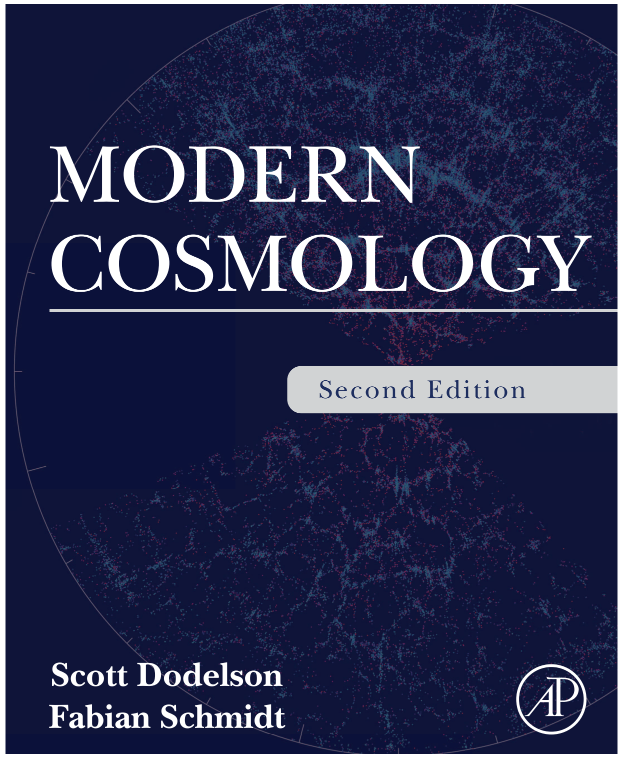
Structure Formation
Lecture 3

Fabian Schmidt
MPA

All figures taken from *Modern Cosmology, Second Edition*, unless otherwise noted



ICTP-SAIFR School on Cosmology, January 2021

Outline of lectures

1. The problem: collisionless Boltzmann equation and fluid approximation
    1. Linear evolution
2. Nonlinear evolution of matter
    1. Perturbation theory
    2. Simulations <- HERE
    3. Phenomenology of nonlinear matter distribution
3. Formation and distribution of galaxies
    1. Galaxy formation in a nutshell
    2. Spherical collapse model
    3. Physical clustering of halos and galaxies; bias
    4. Observed clustering of galaxies
4. Beyond ΛCDM

Notation

$$ds^2 = -(1 + 2\Psi(\boldsymbol{x}, t))dt^2 + a^2(t)(1 + 2\Phi(\boldsymbol{x}, t))d\boldsymbol{x}^2$$

- Comoving coordinates: $d\boldsymbol{r} = a(t)d\boldsymbol{x}$
- Conformal time: $d\eta = \dfrac{dt}{a(t)} = \dfrac{da}{a^2 H(a)} = \dfrac{d\ln a}{aH(a)}.$
  - Comoving distance: $d\chi = -d\eta = \dfrac{dz}{H(z)}$
- Particle velocity/momentum: $\boldsymbol{v} = \dfrac{\boldsymbol{p}}{m} = a\dfrac{d\boldsymbol{x}}{dt} = \boldsymbol{x}'$
- Fluid velocity; divergence: $\boldsymbol{u}; \quad \theta = \partial_i u^i$
- Gravitational potential: Ψ

Notation

- The linear power spectrum (filling in a gap in Lecture 1)

- Almost-scale-invariant power spectrum from inflation: $P_{\mathcal{R}}(k) = 2\pi^2 \mathcal{A}_s k^{-3} \left(\frac{k}{k_\mathrm{p}}\right)^{n_s-1}$

- Relating δ to curvature perturbation:

$$\delta^{(1)}(\boldsymbol{k},\eta) = \frac{2k^2 a(\eta)}{3\Omega_m H_0^2}\Phi(\boldsymbol{k},\eta); \quad \Phi(\boldsymbol{k},\eta) = \frac{3}{5}T(k)\frac{D(a)}{a}\mathcal{R}(\boldsymbol{k})$$

Note: throughout, δ is in synchronous-comoving gauge. Then this relation remains *valid on all scales*.

- Result:

$$P_\mathrm{L}(k,a) = \frac{8\pi^2}{25}\frac{\mathcal{A}_s}{\Omega_\mathrm{m}^2}D_+^2(a)T^2(k)\frac{k^{n_s}}{H_0^4 k_\mathrm{p}^{n_s-1}}$$

$$(D_+ \equiv D)$$

Notation

- The linear power spectrum (filling in a gap in Lecture 1)

- Almost-scale-invariant power spectrum from inflation: $P_{\mathcal{R}}(k) = 2\pi^2 \mathcal{A}_s k^{-3} \left(\frac{k}{k_{\rm p}}\right)^{n_s - 1}$

- Relating δ to curvature perturbation:

Growth factor -> L. 1

$$\delta^{(1)}(\boldsymbol{k}, \eta) = \frac{2k^2 a(\eta)}{3\Omega_m H_0^2} \Phi(\boldsymbol{k}, \eta); \quad \Phi(\boldsymbol{k}, \eta) = \frac{3}{5} T(k) \frac{D(a)}{a} \mathcal{R}(\boldsymbol{k})$$

- Result:

$$P_{\rm L}(k, a) = \frac{8\pi^2}{25} \frac{\mathcal{A}_s}{\Omega_{\rm m}^2} D_+^2(a) T^2(k) \frac{k^{n_s}}{H_0^4 k_{\rm p}^{n_s - 1}}$$

$$(D_+ \equiv D)$$

Notation

- The linear power spectrum (filling in a gap in Lecture 1)
- Almost-scale-invariant power spectrum from inflation: $P_{\mathcal{R}}(k) = 2\pi^2 \mathcal{A}_s k^{-3} \left(\frac{k}{k_\mathrm{p}}\right)^{n_s - 1}$
- Relating δ to curvature perturbation:

$$\delta^{(1)}(\boldsymbol{k}, \eta) = \frac{2k^2 a(\eta)}{3\Omega_m H_0^2} \Phi(\boldsymbol{k}, \eta); \quad \Phi(\boldsymbol{k}, \eta) = \frac{3}{5} T(k) \frac{D(a)}{a} \mathcal{R}(\boldsymbol{k})$$

Transfer function: evolution of perturbations in early Universe

- Result:

$$P_\mathrm{L}(k, a) = \frac{8\pi^2}{25} \frac{\mathcal{A}_s}{\Omega_\mathrm{m}^2} D_+^2(a) T^2(k) \frac{k^{n_s}}{H_0^4 k_\mathrm{p}^{n_s - 1}}$$

$$(D_+ \equiv D)$$

Notation

- The linear power spectrum (filling in a gap in Lecture 1)

- Almost-scale-invariant power spectrum from inflation: $P_{\mathcal{R}}(k) = 2\pi^2 \mathcal{A}_s k^{-3} \left(\frac{k}{k_{\rm p}}\right)^{n_s-1}$

- Relating δ to curvature perturbation:

$$\delta^{(1)}(\boldsymbol{k},\eta) = \frac{2k^2 a(\eta)}{3\Omega_m H_0^2}\Phi(\boldsymbol{k},\eta); \quad \Phi(\boldsymbol{k},\eta) = \frac{3}{5}T(k)\frac{D(a)}{a}\mathcal{R}(\boldsymbol{k})$$

- Result:

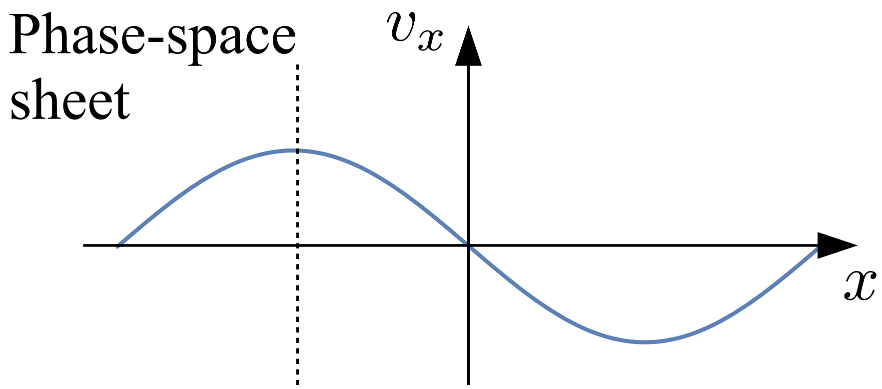
$$P_{\rm L}(k,a) = \frac{8\pi^2}{25}\frac{\mathcal{A}_s}{\Omega_{\rm m}^2} D_+^2(a) T^2(k) \frac{k^{n_s}}{H_0^4 k_{\rm p}^{n_s-1}}$$

Time dependence of linear P(k) ~ D^2, as we showed in L. 1

$$(D_+ \equiv D)$$

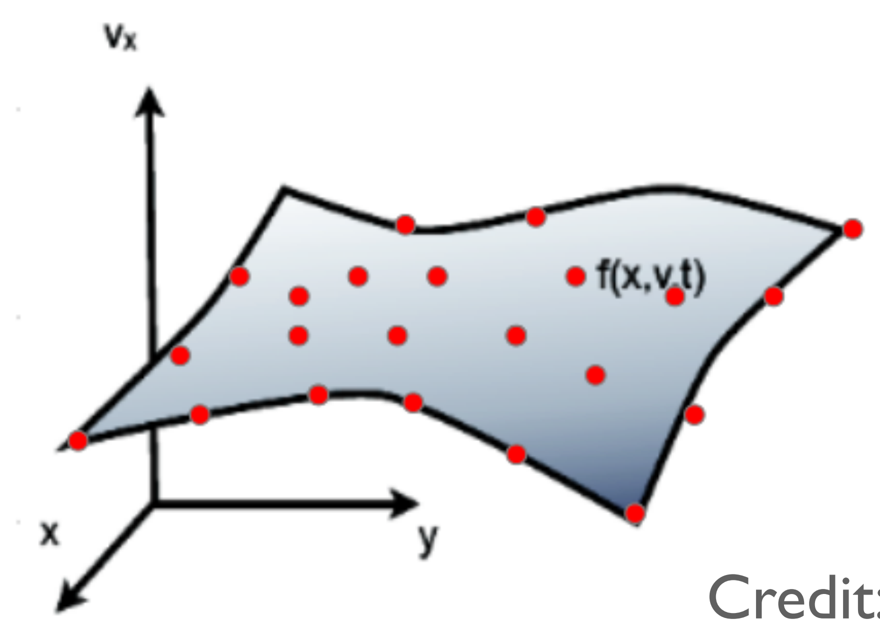
Structure formation beyond perturbation theory

- In order to take phasespace evolution into account properly, need to go beyond fluid picture and perturbation theory.
- Back to collisionless Boltzmann equation!
- Instead of trying to calculate 6D distribution function, we will discretize the thin phasespace sheet that (dark) matter is localized in

