High Performance Computing for Cosmological Simulation

Jun Makino

Interactive Research Center of Science

Tokyo Institute of Technology

Actually...

Actually...

- **Not much cosmology...**

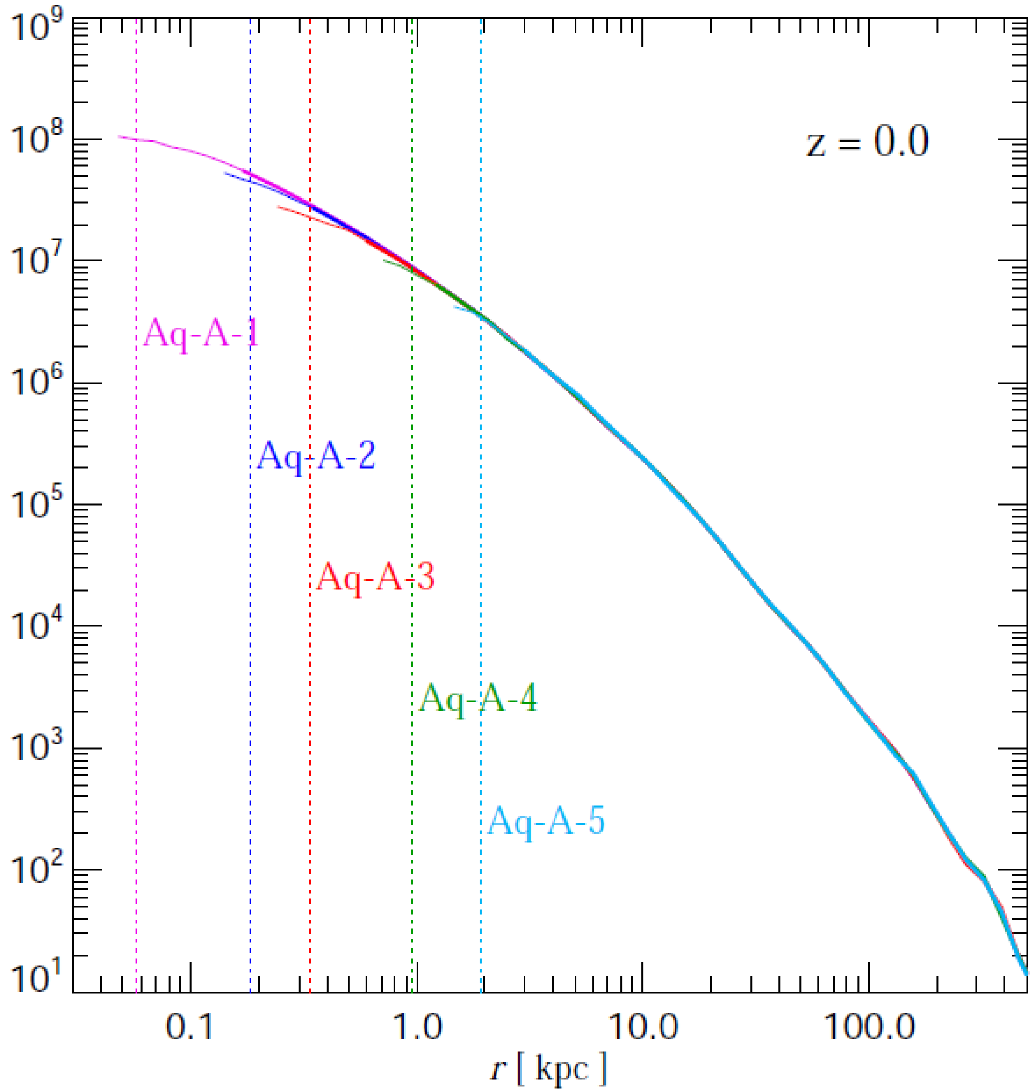
Actually...

