \quad (\Omega_\Lambda = 0)$$

$$\Delta_{vir} \approx (18\pi^2 + 82x - 39x^2)/\Omega_m(t_{vir}) \quad (\Omega_\Lambda \neq 0)$$

Beyond Top-Hat Spherical Collapse

Thus far we have mainly focussed on the non-linear evolution of a spherical top-hat perturbation, embedded in a homogeneous universe.

As we have seen, a mass shell (made out of collisionless matter) oscillates back and forth between $r=0$ and $r=r_{\max}$.

Furthermore, for the top-hat perturbations, all mass shells turn-around and collapse simultaneously.

We now move to more realistic cases. In particular, we will consider perturbations

- with more realistic, initial density profile (Secondary Infall Models)
- with non-zero angular momentum
- without spherical symmetry (Zel'dovich approximation)

Secondary Infall Models

Consider an initial perturbation with a density profile

$$\delta_i \propto r_i^{-3\varepsilon} \propto M^{-\varepsilon}$$

Unless $\varepsilon = 0$ (which corresponds to a top-hat), mass shells will not all collapse at the same time. For realistic profiles ($\varepsilon > 0$), inner shells collapse earlier. Such models, with extended infall of new shells, are called secondary infall models.

Gunn & Gott (1972) assumed that each oscillation the shell expands back out to a radius that is a fixed, constant fraction of r_{\max}

Since a shell spends most of its time near apocenter, Gunn & Gott postulated that the mass enclosed by a shell is the same as its originally enclosed mass (at $t = t_i$)....

$$\rightarrow \rho_{\text{vir}}(r) = \frac{3M}{4\pi r^3(M)}$$

where $r(M) \propto r_{\text{ta}}(M)$ and M is the mass enclosed by the shell at $t = t_i$

For an EdS cosmology, $r_{\text{ta}} \propto r_i/\delta_i \propto M^{1/3+\varepsilon}$, which implies that

$$\rho_{\text{vir}}(r) \propto r^{-\gamma} \quad \text{with} \quad \gamma = \frac{9\varepsilon}{1+3\varepsilon}$$

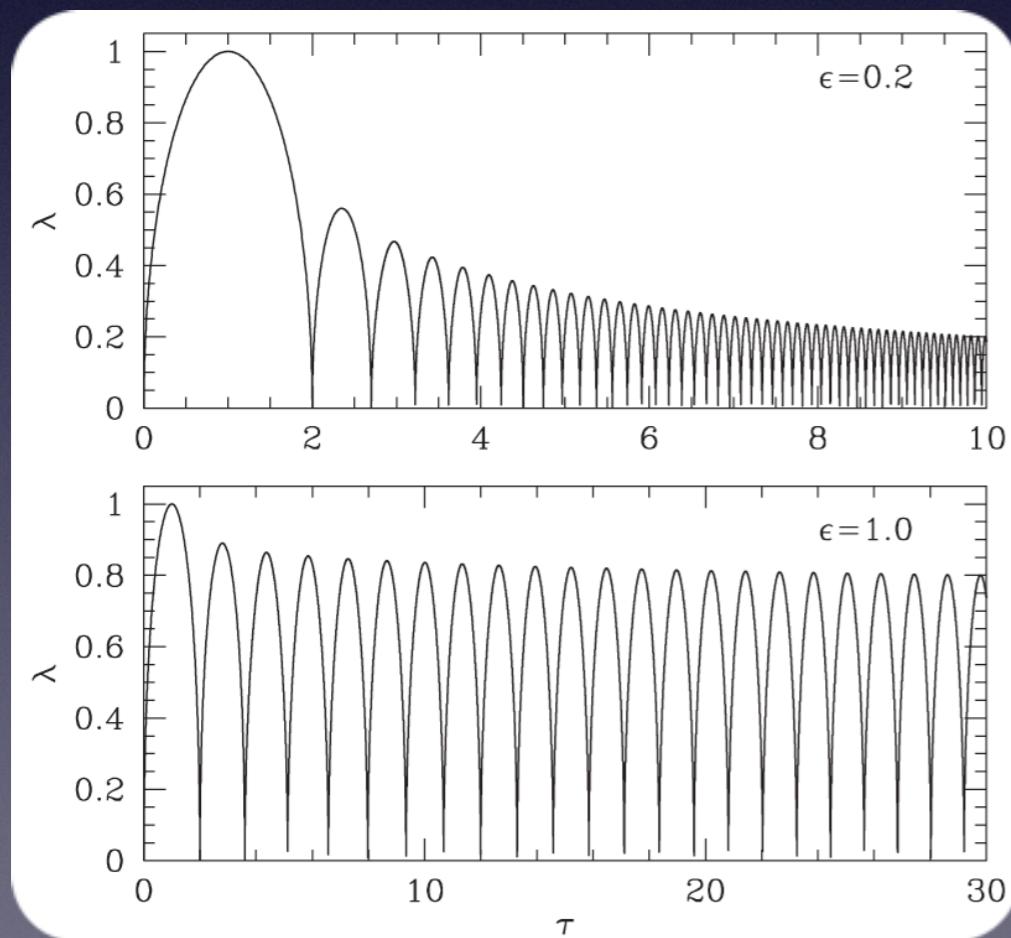
For $\varepsilon = 1$, i.e., point-mass initial conditions, one obtains that $\gamma = -9/4$

Secondary Infall Models

Unfortunately, the Gunn & Gott treatment is not accurate. The total mass within a mass shell at apocenter is not only the mass enclosed initially, but also the mass from shells initially outside it, but with current radii that place it interior....

Hence, to improve upon the Gunn & Gott treatment one needs to solve the equations of motion for all shells simultaneously....this can be done, but only numerically....

Fillmore & Goldreich (1984) and Bertschinger (1985) showed, however, that under certain conditions the problem can be made self-similar, admitting analytical solutions.



Evolution of the scaled radius, $\lambda = r/r_{\max}$, as a function of the scaled time $\tau = t/t_{\max}$, in two self-similar, spherical collapse models (with purely radial orbits).

Note that the shell's apocenter decreases with time, as does its oscillation period. This is due to more and more mass shells having turned-around...

→ each mass shell becomes buried deeper & deeper in the collapsed halo with the passage of time...

Secondary Infall Models

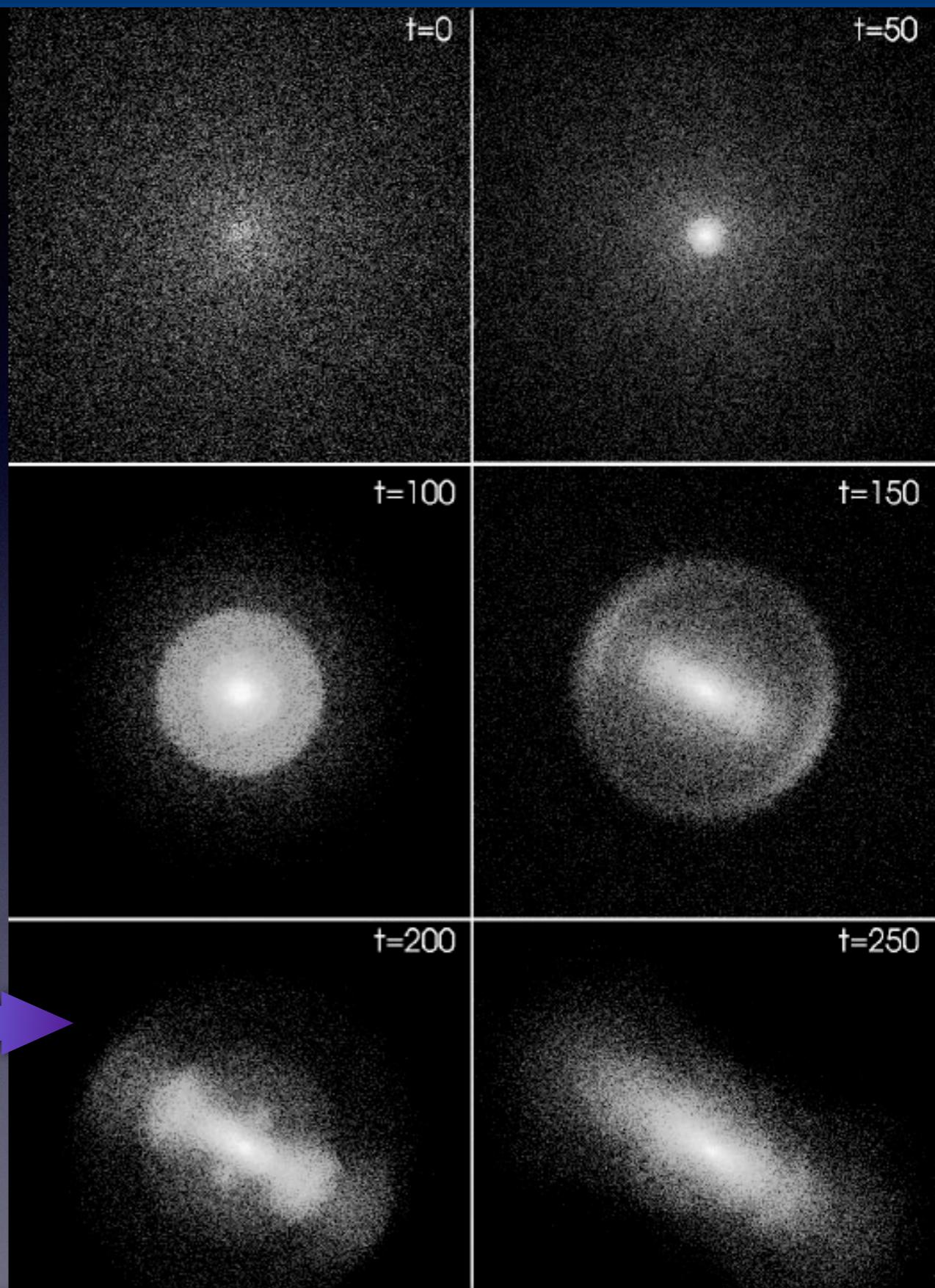
The self-similar solution predicts a final density profile

$$\rho_{\text{vir}}(r) \propto r^{-\gamma} \quad \text{with} \quad \gamma = \begin{cases} 2 & \varepsilon \leq 2/3 \\ \frac{9\varepsilon}{1+3\varepsilon} & \varepsilon > 2/3 \end{cases}$$

(see MBW §5.2.1 for derivation)

A problem with all models discussed thus far is that they only consider purely radial orbits. This has two shortcomings:

- a collisionless system with purely radial orbits suffers from **radial orbit instability** (Antonov 1973) and will rapidly evolve into an elongated bar-shaped configuration (see images).
- in real universe, **tidal torques** from neighboring perturbations will impart **angular momentum** on the particles/mass shells...