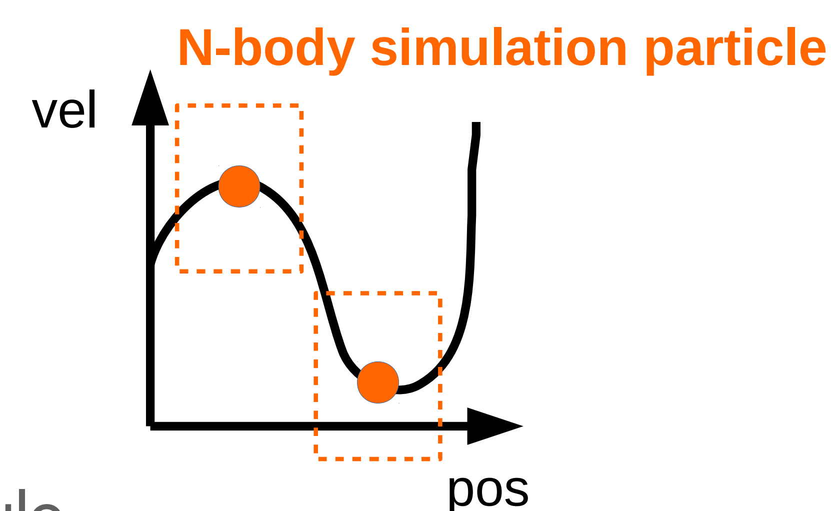


N-body simulations

- Discretize the thin phasespace sheet that (dark) matter is localized in
- Follow the evolution of phasespace elements ("particles") by integrating the geodesic equation (*characteristics* of the PDE)



Credit: Raul Angulo



N-body simulations

- Discretize the thin phasespace sheet that (dark) matter is localized in
- Follow the evolution of phasespace elements ("particles") by integrating the geodesic equation (*characteristics* of the PDE)
- Equations of motion:

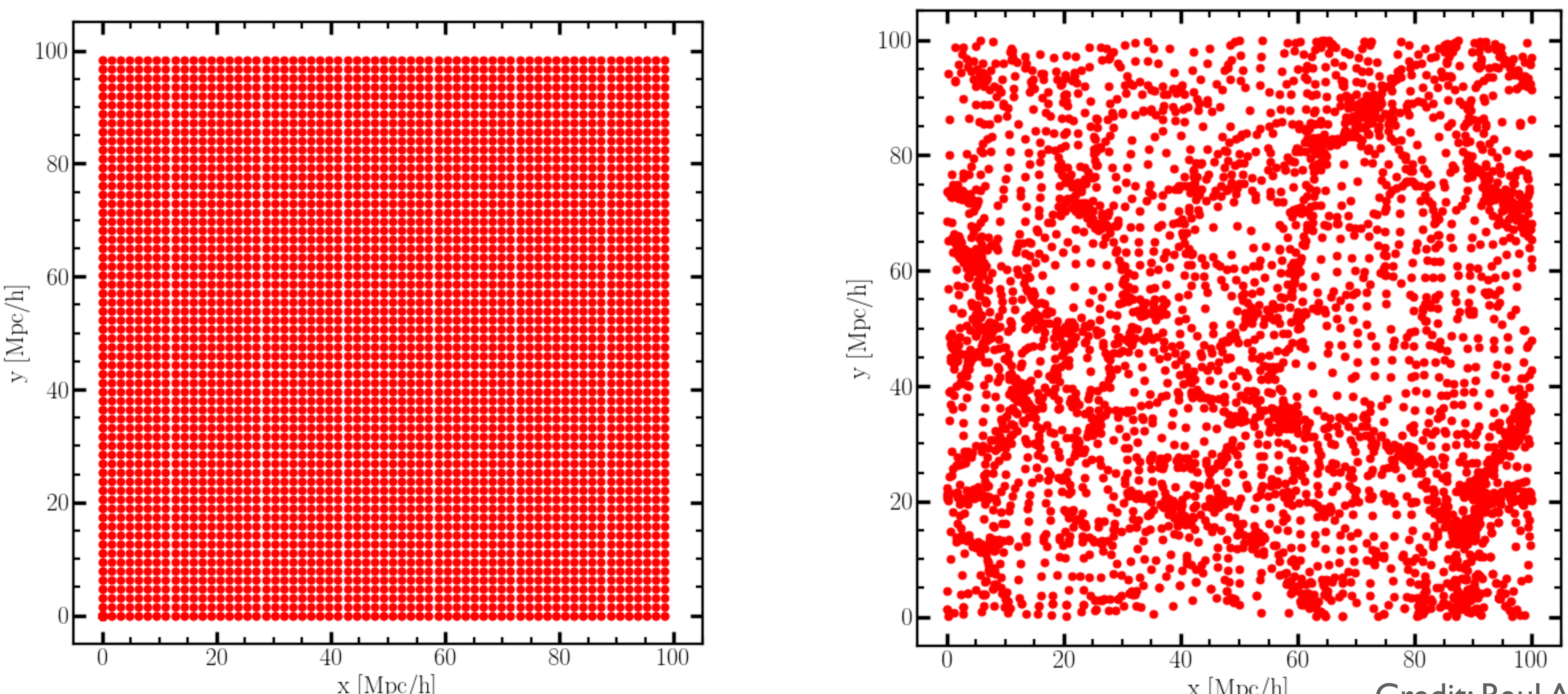
$$\frac{dx^i}{dt} = \frac{p^i}{ma}, \qquad \frac{dp^i}{dt} = -Hp^i - \frac{m}{a}\frac{\partial\Psi}{\partial x^i}. \quad \xrightarrow{p_c^i = ap^i} \quad \frac{dx^i}{dt} = \frac{p_c^i}{ma^2}, \qquad \frac{dp_c^i}{dt} = -m\frac{\partial\Psi}{\partial x^i}.$$

Eq. (12.57)

and $\nabla^2\Psi = \frac{3}{2}\Omega_{\mathrm{m}}(\eta)(aH)^2\delta_{\mathrm{m}}.$

N-body simulations in one step

- Structure appears after just one time step!
- Specifically, solve for particle displacement at linear order: 1st order Lagrangian perturbation theory (Zel'dovich approximation)
- Technique used to generate initial conditions for full simulations at high z



Credit: Raul Angulo

Time integration

- Leapfrog scheme: preserves energy (as well as other constants of motion) to cubic order in time step

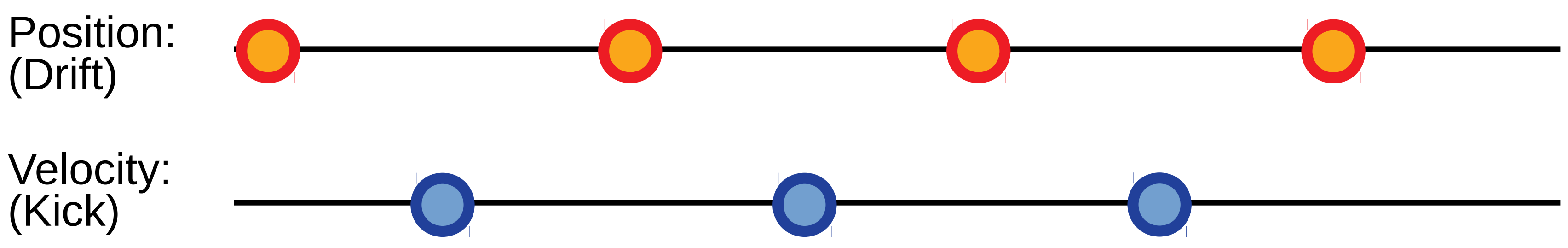
1. Compute the gravitational potential generated by the collection of particles, and take its gradient to obtain $\nabla\Psi(\boldsymbol{x}, t)$
2. Change each particle's momentum ("kick") by

$$\boldsymbol{p}_c^{(i)}(t + \Delta t/2) = \boldsymbol{p}_c^{(i)}(t - \Delta t/2) - m\nabla\Psi(\boldsymbol{x}^{(i)}, t)\Delta t. \tag{12.59}$$

3. Move each particle position ("drift") by

$$\boldsymbol{x}^{(i)}(t + \Delta t) = \boldsymbol{x}^{(i)}(t) + \frac{\boldsymbol{p}_c^{(i)}(t + \Delta t/2)}{ma^2(t + \Delta t/2)}\Delta t. \tag{12.60}$$

4. Repeat.



Credit: Raul Angulo

Solving for the force

- In order to move particles, we need to calculate the force (gradient of the potential) at each particle's position
- As accurately as possible, but with reasonable cost
- Example: could directly sum up forces of all other particles (*direct summation or particle-particle algorithm*)

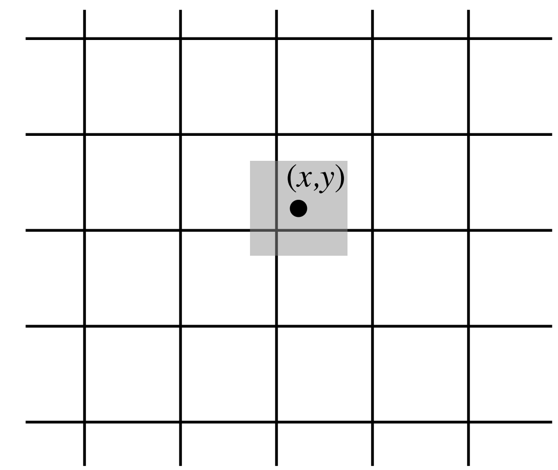
$$\nabla\Psi(\boldsymbol{x}_i) = Gm \sum_{j \neq i} \frac{\boldsymbol{x}_j - \boldsymbol{x}_i}{|\boldsymbol{x}_j - \boldsymbol{x}_i|^3}$$

- Problem: computational cost scales as N^2 for N particles
- Our goal is to use billions of particles, so algorithm has to scale as N or $N \log N$ at most.