- **Not much cosmology...**
- **Mostly on structure and formation of dark matter halos and galaxies.**

Structure of my talk

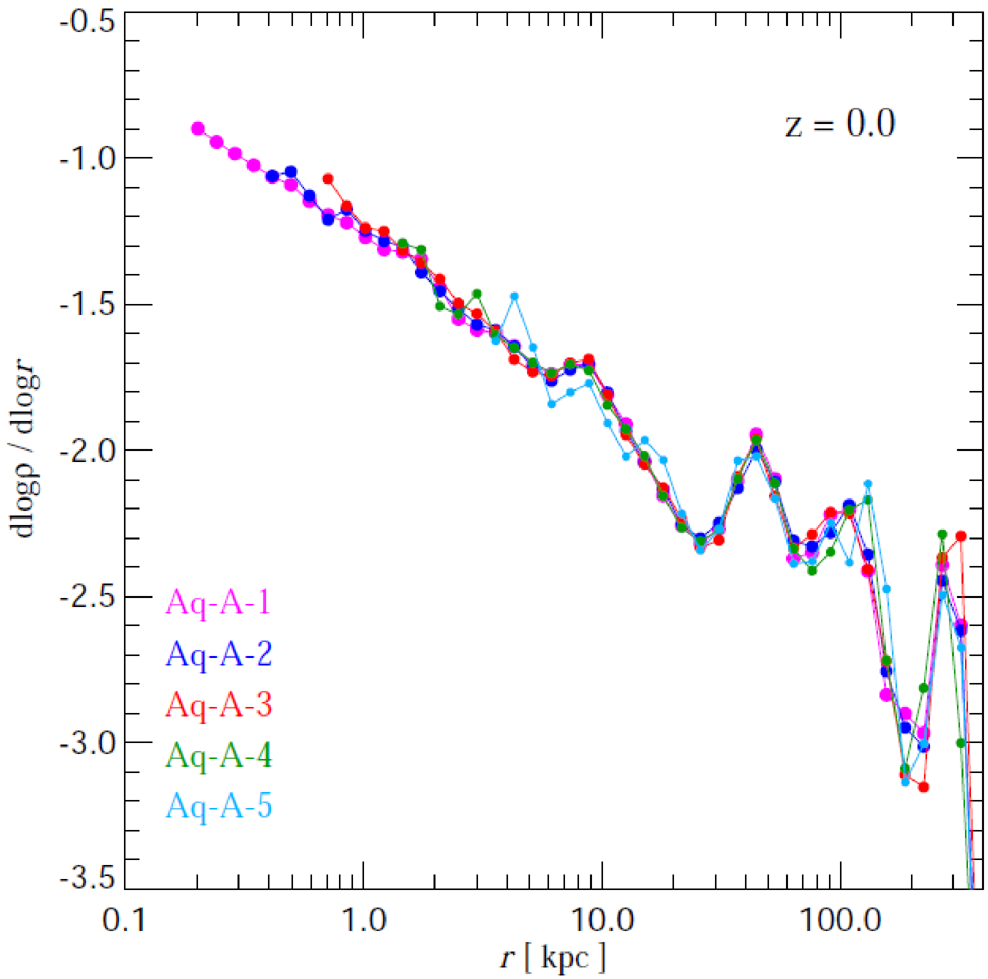
- **Dark Matter Halos**
- **Galactic dynamics**