Non-Radial Secondary Infall Models

White & Zaritsky (1992) considered a SC model with non-radial particle orbits:

Let specific angular momentum of particle in a mass shell be
where $M_{\max} = M(< r_{\max}) = M(< r_i)$.

$$\mathcal{L} = \mathcal{J} \sqrt{GM_{\max}r_{\max}}$$

The equation of motion of the mass shell is $\frac{dr^2}{dt^2} = -\frac{GM}{r^2} - \frac{\mathcal{L}^2}{r^3}$

Because of the centrifugal force, there is now a centrifugal barrier and mass shells can no longer reach $r=0$. Instead, mass shells oscillate between apocenter and pericenter.

Under the assumption that \mathcal{J} is the same for all mass shells, one can again obtain a self-similar solution:

$$\rho_{\text{vir}}(r) \propto r^{-\gamma} \quad \text{with} \quad \gamma = \frac{9\varepsilon}{1 + 3\varepsilon} \quad (\varepsilon > 0)$$

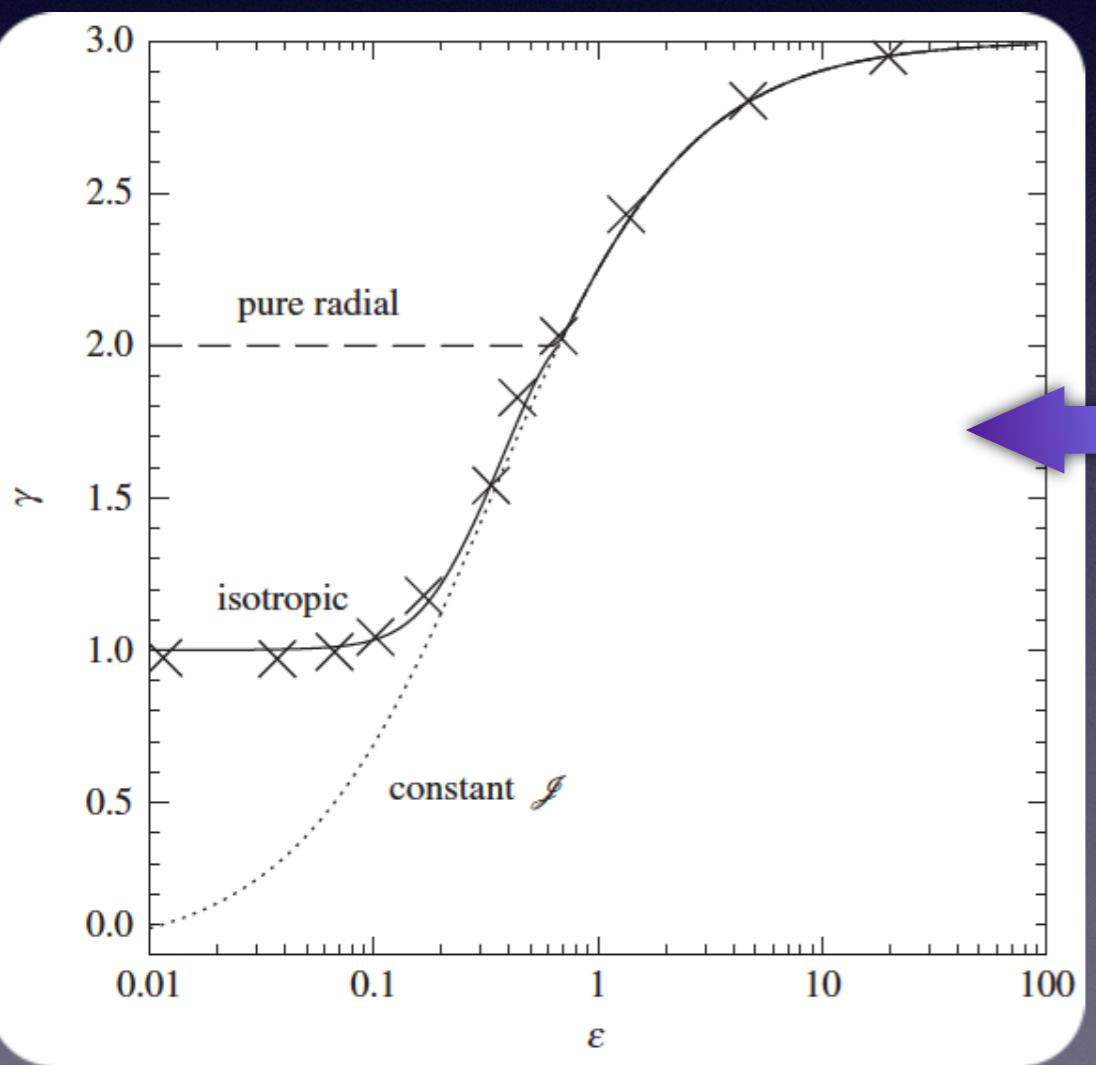
(Nusser 2001)

Note that, somewhat surprisingly, this is exactly the Gunn & Gott solution....

Non-Radial Secondary Infall Models

If \mathcal{J} is different for different mass shells, there are no longer any self-similar, analytical solutions. Rather, one has to resort to numerical simulations (in 1D) to follow the ‘orbits’ of each individual mass shell....

Lu et al. (2006) performed such a calculation in which they assumed that the velocity of each particle is isotropized during the collapse.



Their results suggest that the end-state of collapse is a power-law density profile

$$\rho_{\text{vir}}(r) \propto r^{-\gamma} \quad \text{with} \quad \gamma = \begin{cases} 1 & \varepsilon \leq 1/6 \\ \frac{9\varepsilon}{1+3\varepsilon} & \varepsilon > 1/6 \end{cases}$$

As we will see later, virialized dark matter haloes in realistic (3D) N-body simulations have $\rho \propto r^{-1}$ in their inner regions. The results of Lu et al. suggest that this may well be a manifestation of orbit isotropization during the early (rapid) collapse phase of the halo...

The Zel'dovich Approximation

So far we considered perturbations in Eulerian ('grid') coordinates. Individual overdensities stay at a fixed (comoving) position and grow or decay in amplitude....

We now switch to Lagrangian description, which follows motion of individual particles. This gives insights into dynamics of structure formation process, and, unlike its Eulerian counterpart, remains (fairly) accurate in the mildly non-linear regime...

It is easy to see that Eulerian description brakes down in mildly non-linear regime: Once overdensities ($\delta_i > 0$) reach amplitudes of order unity, the underdensities ($\delta_i < 0$) have grown to $\delta < -1$, which would imply a negative (=unphysical) density...

Zel'dovich (1970) came up with a Lagrangian formalism that is based on the following approximation (known as Zel'dovich Approximation, ZA):

particles continue to move in the direction of their initial displacement

$$\rightarrow \vec{x}(t) = \vec{x}_i - c(t) \cdot \vec{f}(\vec{x}_i)$$

\vec{x}_i initial (Lagrangian), comoving coordinates

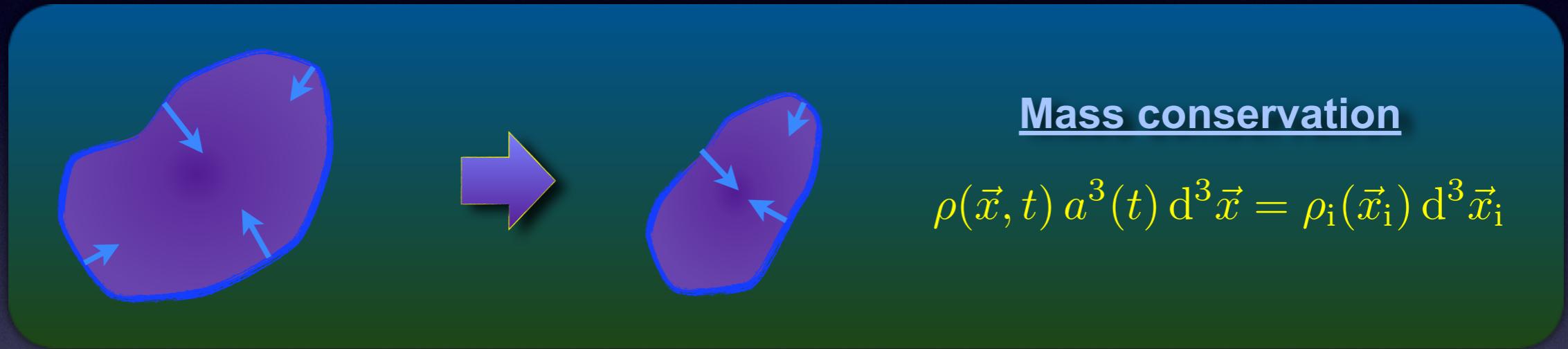
$c(t)$ function of time, to be determined below

$\vec{f}(\vec{x}_i)$ vector function of initial coordinates, specifying direction of velocity

The Zel'dovich Approximation

$$\text{ZA: } \vec{x}(t) = \vec{x}_i - c(t) \cdot \vec{f}(\vec{x}_i)$$

Note: the ZA is exact if perturbation is a 1D sheet in an otherwise homogeneous universe; in that case direction of velocity remains fixed...