Reminder: N-body particles are *not actual particles!*

Solving for the force

- Simplest solution: solve potential on a fixed grid: *particle-mesh (PM) algorithm*



1. Assign particles to grid to obtain density (most commonly using cloud-in-cell [CIC] assignment)
2. Solve Poisson equation on grid, via discrete Fast Fourier Transform (FFT)

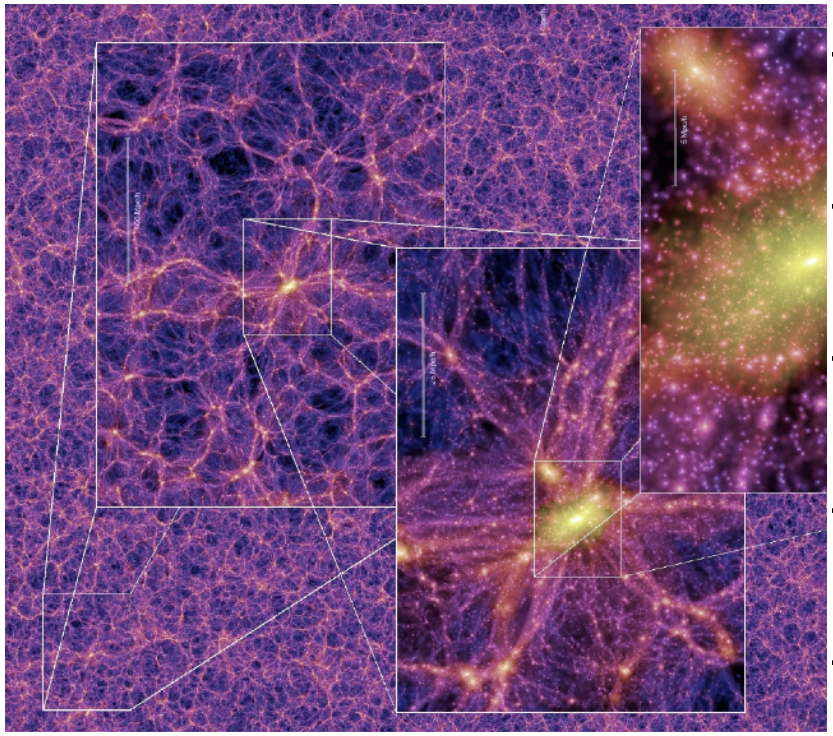
$$\nabla^2 \Psi = \frac{3}{2} \Omega_{\mathrm{m}}(\eta)(aH)^2 \delta_{\mathrm{m}}.$$

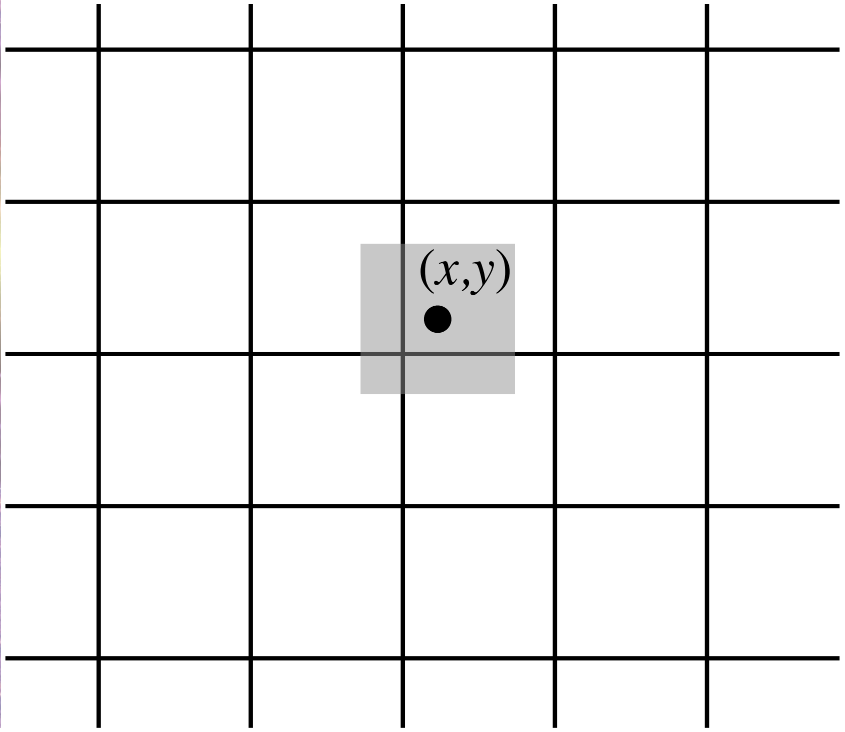
3. Interpolate gradient of resulting potential to each particle's position (using same interpolation scheme)

- Cost: *N log N* due to FFT

The resolution problem

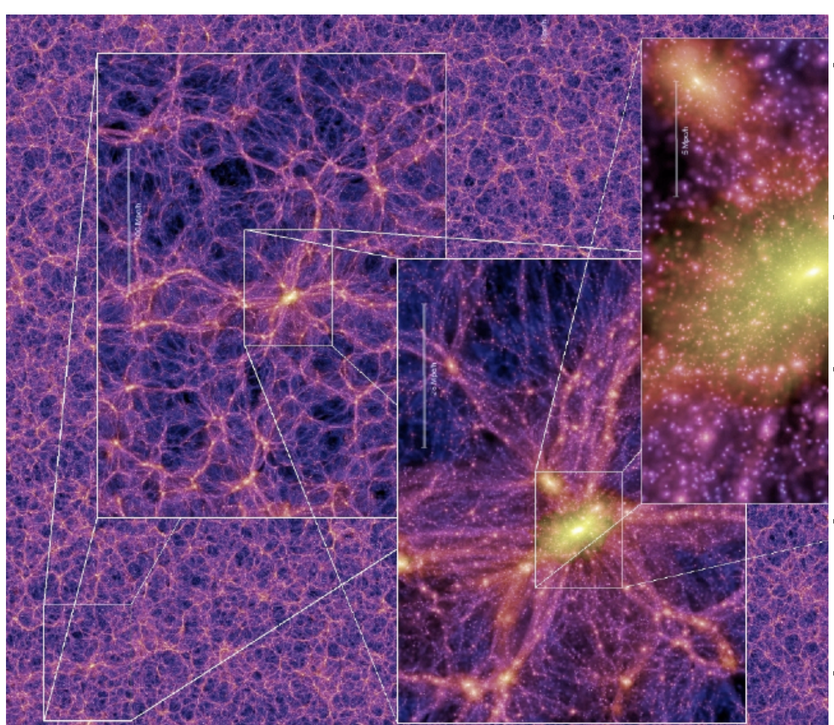
- Issue: fixed grid means fixed resolution
- Cannot resolve small-scale structure below grid resolution, and increasing resolution is memory- and CPU-intensive ($\sim N_{grid}^3$)
- On the other hand, we don't need high resolution for a large fraction of the volume

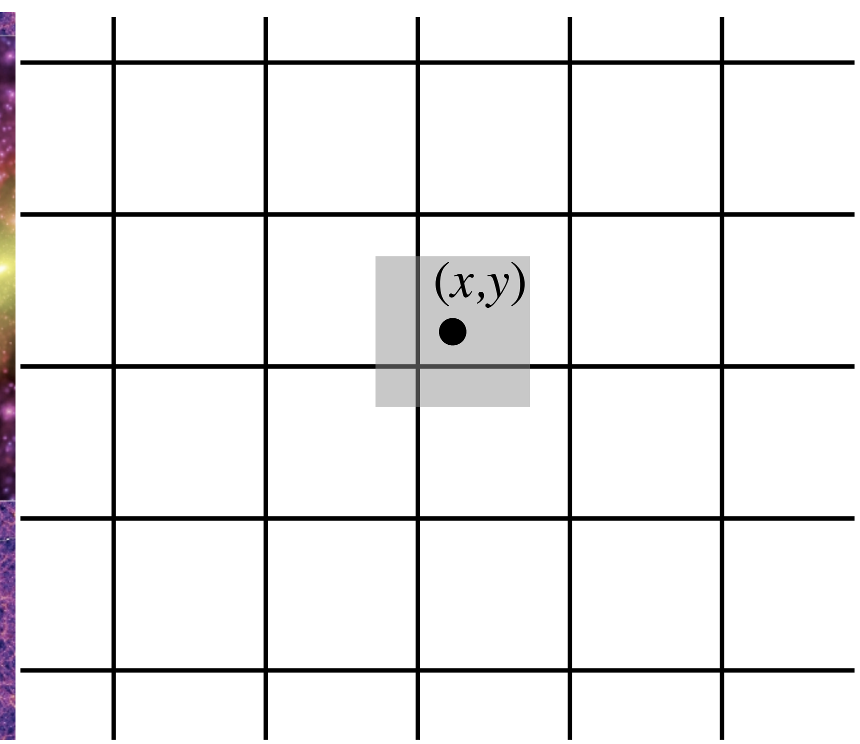




The resolution problem

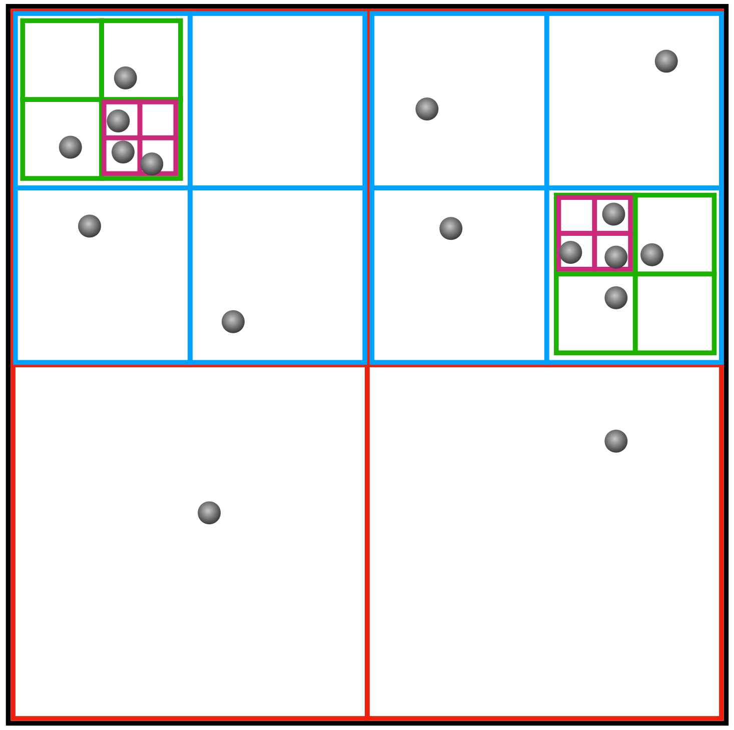
- Issue: fixed grid means fixed resolution
- On the other hand, we don't need high resolution for a large fraction of the volume
- Solution: *adaptive* algorithms which go to higher resolution only where necessary





Adaptive mesh refinement

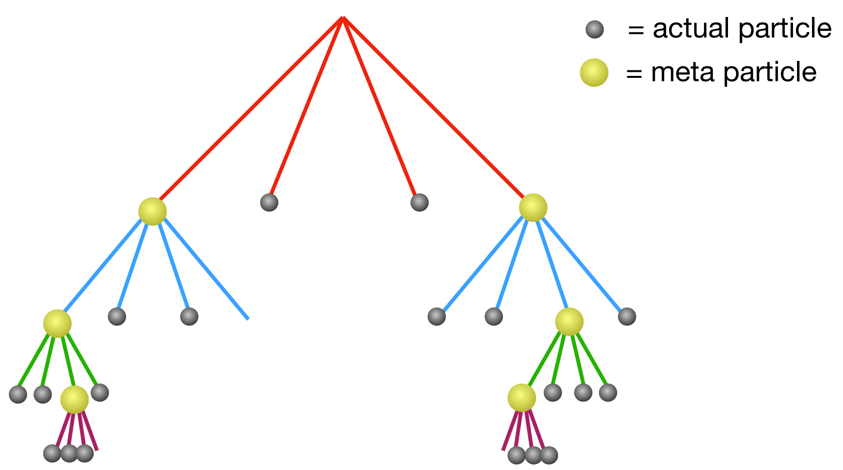
- Start with regular base grid
- Split any cell that crosses a certain particle number threshold into 8 sub-cells; repeat process until particle number sufficiently small in all cells
- Advantage: can use same grid for hydrodynamics
- Disadvantage: need relaxation method to solve Poisson equation on subgrids to incorporate boundary conditions