Structure of DM halo



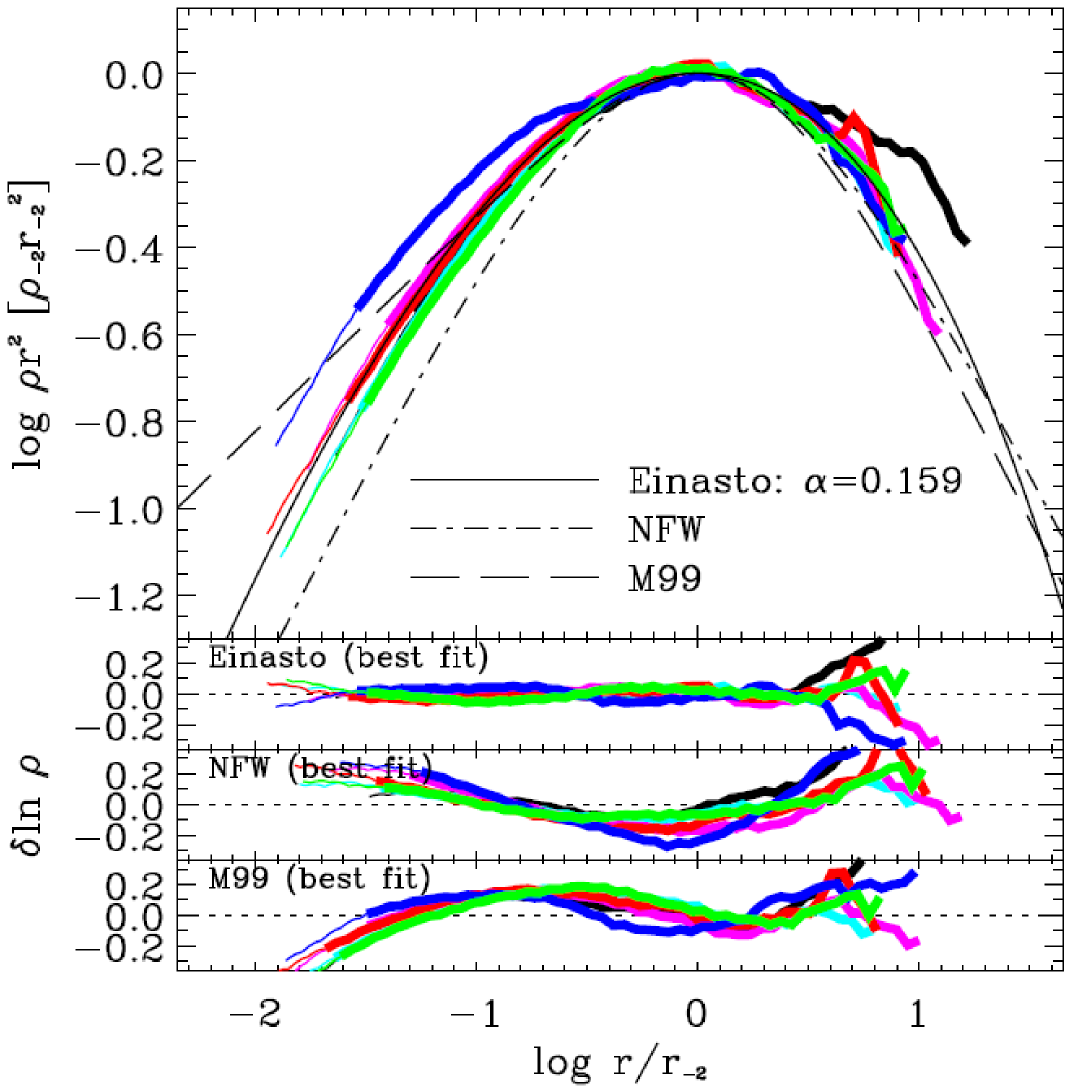
Highest-resolution run done so far : Springel et al (2008)
Mass resolution changed by three orders of magnitude
Sign of convergence?

Density Slope



Not a single power...

NFW and other profiles



Navarro et al. 2008

NFW: $\dfrac{1}{r(1+r)^2}$

Moore99:

$$\frac{1}{r^{1.5}(1+r^{1.5})}$$

Einasto:

$\exp[(-2/\alpha)(r^{\alpha}-1)]$

Historical perspective

- **NFW (1996): Density structure of DM halo is expressed by the NFW profile, independent of initial power spectrum or cosmological parameters.**
  - **Simulation used 10-20k particles/halo**
  - **Two-body relaxation affected the central structure**
- **Fukushige and JM (1997) : Central slope is steeper for 1M particle simulation.**
- **Moore et al. (1999) Slope is steep with 3M particles, Moore99 profile**
- **Now: $>$ 1B particles: NFW fit is not good, but slope becomes shallower at the very center.**

To summarize...

- **Central structure of DM halos is quite strange**
- **It is not a single power-law cusp. In simulations, slope becomes shallower as we improve resolution**
- **Numerical result is "reliable"**

Naively:

- **Initial density fluctuation is power-law**
- **No characteristic scale other than particle mass**

So why not a single power?