Here $a(t)$ is the scale-factor normalized to unity at the initial time t_i : the scaling with $a^3(t)$ is required since \vec{x} are comoving coordinates. The equation of mass conservation is valid (up to orbit crossing) for any geometry; no spherical symmetry is required!!

Using Linear Algebra:

$$\rho(\vec{x}, t) = \rho_i(\vec{x}_i) a^{-3} \left\| \frac{d\vec{x}}{d\vec{x}_i} \right\|^{-1}$$

Here $\|A\| = \det(A) = \prod_i \lambda_i$ with λ_i the eigenvalues of the matrix A

The Zel'dovich Approximation

Using that the tensor $\left(\frac{d\vec{x}}{d\vec{x}_i} \right)_{jk} = \delta_{jk} - c(t) \frac{\partial f_j}{\partial x_k}$ we have that

$$\rho(\vec{x}, t) = \rho_i(\vec{x}_i) a^{-3} \frac{1}{(1 - c\lambda_1)(1 - c\lambda_2)(1 - c\lambda_3)}$$

where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the eigenvalues of the deformation tensor $\partial f_i / \partial x_j$

Using that $\rho_i(\vec{x}_i) = \bar{\rho}_i [1 + \delta_i(\vec{x}_i)] \simeq \bar{\rho}_i$ and that $\bar{\rho}(t) a^3 = \bar{\rho}_i a_i^3$ this yields

$$1 + \delta(\vec{x}, t) = \frac{\rho(\vec{x}, t)}{\bar{\rho}(t)} = \frac{1}{(1 - c\lambda_1)(1 - c\lambda_2)(1 - c\lambda_3)}$$

(recall that $a_i = 1$)

We can gain some useful insight from this equation (using that $c(t) > 0$) :

- if $\lambda_i > 0$ this implies collapse in the direction of the i^{th} eigenvector.
- if $\lambda_i < 0$ this implies expansion in the direction of the i^{th} eigenvector.
- if $c(t) = 1/\lambda_i$ 'shell' crossing happens along the direction of the i^{th} eigenvector.
- as long as $c\lambda_1 \ll 1$ the perturbation is still in the linear regime.

The Zel'dovich Approximation

Linearization of the equation for the density perturbation yields

$$1 + \delta(\vec{x}, t) = \frac{1}{1 - c(\lambda_1 + \lambda_2 + \lambda_3)} \simeq 1 + c(\lambda_1 + \lambda_2 + \lambda_3)$$

Hence, we have that, in the linear regime $\delta(\vec{x}, t) = c(t) \operatorname{Tr}(\partial f_i / \partial x_j) = c(t) \vec{\nabla} \cdot \vec{f}$

If we compare this to the fact that, in the linear regime, $\delta(\vec{x}, t) = D(t) \delta_i(\vec{x}_i)$ we see that $c(t) = D(t)$ and $\vec{\nabla} \cdot \vec{f} = \delta_i$.

Using the Poisson equation, according to which $\delta_i = \nabla^2 \Phi_i / 4\pi G \bar{\rho}_i$ (recall that $a_i = 1$) and the fact that $\nabla^2 \Phi = \vec{\nabla} \cdot \vec{\nabla} \Phi$, we finally see that $\vec{f} = \vec{\nabla} \Phi_i / 4\pi G \bar{\rho}_i$



$$\vec{x}(t) = \vec{x}_i - \frac{D(a)}{4\pi G \bar{\rho}_i} \vec{\nabla} \Phi_i$$

Zel'dovich Approximation



- As we will see in Lecture 18, this ZA is ideally suited to set up ICs for Nbody simulations...
- In Problem Set 3 we will use a different derivation, based on linearized Euler equation....

Zel'dovich Pancakes

The ZA describes the non-linear evolution of density perturbations. It has two important advantages over the spherical collapse model:

- it makes no oversimplified assumptions about geometry
- it remains accurate well into the quasi-linear regime

To understand why the ZA is more accurate in the quasi-linear regime (breaks down at a later stage), have a look at its predicted evolution for an overdensity:

$$1 + \delta(\vec{x}, t) = \frac{\rho(\vec{x}, t)}{\bar{\rho}(t)} = \frac{1}{(1 - c\lambda_1)(1 - c\lambda_2)(1 - c\lambda_3)}$$

It is clear from this equation that collapse happens first along the axis associated with the first (largest) eigenvalue, $\rightarrow \lambda_1$ gravity accentuates asphericity!

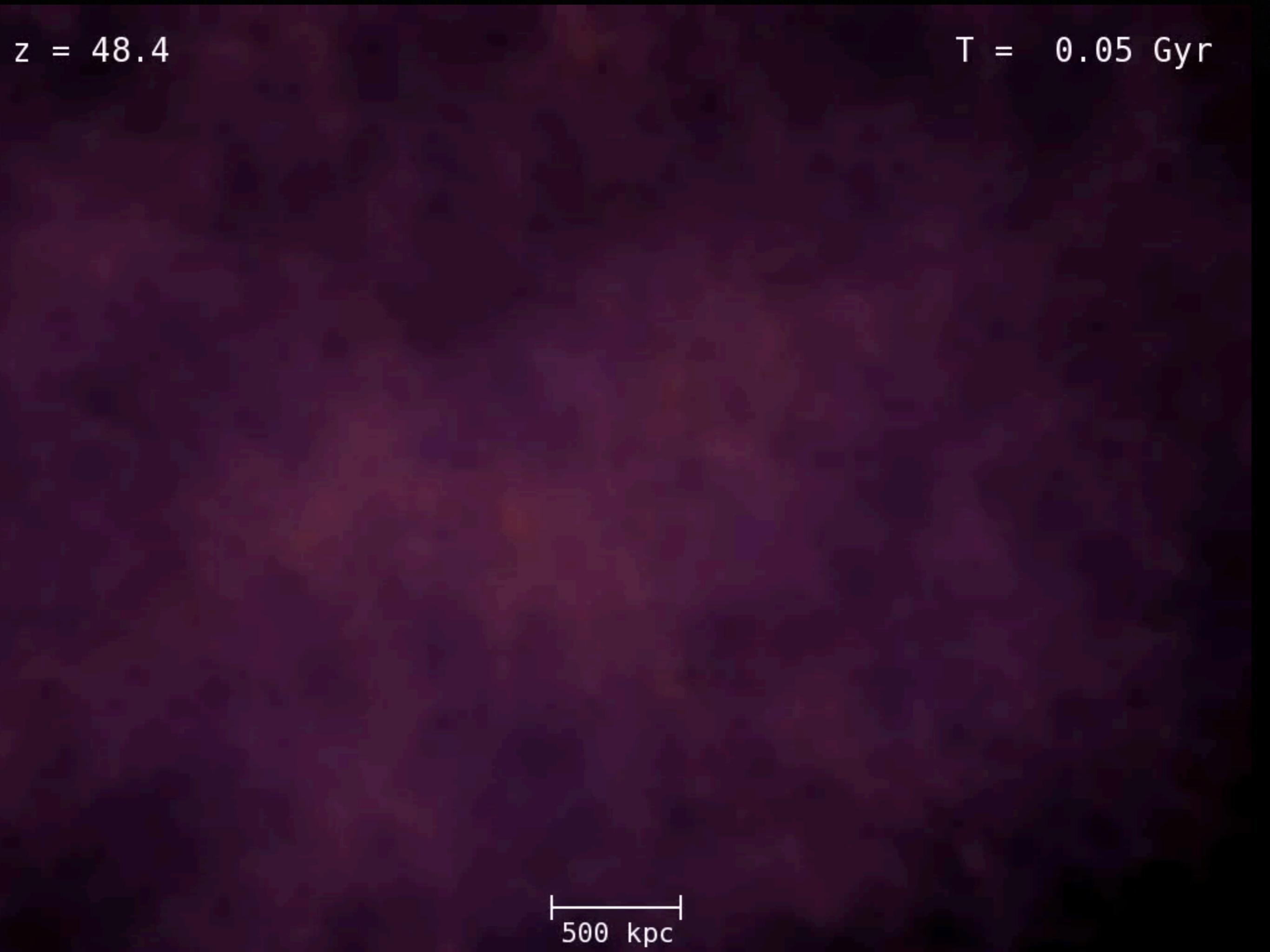
Hence, collapse leads to flattened structures, called (Zel'dovich) pancakes. The ZA approximation is so accurate simply because, as mentioned above, it becomes exact in the limit of planar perturbations...

Because ZA is so accurate, it is often used in setting up the initial conditions for N-body simulations.