Examples: RAMSES, ART

Tree algorithm

- Get rid of grid: instead, lump particles together to compute their effect on distant other particles



Credit: Oliver Hahn's lecture notes

Tree algorithm

- Get rid of grid: instead, lump particles together to compute their effect on distant other particles

At each node in the tree, one inserts a 'meta' particle that carries the mass and sits at the centre of mass of the branch of the tree. The total interaction

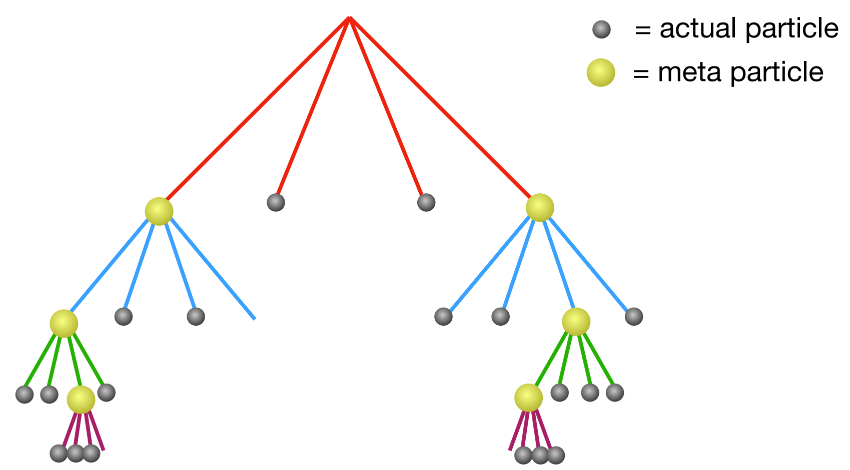
$$\phi(\mathbf{x}) \propto \sum_j \frac{1}{|\mathbf{x}-\mathbf{x}_j|} = \sum_j \frac{1}{|(\mathbf{x}-\boldsymbol{\lambda})-(\mathbf{x}_j-\boldsymbol{\lambda})|}$$

sum can be multipole expanded using

$$\frac{1}{|\mathbf{y}+\boldsymbol{\lambda}-\mathbf{x}_j|} \simeq \frac{1}{|y|} - \mathbf{y}\cdot\frac{\boldsymbol{\lambda}-\mathbf{x}_j}{|\mathbf{y}|^3} + \ldots.$$

If we set λ to be the centre of mass, then the dipole vanishes. The trick is now that since we know that $|mathbf{x_j} - \boldsymbol{\lambda}|$ is bounded by the space partitioning cell size, we can directly control the accuracy by accepting a meta particle as a valid approximation for the entire branch if

$$\theta = \frac{\ell}{|\mathbf{y}|} < \theta_c, \qquad (3.28)$$

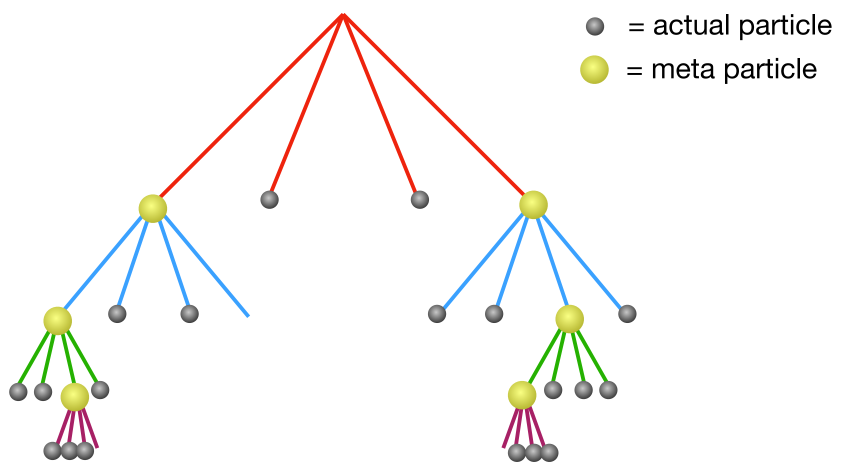


On small scales, use *force softening* to avoid particle "collisions", hard scattering.

Credit: Oliver Hahn's lecture notes

Tree algorithm

- Get rid of grid: instead, lump particles together to compute their effect on distant other particles
- Advantage: elegant and direct, no need for relaxation
- Disadvantage: need top-level grid for periodic boundary conditions, and hence split into grid and tree forces (*Tree-PM*)



Examples: Gadget, PKDgrav

Brief history of cosmological simulations

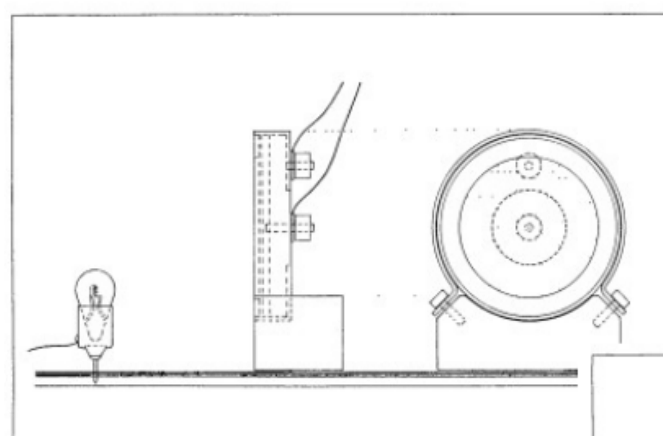
First N-body calculation: light bulbs, photocells and galvanometers

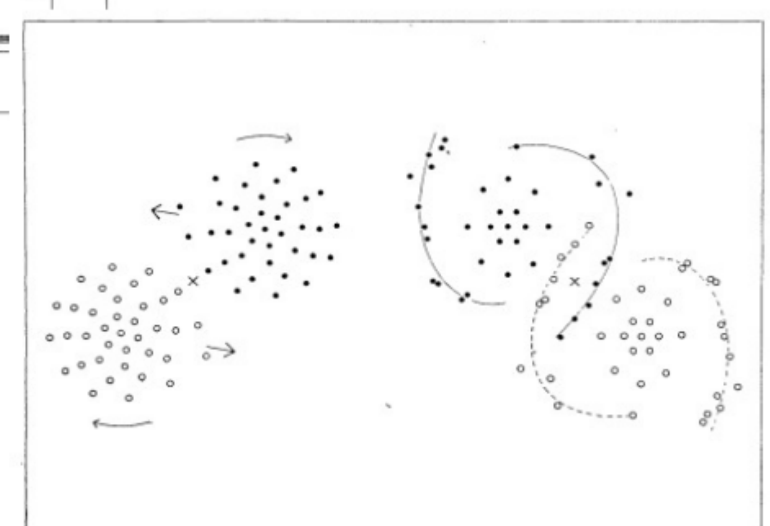
VOLUME 94 NOVEMBER 1941 NUMBER 3

ON THE CLUSTERING TENDENCIES AMONG THE NEBULAE

II. A STUDY OF ENCOUNTERS BETWEEN LABORATORY MODELS OF STELLAR SYSTEMS BY A NEW INTEGRATION PROCEDURE

ERIK HOLMBERG





Tidal features appear in interacting nebulae

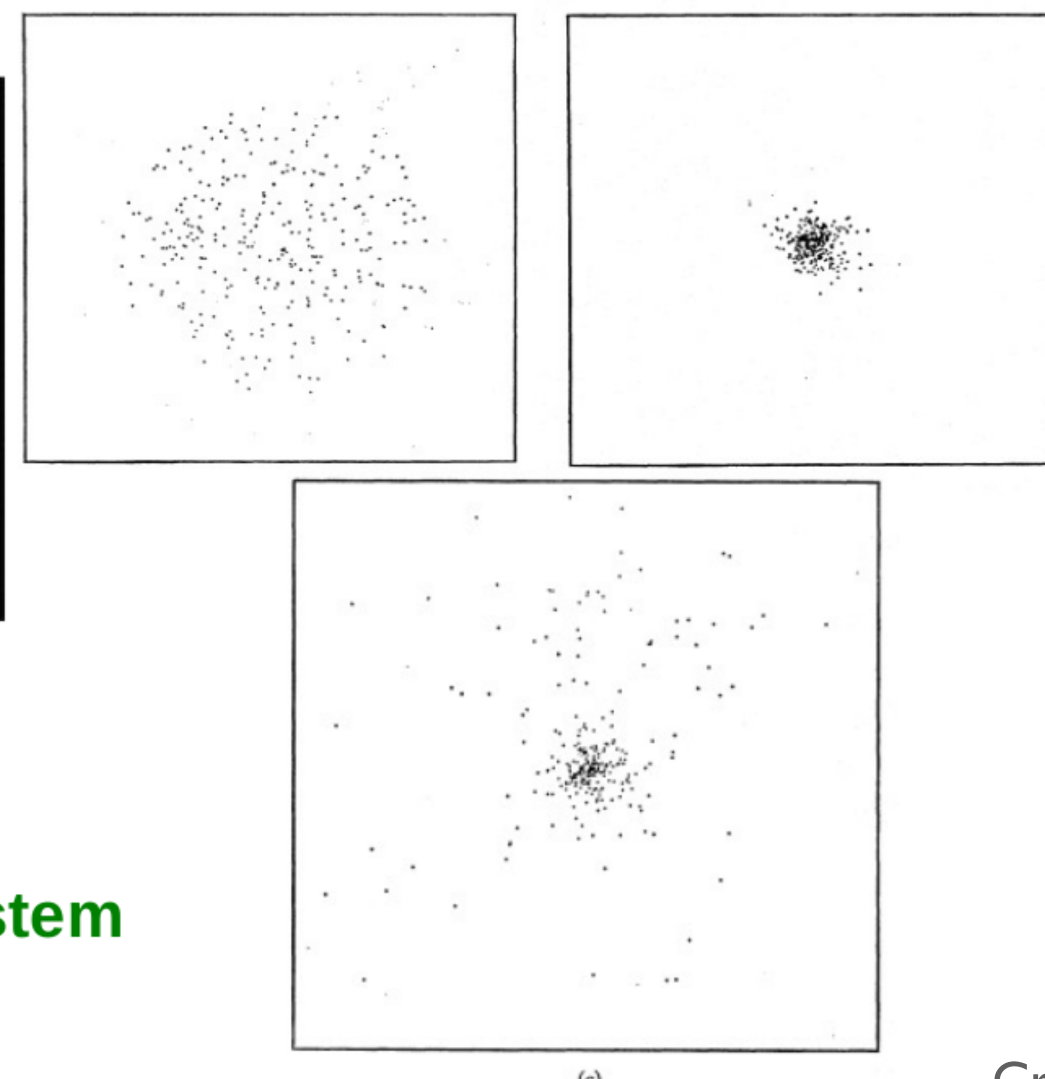
Credit: Raul Angulo

Brief history of cosmological simulations

N-body calculation supports the idea of dark Matter

A rotating group of 300 bodies ends up *too* concentrated.

1970-1974, Princeton

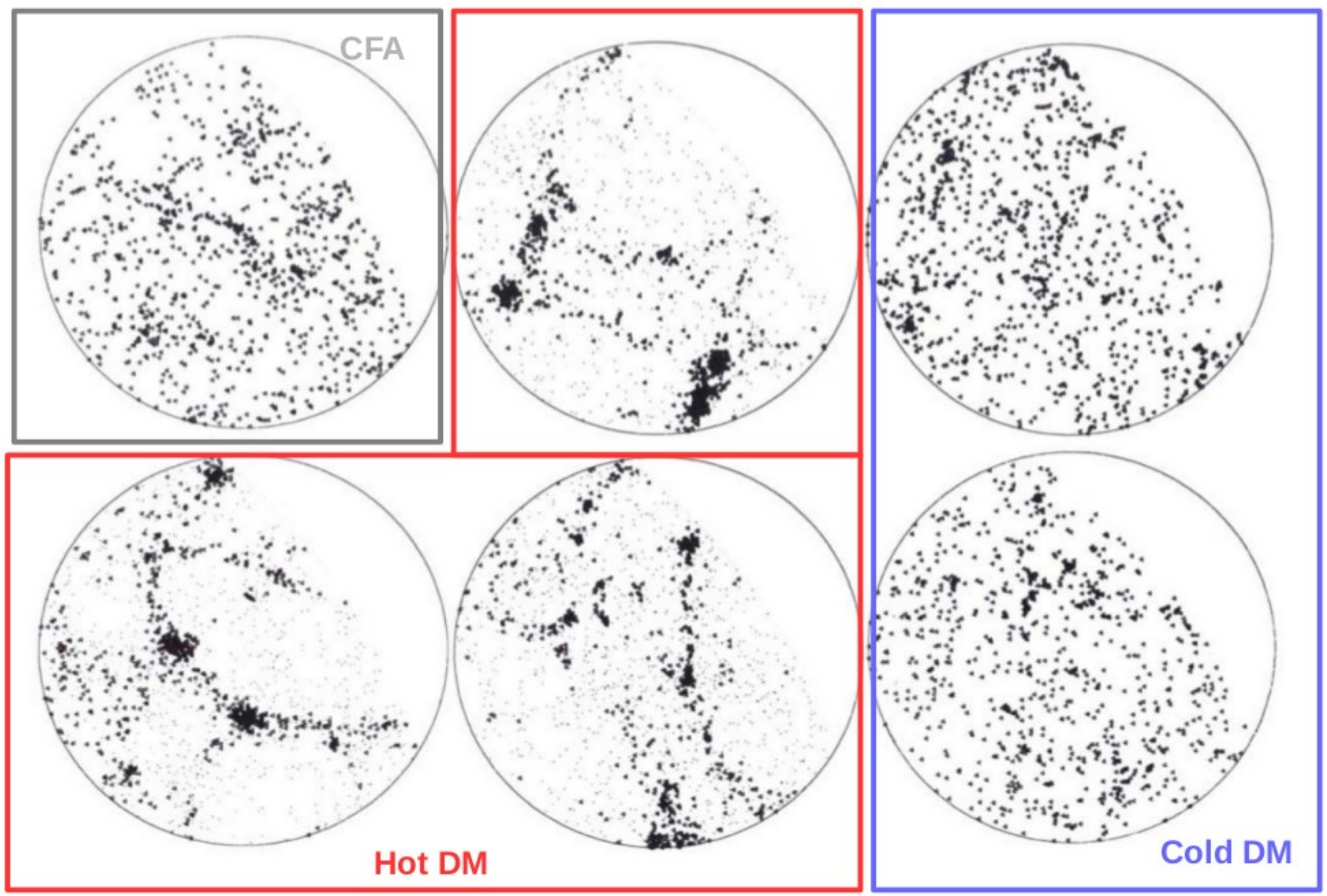


Dark matter (M/L=10) is needed to stabilize the system

Credit: Raul Angulo

Brief history of cosmological simulations

1985: The CDM model plus gravitational instability can explain qualitatively the observed universe

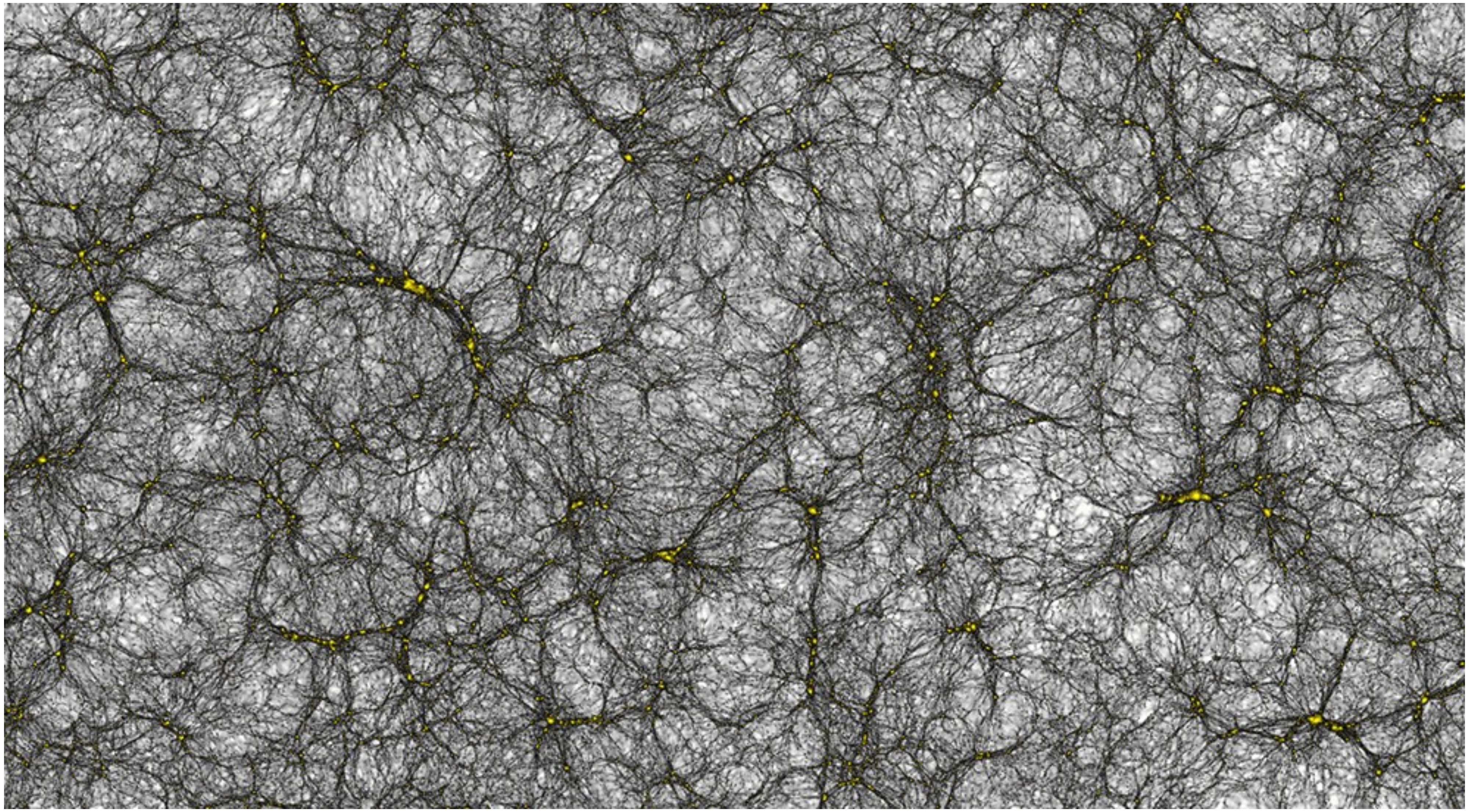


Davis, Efstathiou, Frenk & White 1985

Credit: Raul Angulo

Brief history of cosmological simulations

2017: EUCLID Flagship Simulation



Potter et al 2017

Credit: Raul Angulo

Result of an N-body simulation

- Catalog of "particle" positions and velocities at various points in time
- Basically, position of the phase-space sheet after gravitational evolution
- Use this to degenerate density field, for example

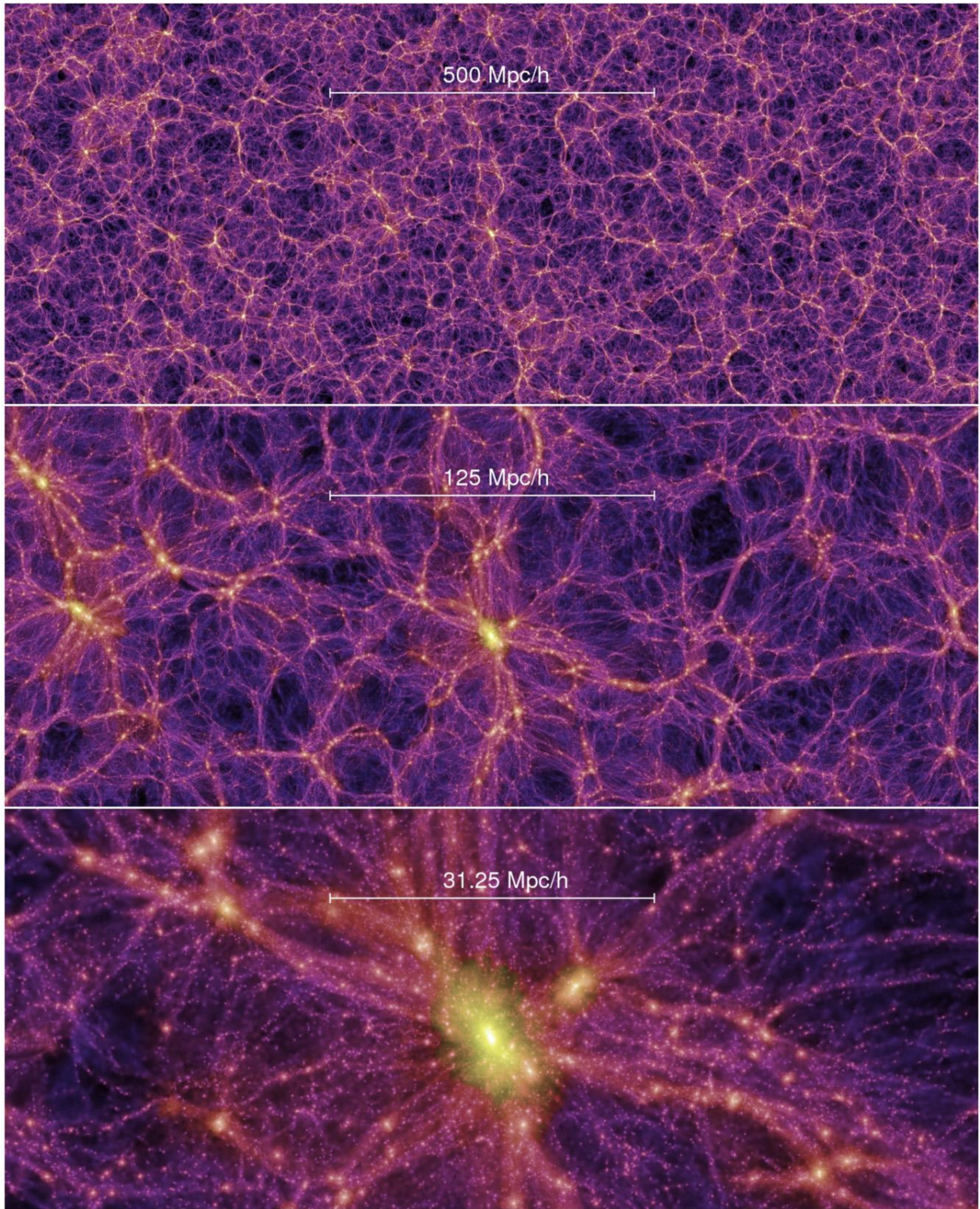


FIGURE 12.6 Slices of width $15\,h^{-1}$ Mpc through the density field at redshift zero in the *Millennium* N-body simulation which follows 10^{10} particles (i.e., phase-space elements). From top to bottom, the different panels zoom in to show the hierarchical nature of the matter distribution in a ΛCDM cosmology. The spatial scale is labeled in each panel. The color scale denotes density in logarithmic units. The simulations shown here are described in Springel et al. (2005).