$z = 48.4$

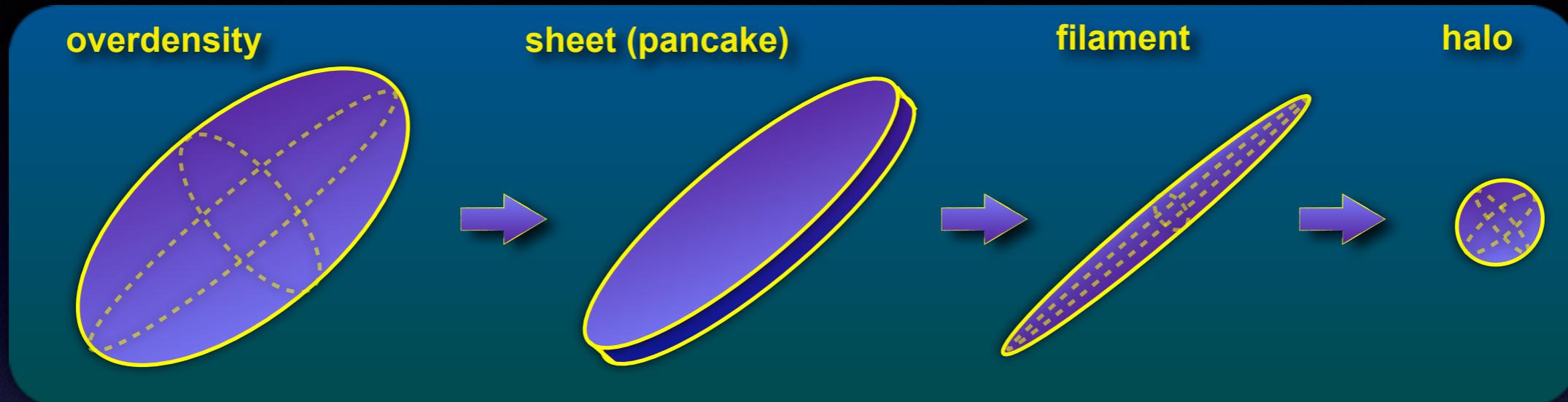
$T = 0.05 \text{ Gyr}$



500 kpc

Ellipsoidal Collapse

As is evident from the ZA, in general density perturbations will collapse according to:



For a uniform, ellipsoidal overdensity in homogeneous universe (ellipsoidal top-hat) one can obtain analytical approximations for time evolution of its 3 principal axes (see MBW §5.3).

This can be used to compute the critical overdensity for collapse (of the longest axis = 'halo formation') in linear theory. The result can be obtained by solving

$$\frac{\delta_{\text{ec}}}{\delta_{\text{sc}}} \approx 1 + 0.47 \left[5(e^2 \pm p^2) \frac{\delta_{\text{ec}}^2}{\delta_{\text{sc}}^2} \right]^{0.615}$$

Sheth, Mo & Tormen (2001)

Here $\delta_{\text{ec}} = \delta_{\text{ec}}(e, p)$ is the critical overdensity for ellipsoidal collapse, $\delta_{\text{sc}} = \delta_c \simeq 1.686$ is the critical overdensity for spherical collapse, and the plus (minus) sign is used if p is negative (positive)....

continued on next page...

Ellipsoidal Collapse

$$\frac{\delta_{\text{ec}}}{\delta_{\text{sc}}} \approx 1 + 0.47 \left[5(e^2 \pm p^2) \frac{\delta_{\text{ec}}^2}{\delta_{\text{sc}}^2} \right]^{0.615}$$

Ellipsoidal collapse

The parameters **e** and **p** characterize the asymmetry of the initial tidal field:

$$e \equiv \frac{\lambda_1 - \lambda_3}{2(\lambda_1 + \lambda_2 + \lambda_3)} \quad p \equiv \frac{\lambda_1 + \lambda_3 - 2\lambda_2}{2(\lambda_1 + \lambda_2 + \lambda_3)}$$

Note that for a spherical system $\lambda_1 = \lambda_2 = \lambda_3 \rightarrow e = p = 0 \rightarrow \delta_{\text{ec}} = \delta_{\text{sc}} \simeq 1.686$

In general, however, $\lambda_1 > \lambda_2 > \lambda_3$ which results in $\delta_{\text{ec}} > \delta_{\text{sc}}$, which implies that structures collapse later under ellipsoidal collapse conditions (more realistic) than under spherical collapse conditions.

Note, though, that this depends on how 'collapse' of ellipsoid is defined:
Here we associated collapse with that of the **longest** axis. If using collapse along the **shortest** axis instead, one finds the opposite: ellipsoidal structures collapse earlier than spherical ones....

As a final remark, as we will see later, less massive structures are more strongly influenced by tides and therefore more ellipsoidal...This has important implications....

Relaxation & Virialization

More details
in MBW §5.4

Relaxation: the process by which a physical system acquires equilibrium or returns to equilibrium after a disturbance. Often, but not always, relaxation erases the system's "knowledge" of its initial conditions.

Virialization: the process by which a physical system settles in virial equilibrium

Virial Equilibrium: A system is said to be in virial equilibrium if

$$2K + W + \Sigma = 0$$

Often, Σ can be ignored, in which case virial equilibrium implies that $E = -K = W/2$

K = kinetic energy
 W = potential energy
 Σ = work done by surface pressure

Two-body relaxation time: the time required for a particle to change its kinetic energy by about its initial amount due to two-body interactions

As you learn in Galactic Dynamics, the two-body relaxation time,

$$t_{\text{relax}} \simeq \frac{N}{10 \ln N} t_{\text{cross}}$$

Here N is the number of particles and $t_{\text{cross}} \sim R/v$ is the system's crossing time.

For almost all collisionless systems of interest to us (galaxies, dark matter haloes) it is easy to show that $t_{\text{relax}} \gg t_{\text{Hubble}} \simeq 1/H_0$

Relaxation & Virialization

More details
in MBW §5.5

PUZZLE: if galaxies and haloes have two-body relaxations times that are orders of magnitude larger than the Hubble time, how can galaxies (and haloes) appear relaxed?



Collisionless systems such as galaxies and dark matter haloes do not relax via two-body interactions, but rather by a combination of four other mechanisms:

Phase-mixing

the spreading of neighboring points in phase-space due to the difference in frequencies between neighboring orbits

Chaotic mixing

the spreading of neighboring points in phase-space due to the chaotic nature of their orbits

Violent Relaxation

the change in energy of individual particles due to changes in the overall potential

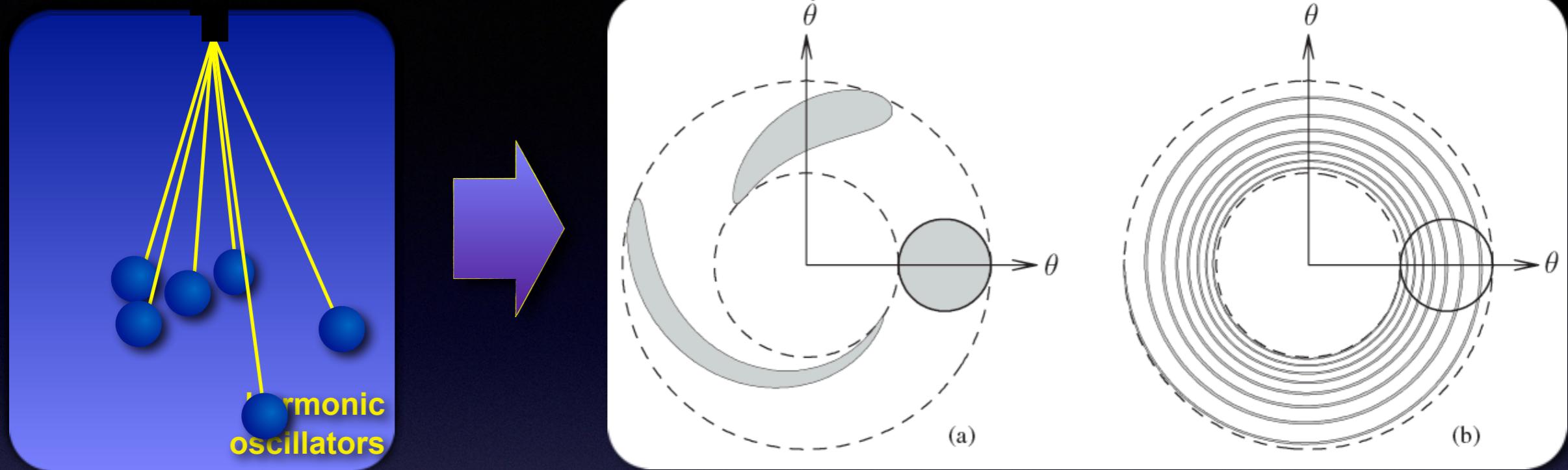
Landau damping

the damping and decay of perturbations due to decoherence between particles and waves (recall free streaming)

In what follows, we briefly discuss each of these in turn. As we will see violent relaxation and Landau damping are basically specific examples of phase mixing....