Result of an N-body simulation

- Catalog of "particle" positions and velocities at various points in time
- Basically, position of the phase-space sheet after gravitational evolution
- Use this to generate density field, for example

Millennium simulation (Volker Springel)
Logarithmic color scale, comoving length units

Result of an N-body simulation

- Catalog of "particle" positions and velocities at various points in time
- Basically, position of the phase-space sheet after gravitational evolution
- Use this to generate density field, for example

Millennium simulation (Volker Springel)
Logarithmic color scale, comoving length units

Phenomenology of nonlinear structure

- Small-scale density fluctuations are largest: small-scale structure forms first
- Then, structure successively assembles to large-mass objects
- Topologically, we have
  - "3D:" voids - large underdense regions
  - "2D:" sheets, or "pancakes"
  - "1D:" filaments
  - "0D:" bound structures - halos

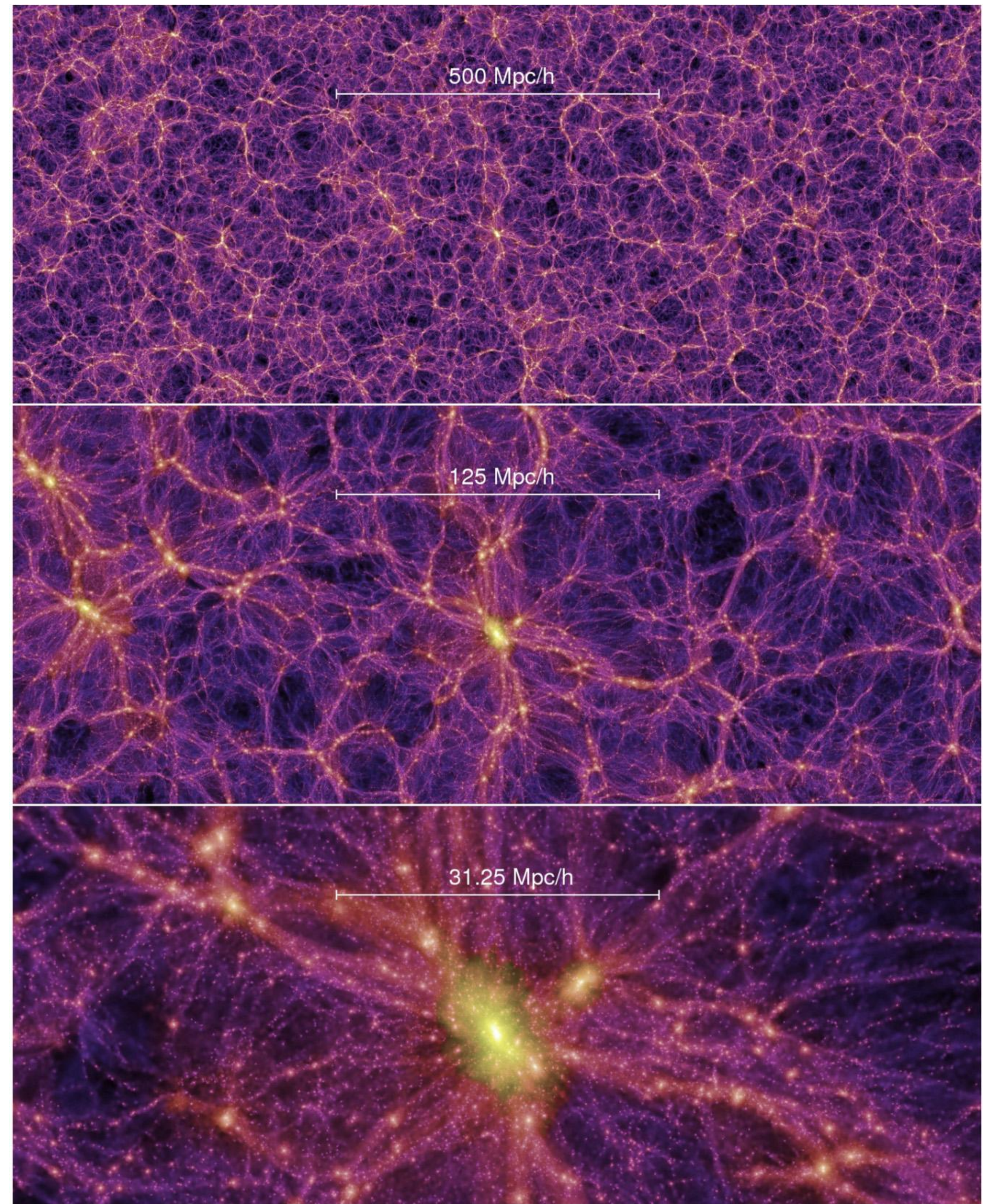


FIGURE 12.6 Slices of width $15\,h^{-1}\,\mathrm{Mpc}$ through the density field at redshift zero in the *Millennium* N-body simulation which follows 10^{10} particles (i.e., phase-space elements). From top to bottom, the different panels zoom in to show the hierarchical nature of the matter distribution in a ΛCDM cosmology. The spatial scale is labeled in each panel. The color scale denotes density in logarithmic units. The simulations shown here are described in Springel et al. (2005).

Phenomenology of nonlinear structure

- Small-scale density fluctuations are largest: small-scale structure forms first
- Then, structure successively assembles to large-mass objects
- Topologically, we have
  - "3D:" voids - large underdense regions
  - "2D:" sheets, or "pancakes"
  - "1D:" filaments
  - "0D:" bound structures - halos

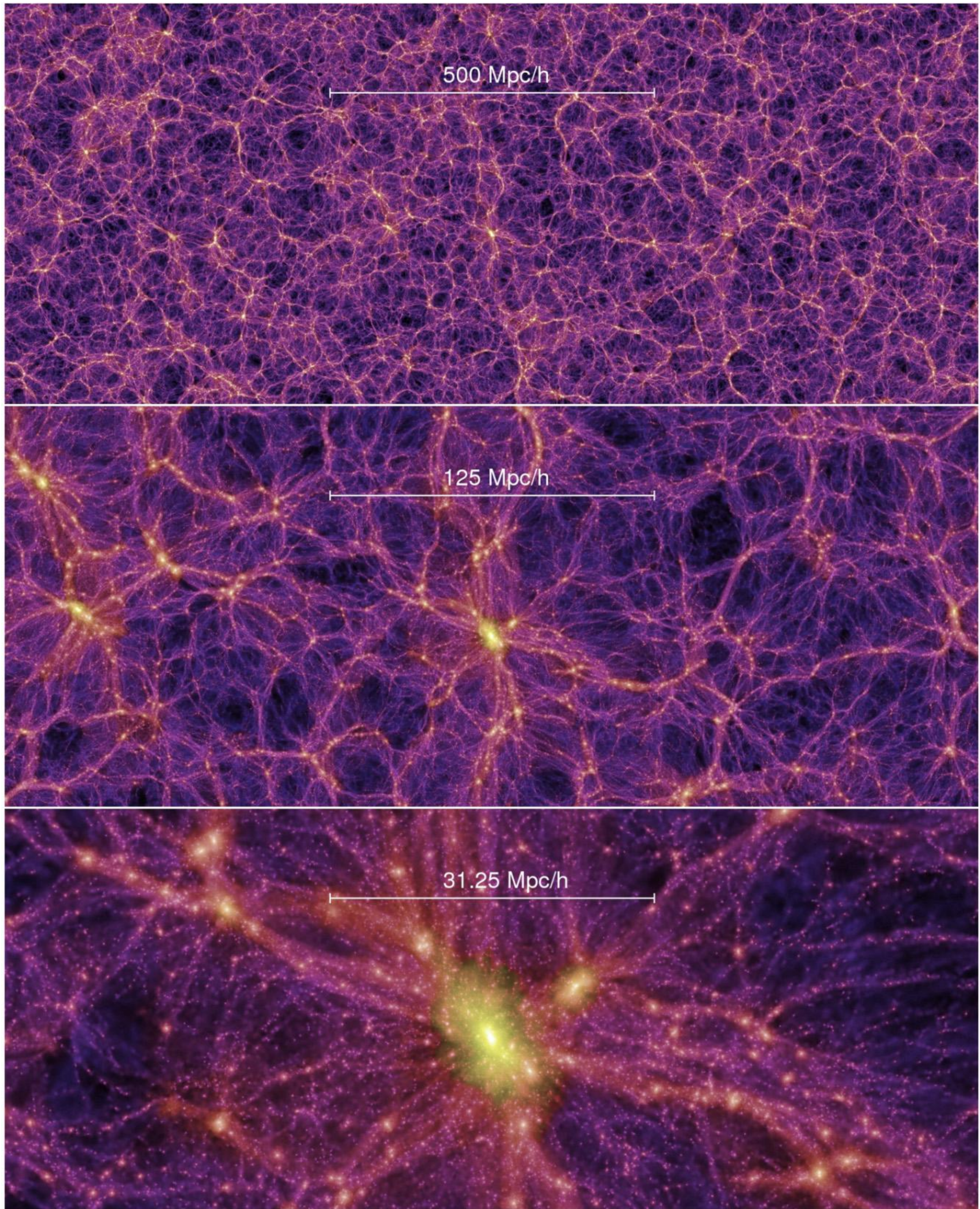


FIGURE 12.6 Slices of width $15\,h^{-1}\,\mathrm{Mpc}$ through the density field at redshift zero in the *Millennium* N-body simulation which follows 10^{10} particles (i.e., phase-space elements). From top to bottom, the different panels zoom in to show the hierarchical nature of the matter distribution in a ΛCDM cosmology. The spatial scale is labeled in each panel. The color scale denotes density in logarithmic units. The simulations shown here are described in Springel et al. (2005).