Phase Mixing

More details
in MBW §5.5



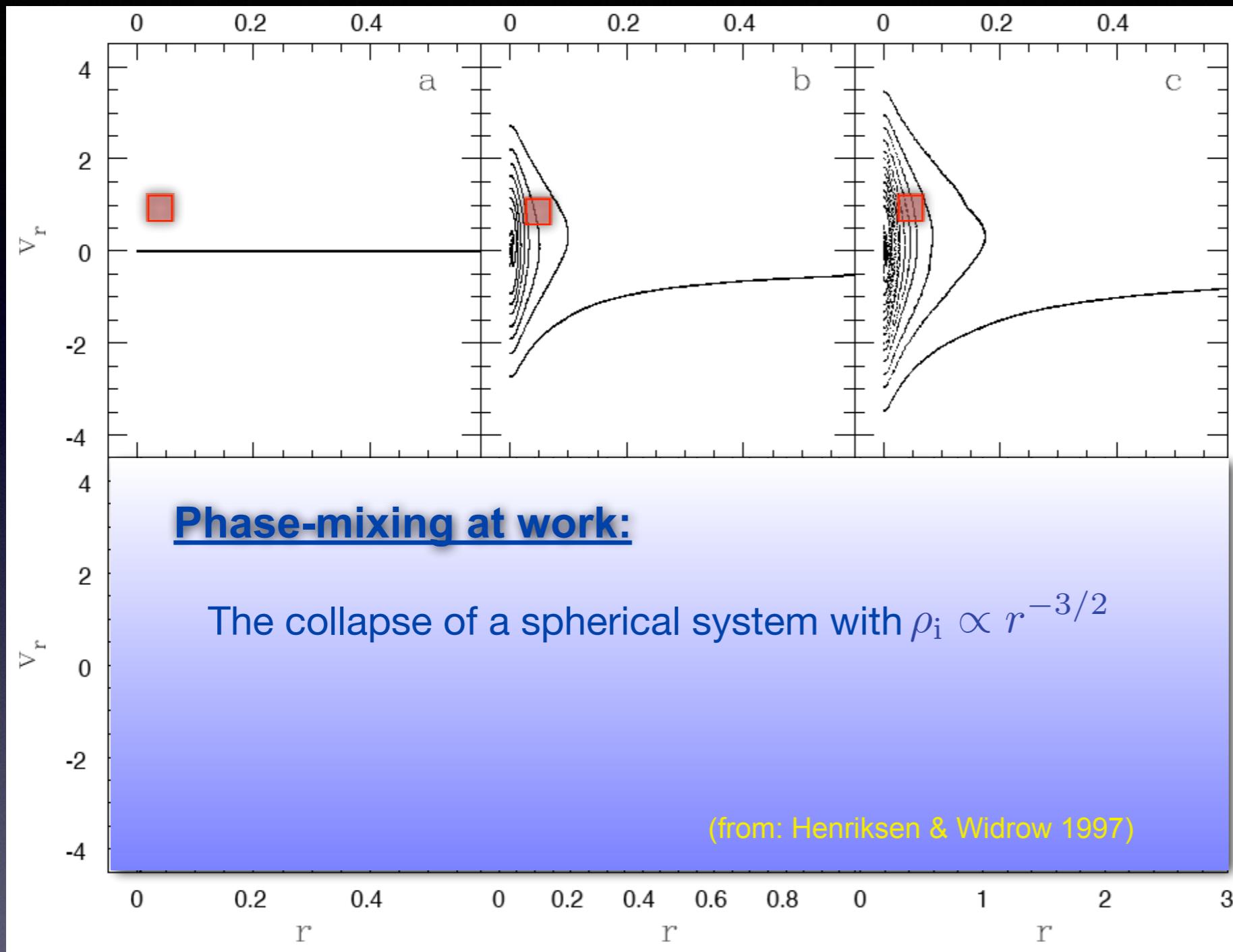
Consider a large number of harmonic oscillators, all with slightly different frequencies (i.e., with slightly different sling-lengths). If they are close to each other initially, they will, over time, **phase-mix** (the overall system appears more relaxed).

Let ϕ_i and ω_i be the phase and frequency of oscillator i , then oscillators i and j separate at a rate $(\Delta\phi)_{ij}(t) = 2\pi(\Delta\omega)_{ij}t$: \rightarrow phase mixing scales linearly with time.

According to the collisionless Boltzmann equation, the (fine-grained) DF $f(\vec{x}, \vec{v})$ remains constant. However, the coarse-grained DF, $f_c(\vec{x}, \vec{v})$, measured at the initial region of phase-space, decreases as a function of time, as more and more “vacuum” is mixed in.

Phase Mixing

More details
in MBW §5.5



Phase-mixing of dark matter particles in a numerical N-body simulation. The particles are initially placed in a stratified sphere with zero-velocities. Collapse rapidly phase mixes the particles

Note how the number of particles in the red box, representing the coarse-grained DF, f_c , becomes more and more similar to that of neighboring boxes; the system is relaxing...

Note that phase-mixing is a relaxation process that does not cause any loss of information: at the fine-grained level, phase-mixing is perfectly reversible and preserves all knowledge of the initial conditions....

Chaotic Mixing

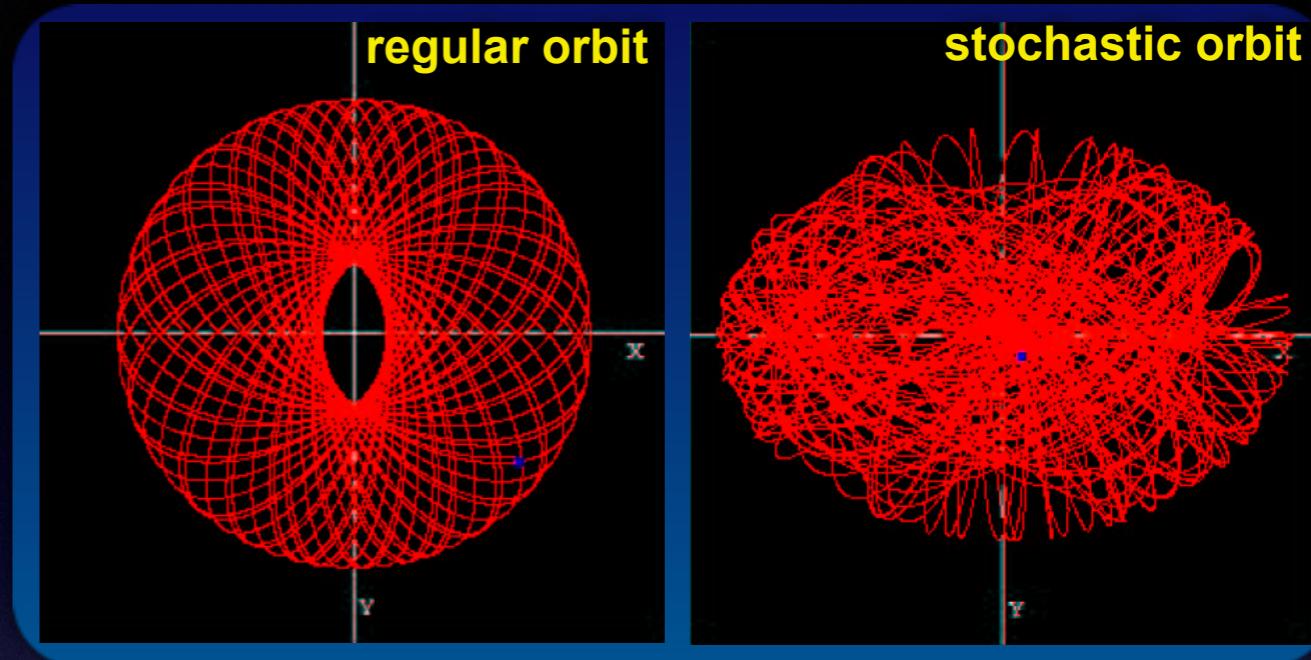
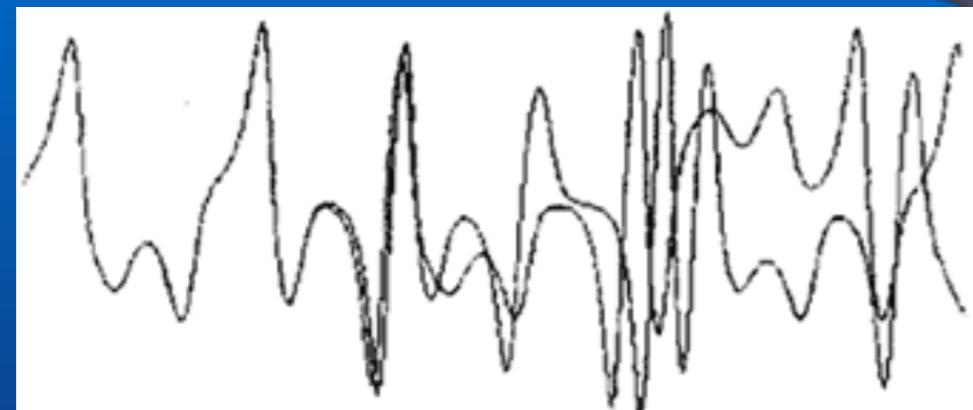
More details
in MBW §5.5

Some particles are on stochastic (or ‘chaotic’), rather than regular orbits (see MBW §5.4.5 for a detailed description). Such particles experience **chaotic mixing** (in addition to **phase mixing**).

Chaotic mixing arises from the fact that stochastic orbits separate exponentially with time (they have non-zero Lyapunov exponents)

Time-scale for **chaotic mixing**, however, is often much longer than Lyapunov time scale due to Arnold web diffusion...

divergence of
two trajectories
due to chaotic
behavior



Unlike for **phase-mixing**, **chaotic mixing** is **irreversible** in the sense that an infinitely precise fine-tuning of the phase-space coordinates is required to undo its effects...

Unlike for **phase-mixing**, which operates in all dynamical systems, **chaotic mixing** is only important if a significant fraction of phase-space is occupied by stochastic orbits.

Violent Relaxation

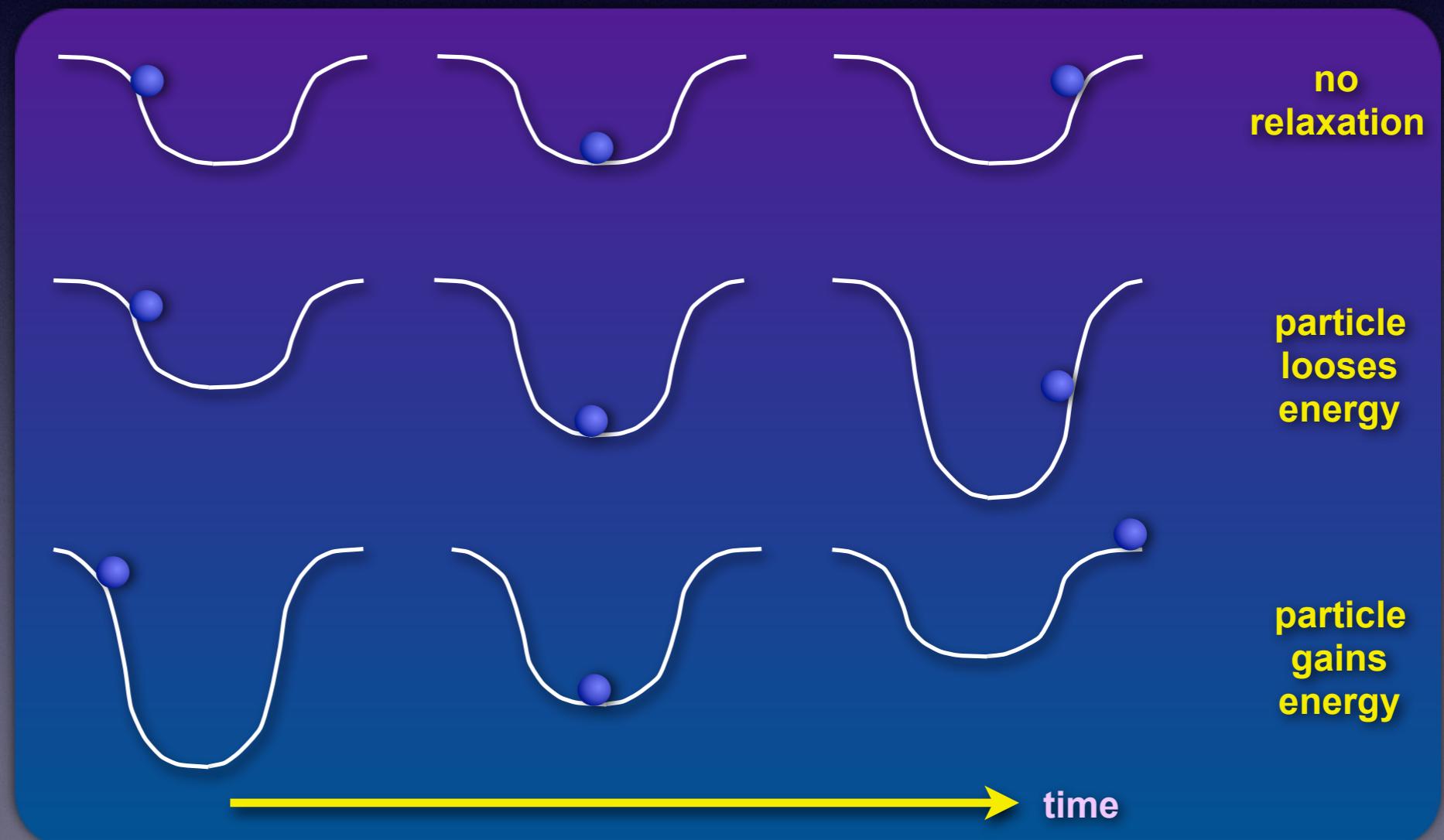
More details
in MBW §5.5

Since $E = v^2/2 + \Phi$ and $\Phi = \Phi(\vec{x}, t)$ we have that:

$$\frac{dE}{dt} = \frac{\partial E}{\partial \vec{v}} \frac{d\vec{v}}{dt} + \frac{\partial E}{\partial \Phi} \frac{d\Phi}{dt} = -\vec{v} \cdot \vec{\nabla} \Phi + \frac{d\Phi}{dt} = -\vec{v} \cdot \vec{\nabla} \Phi + \frac{\partial \Phi}{\partial t} + \frac{\partial \Phi}{\partial \vec{x}} \cdot \frac{d\vec{x}}{dt} = \frac{\partial \Phi}{\partial t}$$

Thus we see that the only way in which a particle's energy can change in a collisionless system is by having a time-dependent potential.

Exactly how a particle's energy changes due to violent relaxation depends in a complex way on the particle's initial position and energy: particles can gain or loose energy. Overall, however, violent relaxation increases the width of the energy distribution...



Violent Relaxation

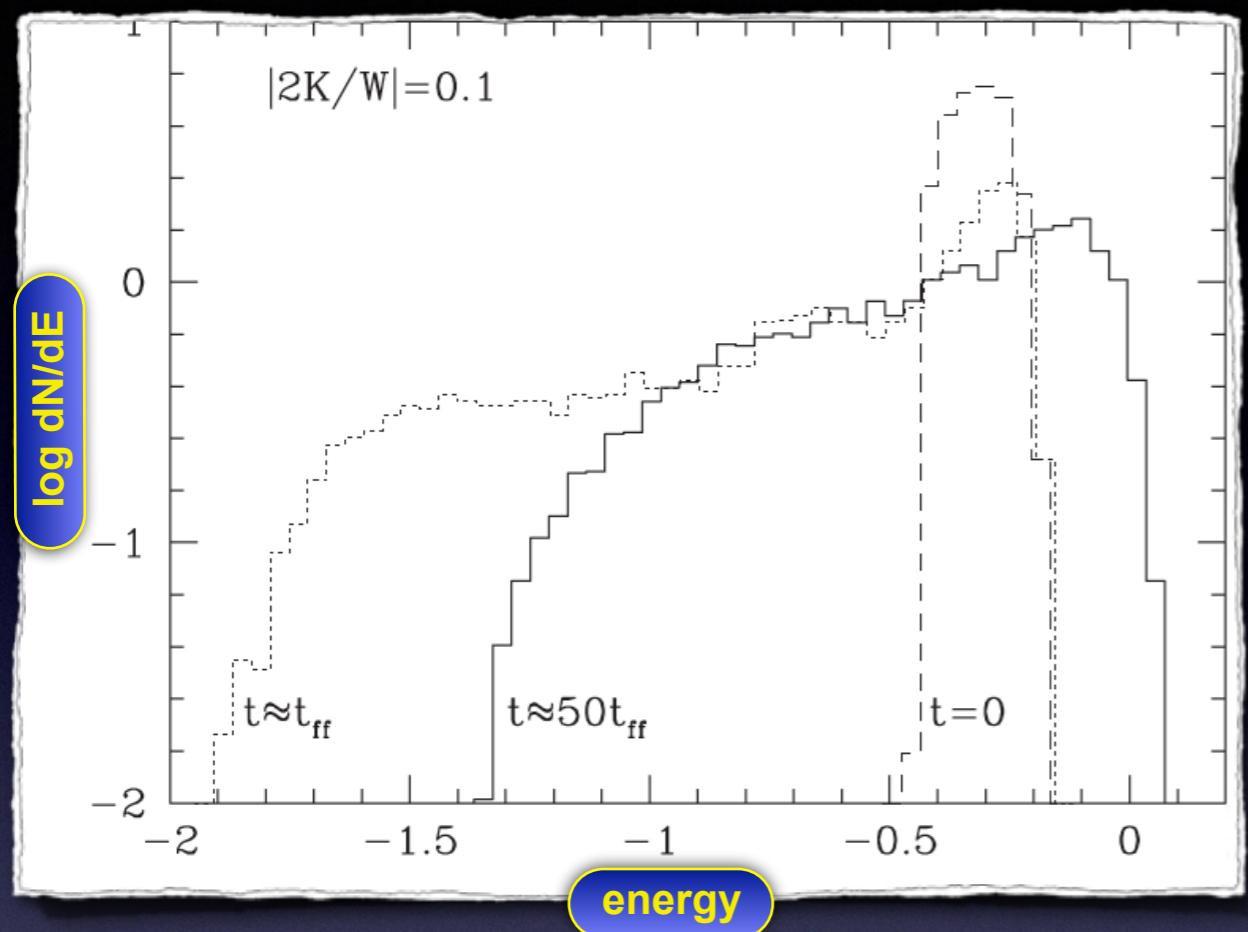
More details
in MBW §5.5

A few remarks about violent relaxation:

- Note that dE/dt is independent of particle mass; hence, **violent relaxation** has no tendency to **segregate** particles by mass (in fact, it will undo any pre-existing segregation). This is very different from **collisional relaxation**, where momentum exchange during collisions drives system towards equipartition of kinetic energy: more massive particles end up with lower velocities → **mass segregation**.
- During collapse of a **collisionless** system the **CBE** is still valid, i.e., the **fine-grained** DF does not evolve $df/dt = 0$ → **violent relaxation** only mixes at the **coarse-grained** level. Note, though, that unlike for a **steady-state** system, $\partial f / \partial t \neq 0$
- The time scale for **violent relaxation** is of order the time scale on which the potential changes by its own amount. This is basically the collapse time scale (\approx free fall time) → **violent relaxation** is very fast, hence its name
- **Violent relaxation** is **self-limiting**: as soon as a system approaches **any** equilibrium, the large-scale potential fluctuations vanish; the **mixing** due to violent relaxation destroys the coherence that drives potential fluctuations → **violent relaxation** does not run to completion; not all knowledge of initial conditions is erased