Phenomenology of nonlinear structure

- Small-scale density fluctuations are largest: small-scale structure forms first
- Then, structure successively assembles to large-mass objects
- Topologically, we have
    - “3D:” voids - large underdense regions
    - “2D:” sheets, or “pancakes”
    - “1D:” filaments
    - “0D:” bound structures - halos

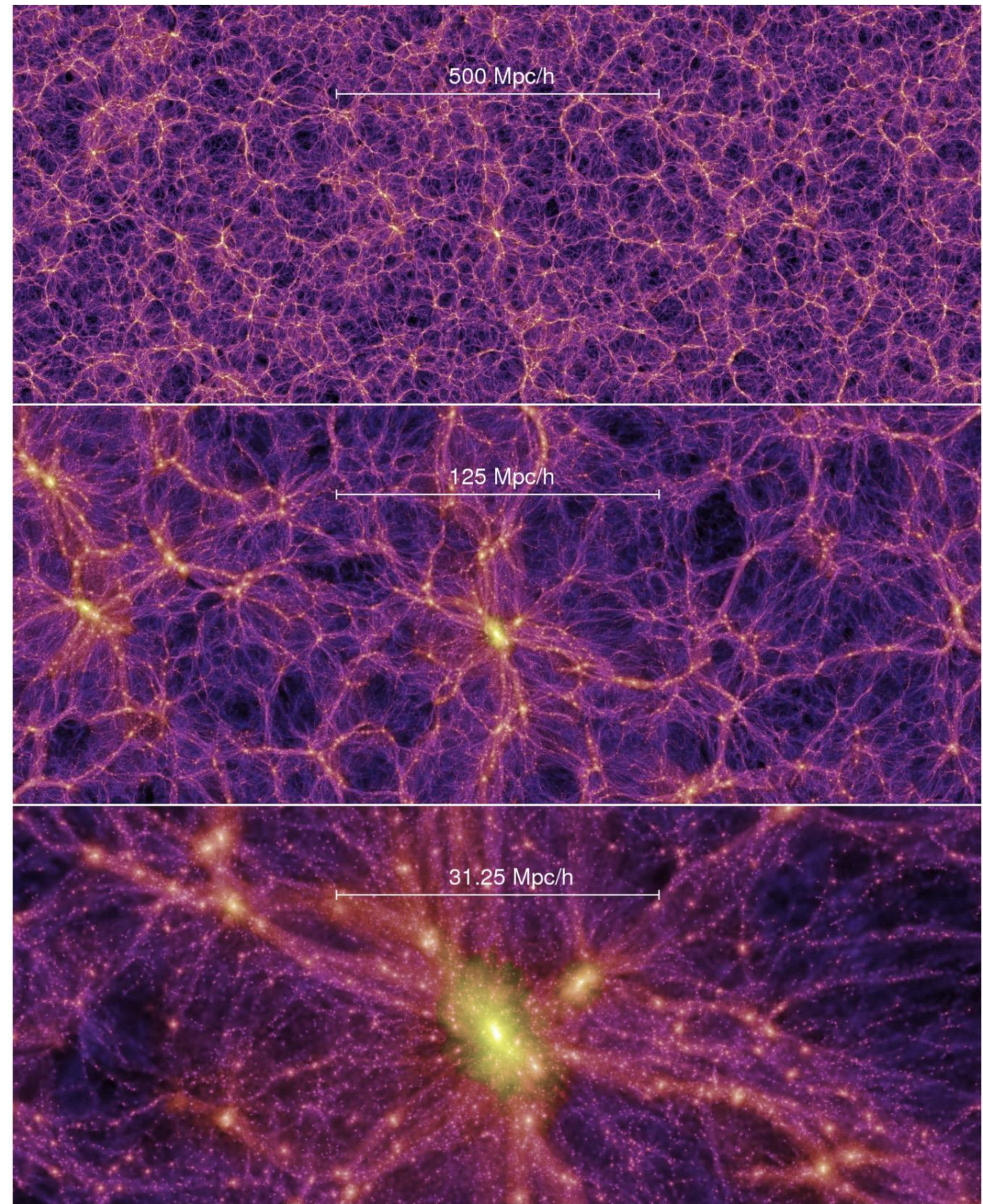


FIGURE 12.6 Slices of width $15\,h^{-1}$ Mpc through the density field at redshift zero in the *Millennium* N-body simulation which follows 10^{10} particles (i.e., phase-space elements). From top to bottom, the different panels zoom in to show the hierarchical nature of the matter distribution in a ΛCDM cosmology. The spatial scale is labeled in each panel. The color scale denotes density in logarithmic units. The simulations shown here are described in Springel et al. (2005).

Dark matter halos

- Bound structures of dark matter and baryons
- Densest regions in the universe (from a cosmologist's viewpoint…)
- All galaxies are believed to be hosted by dark matter halos
  - Strong observational evidence for this from *dynamics* (velocities of gas, galaxies) and *gravitational lensing*, both of which *probe all matter*
- The most massive halos are associated with galaxy clusters
- Still, halos are mostly studied as objects in simulations

Finding dark matter halos

- Halos are found using tools called *halo finders* which work on the catalog of particle positions
- Start from density maxima in the density field
- Determine whether particles are bound by comparing velocity w.r.t center of mass with local escape velocity
- Repeat this iteratively, since center of mass changes when particles are added
- Algorithms differ in detail

The issue of halo mass

- Strict definition, counting all particles that are bound, is not very practical: affected by numerical noise, and we don't observe dark matter anyway
- Definition based on maximum radius is more practical; however, no well-defined radius exists, since halo profiles smoothly transition to surrounding structure
- Instead, define mass and radius which enclose fixed density Δ times cosmic mean:

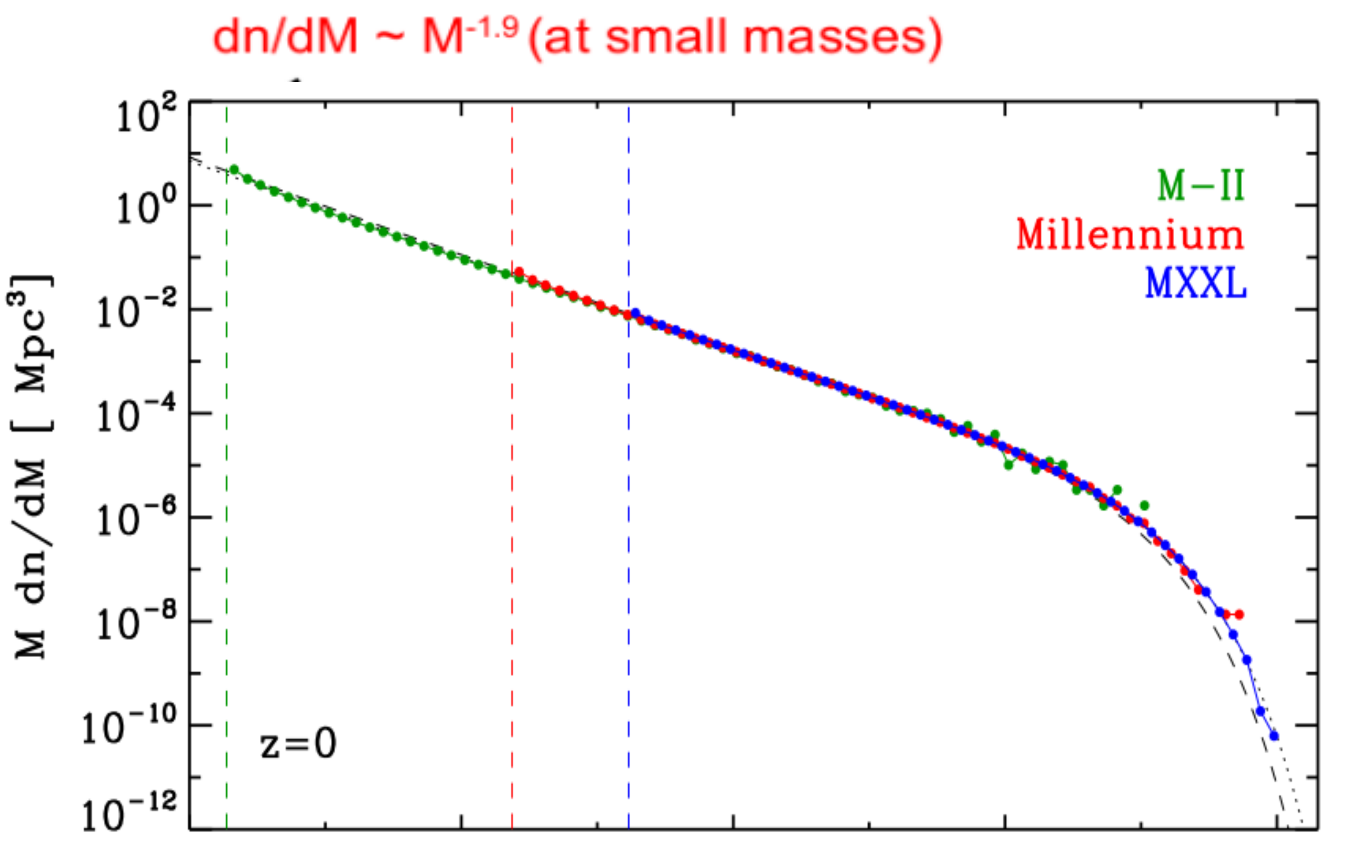
$$\frac{M(< R_\Delta)}{4\pi R_\Delta^3/3} = \Delta \times \rho_{\rm m}(t_0), \qquad M_\Delta = M(< R_\Delta)$$

- Special case Δ=1: *Lagrangian radius R_L*. Comoving size of region from which particles originated in the initial conditions. Important!

$$M = \frac{4\pi}{3}\bar{\rho}_{\rm m}(t_0) R_L^3$$

Halo abundance

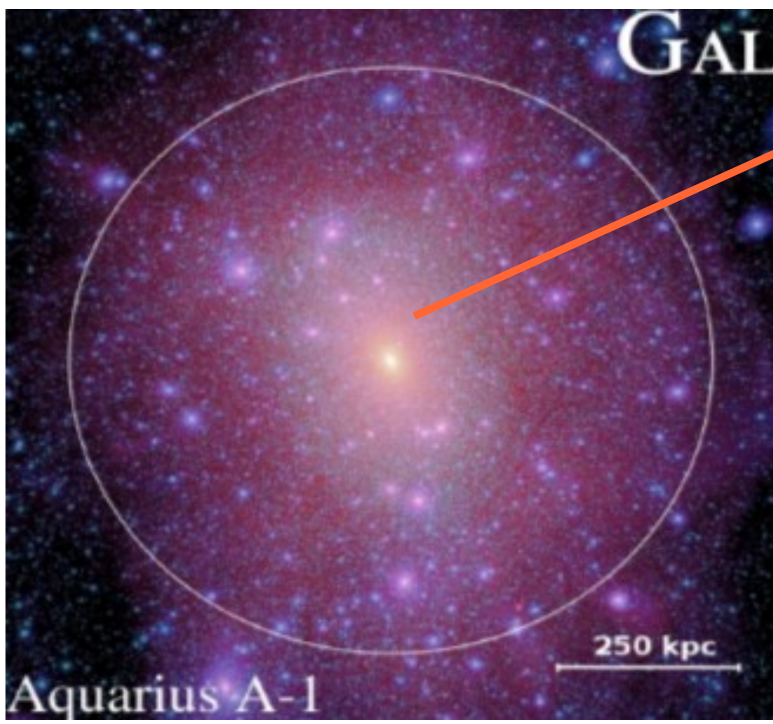
- Mean number density of halos in logarithmic mass bins
- Power-law at small masses
- Exponential cutoff at high masses - reflecting Gaussian statistics of initial density field