Violent Relaxation

More details
in MBW §5.5

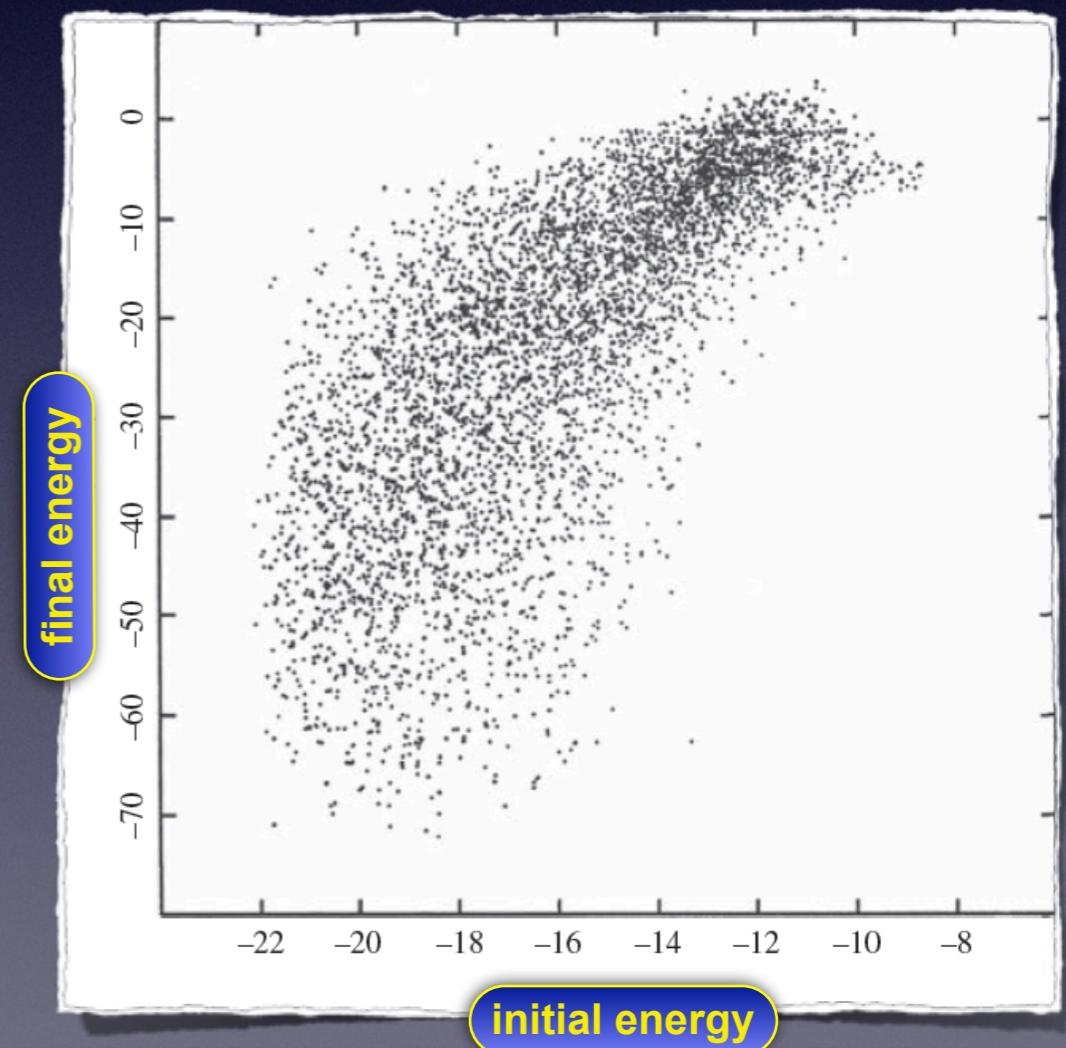


Differential energy distributions of particles in Nbody simulation of gravitational collapse. Note how violent relaxation broadens the energy distribution with time.

(from: van Albada 1982)

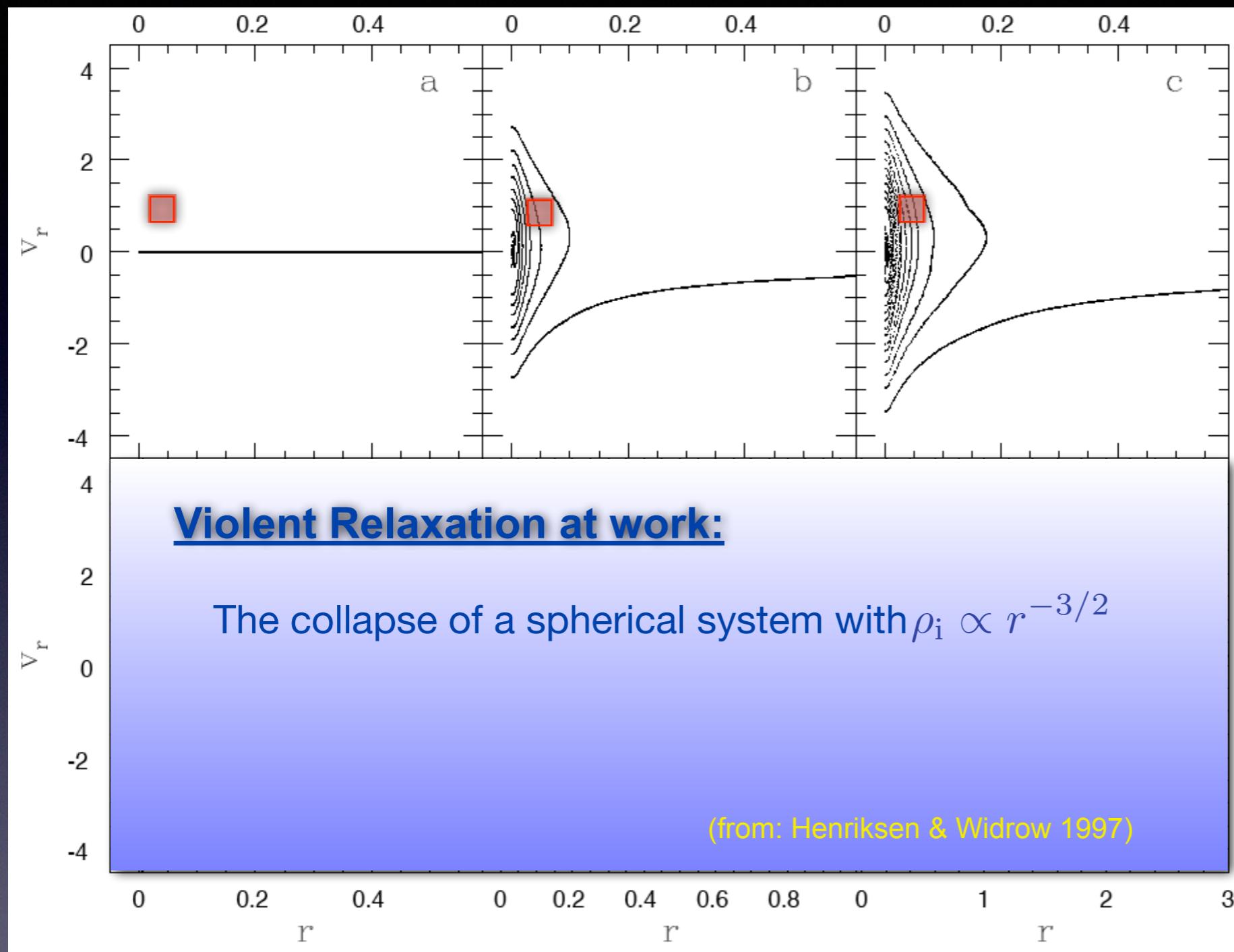
Scatter plot of final vs initial energies of the particles in the above Nbody simulation. Note that the correlation is significant, indicating that violent relaxation has not completely erased memory of the system's initial conditions.

(from: van Albada 1982)



Violent Relaxation

More details
in MBW §5.5



Note how **phase-mixing** is the dominant **relaxation** mechanism during the initial phases of the collapse.

After some time there is a transition to a more “erratic” flow: due to the time-varying potential phase-space streams start to undergo complicated bends and wiggles. This is **violent relaxation** at work!

Note how the number of particles in the **red box**, representing the coarse-grained DF, f_c , becomes more and more similar to that of neighboring boxes; the system is **relaxing**...

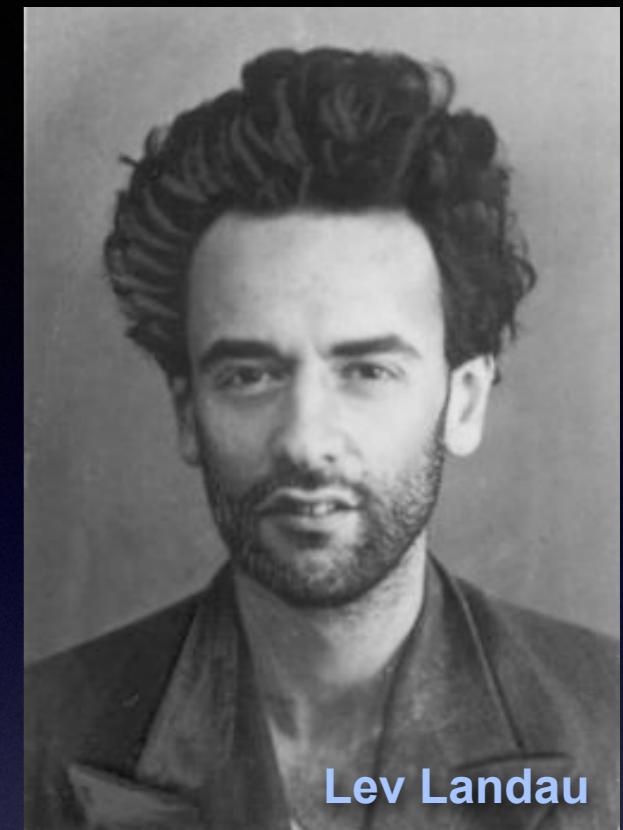
Violent relaxation leads to efficient coarse-grain mixing of the DF and erases the system’s memory of its initial conditions in a non-reversible way.