Angulo et al 2012

Inner structure of halos

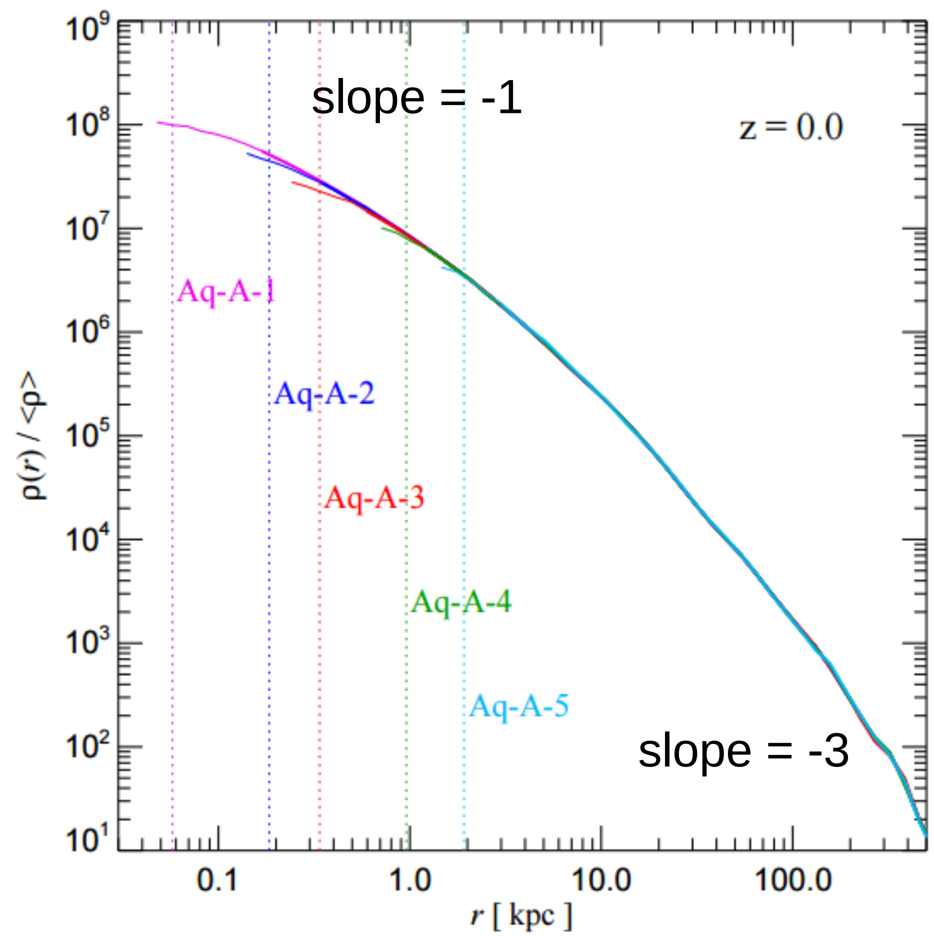
- Spherically-averaged density profile: Navarro-Frenk-White (1996) (NFW) form is universal



Springel et al 2008

Smooth distribution

Density profile is described by NFW/Einasto functional form, independent of mass, Cosmology, etc



Springel et al 2008

Eq. (12.62)

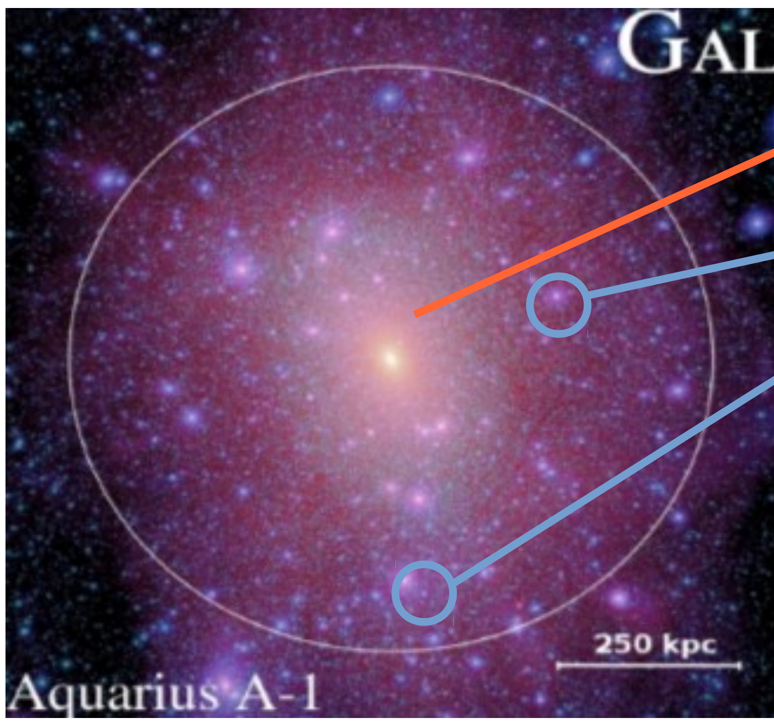
$$\rho(r) = \frac{\rho_s}{(r/r_s)(1+(r/r_s))^2}$$

$$\ln \rho(r)/\rho_{-2} = (-2/\alpha)\,(r/r_{-2})^{\alpha}$$

Slide credit: Raul Angulo

Inner structure of halos

- However, halos formed from smaller previous formed halos, which survive as substructure (*subhalos*)

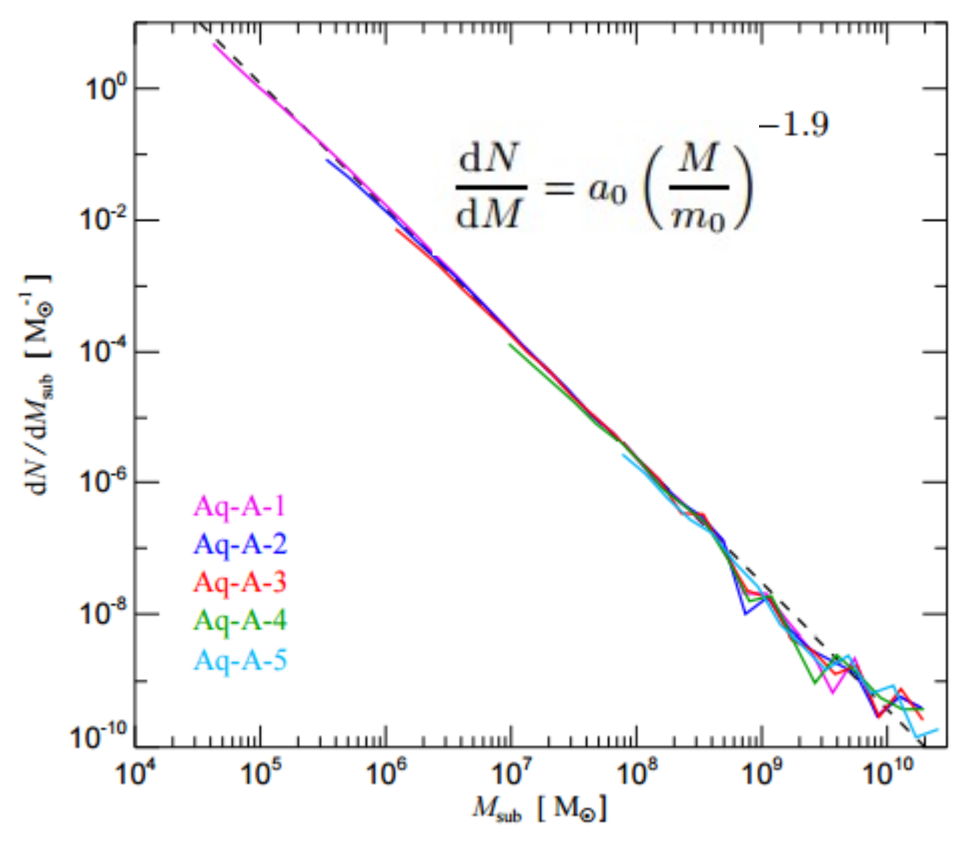


Springel et al 2008

Smooth distribution

Hierarchy of substructures

→ **Abundance**

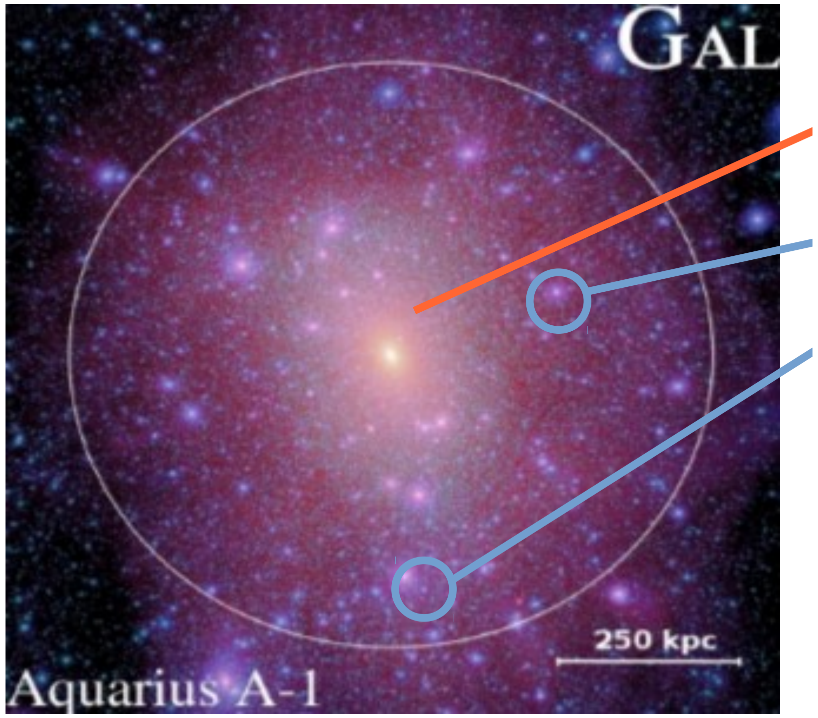


Springel et al 2008

Slide credit: Raul Angulo

From halos to galaxies

- We think that galaxies reside in these substructures of halos - but which ones…?
- Galaxy formation and (effective field) theory of galaxy clustering: next lecture!



Springel et al 2008