Landau Damping

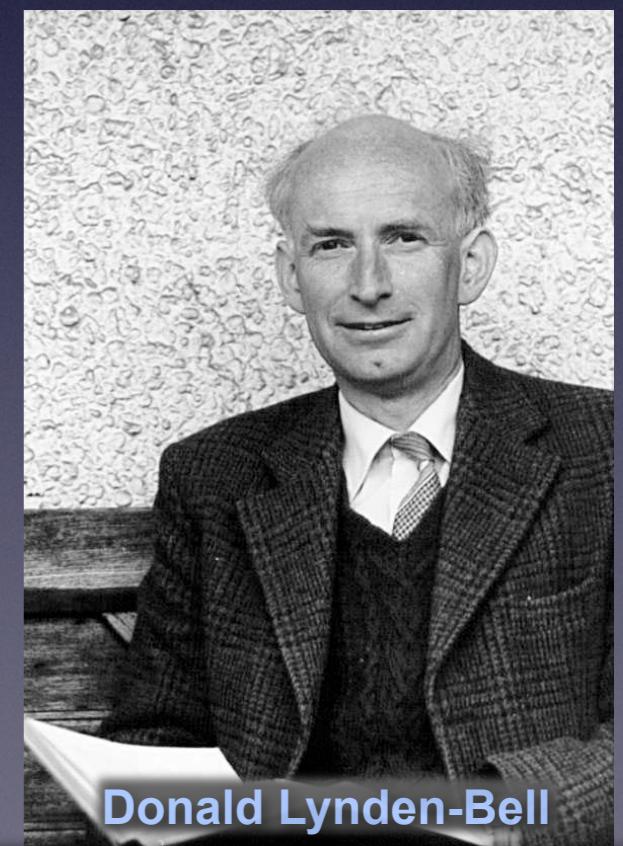
More details
in MBW §5.5

A detailed treatment of Landau damping is beyond the scope of these lectures.

- Landau damping is the damping of plasma waves due to interactions between wave and background particles.
- Lynden-Bell showed that a similar phenomenon also applies to a collisionless, gravitational system.
- It is basically the mechanism behind free-streaming damping of perturbations in a collisionless fluid.
- It arises due to decoherence between the particle velocities and the group velocity of the wave...
- Its main effect is to convert the energy in the wave, which is created due to some gravitational interaction/disturbance, into random motions of the background particles.



Lev Landau



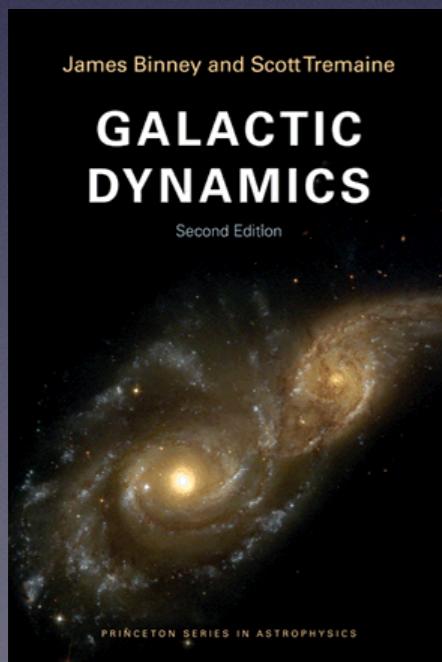
Donald Lynden-Bell

The End State of Relaxation

Several people have tried to use the principles of statistical mechanics to predict the end-state of a relaxed, collisionless system. Unfortunately, very little progress has been made, and there is currently no clear understanding on how to proceed, other than use numerical N-body simulations.

The main problem is that because of the long-range nature of gravity, it isn't clear how (best?) to define the entropy of a gravitational system...

Students: read MBW §5.5.5 for a discussion as to why a statistical mechanics treatment fails...



An excellent treatment of relaxation mechanisms in collisionless systems can be found in a review article by Merritt (1998): <http://arxiv.org/abs/astro-ph/9810371>

For a more in-depth treatment, consult the excellent textbook “Galactic Dynamics” by Binney & Tremaine

Lecture 8

SUMMARY

Summary: key words & important facts

Key words	
Spherical/Ellipsoidal collapse	Mode coupling
Secondary Infall model	Violent relaxation
Zel'dovich approximation	Phase Mixing
critical overdensity	Virial Theorem
shell crossing	Two-body relaxation

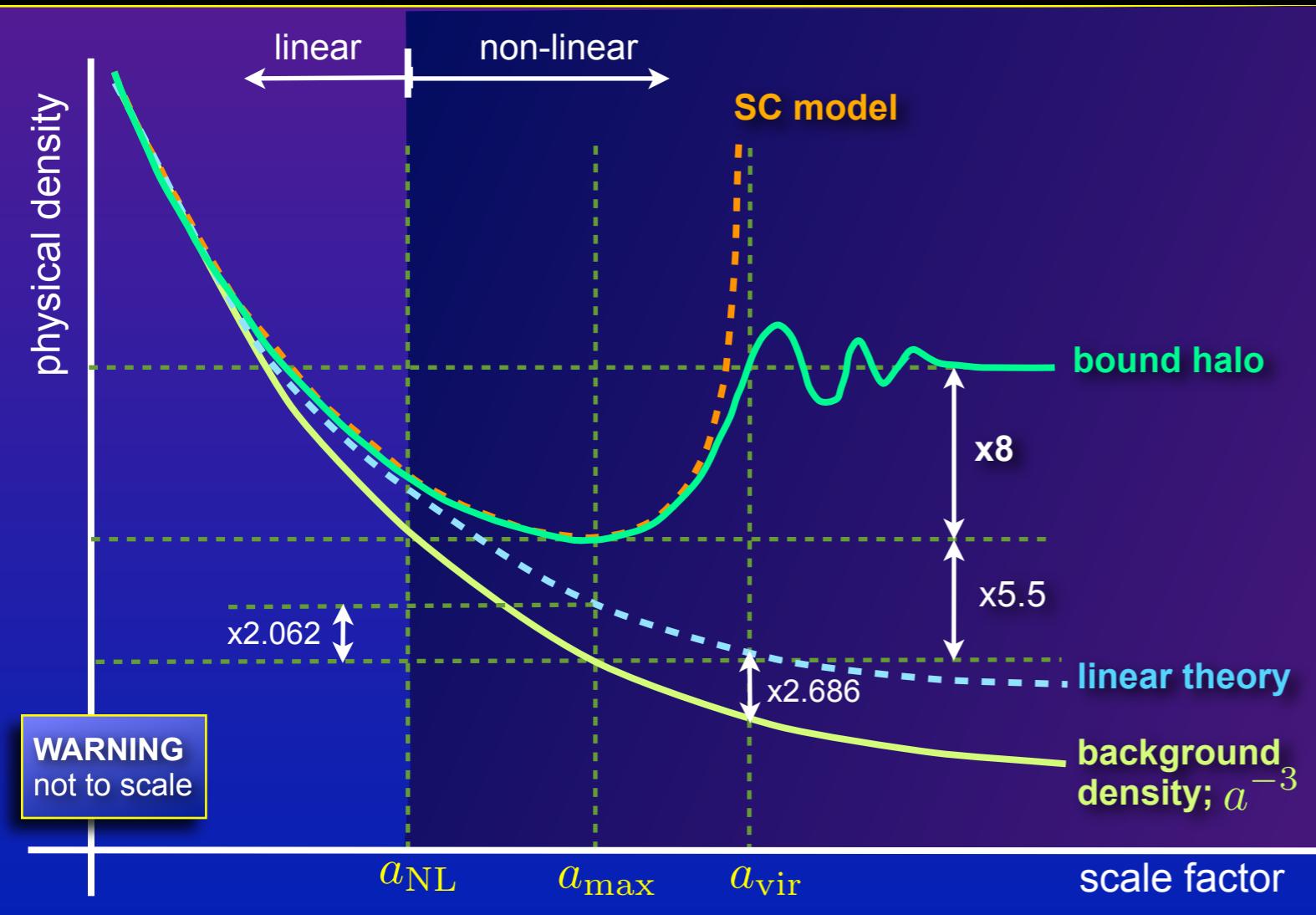
- In the **non-linear** regime ($\delta > 1$) perturbation theory is no longer valid. Modes start to couple to each other, and one can no longer describe the evolution of the density field with a simple growth rate: in general, no analytic solutions exist...
- Because of this **mode-coupling**, the density field loses its Gaussian properties, i.e., in the **non-linear** regime, density field cannot remain Gaussian.
- **Spherical Collapse** (SC) model can be used to ‘identify’ when and **where** collapsed objects will appear. **Ellipsoidal Collapse** model improves upon SC by accounting for the impact of tides, which typically are more important for less massive objects
- The **Zel'dovich approximation** is a Lagrangian treatment of the **displacement field**. It remains accurate in the quasi-linear regime, up to first shell crossing.

Summary: key words & important facts

Key words	
Spherical/Ellipsoidal collapse	Mode coupling
Secondary Infall model	Violent relaxation
Zel'dovich approximation	Phase Mixing
critical overdensity	Virial Theorem
shell crossing	Two-body relaxation

- There are four relaxation mechanisms for collisionless systems:
 - phase mixing
 - chaotic mixing
 - violent relaxation
 - Landau damping
- The only way in which a particle's energy can change in a collisionless system is by having a time-dependent potential.
- Unlike **collisional** relaxation, violent relaxation does not cause mass segregation
- Violent relaxation operates on the free-fall time, only mixes at the course-grain level of the distribution function, and is self-limiting.

Summary: key equations & expressions



$\delta = \rho/\bar{\rho} - 1$	turn-around	collapse
SC model	4.55	∞
linear model	1.062	1.686

Spherical Collapse model

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

Linear theory

$$\delta_{\text{lin}} = \frac{3}{20} (6\pi)^{2/3} \left(\frac{t}{t_{\max}} \right)^{2/3}$$

Virialization:

$$r_{\text{vir}} = r_{\text{ta}}/2$$

$$1 + \Delta_{\text{vir}} = 18\pi^2 \simeq 178 \sim 200$$

Zel'dovich Approximation

$$\vec{x}(t) = \vec{x}_i - \frac{D(a)}{4\pi G \bar{\rho}_i} \vec{\nabla} \Phi_i$$

two-body relaxation time: $t_{\text{relax}} \simeq \frac{N}{10 \ln N} t_{\text{cross}}$