Theoretical difficulties

- **We cannot understand numerical result**
  - **Well, though numerical result is "reliable", it is pretty hard to believe...**
  - **Result depends strongly on resolution**
- **Isn't there something wrong in our simulations?**

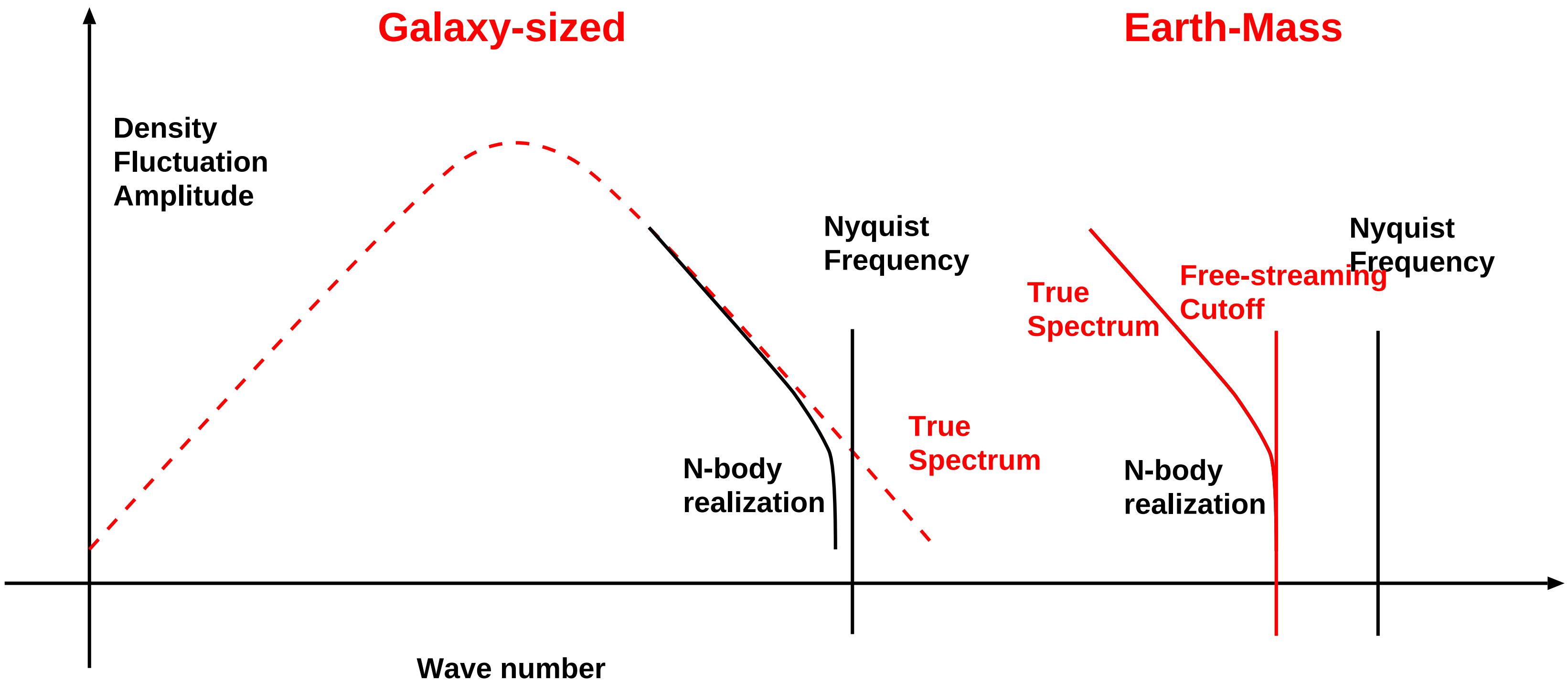
Fundamental problem?

Cosmological N-body simulations are not "correct" simulation of collisionless N-body systems.

Construction of initial condition:

1. Place particles uniformly
2. Add random fluctuations to position and velocity of particles according to the power spectrum of density fluctuation (cutoff: order of interparticle distance)

Smallest structures



Small halos form first in CDM = first "halos" contain $\sim$ 10 particles

Common belief: Since the process of hierarchical merging determines the structure, the structure of halos in the smallest scale should not affect the result.

No one has confirmed this belief, though.

Can we confirm this belief?

In principle, it is easy to confirm.

- **Fix the cutoff scale of initial power spectrum**
- **vary the mass resolution**

Practical difficulties:

- **Requires huge amount of computing resources.**
- **Not clear if the result is of any scientific interest...**

Smallest-mass halos

Actually, it might have some scientific importance:

Cold dark matter has free-streaming cutoff

Typical mass scale (depending on the nature of DM particle): Earth mass

Smallest halo: Earth mass, size 100AU.

- **Structure of these halos?**
- **Do they survive in galaxies?**

Importance of the structure of smallest halos

Primary question: Have they survived?

Important for detection of CDM particles

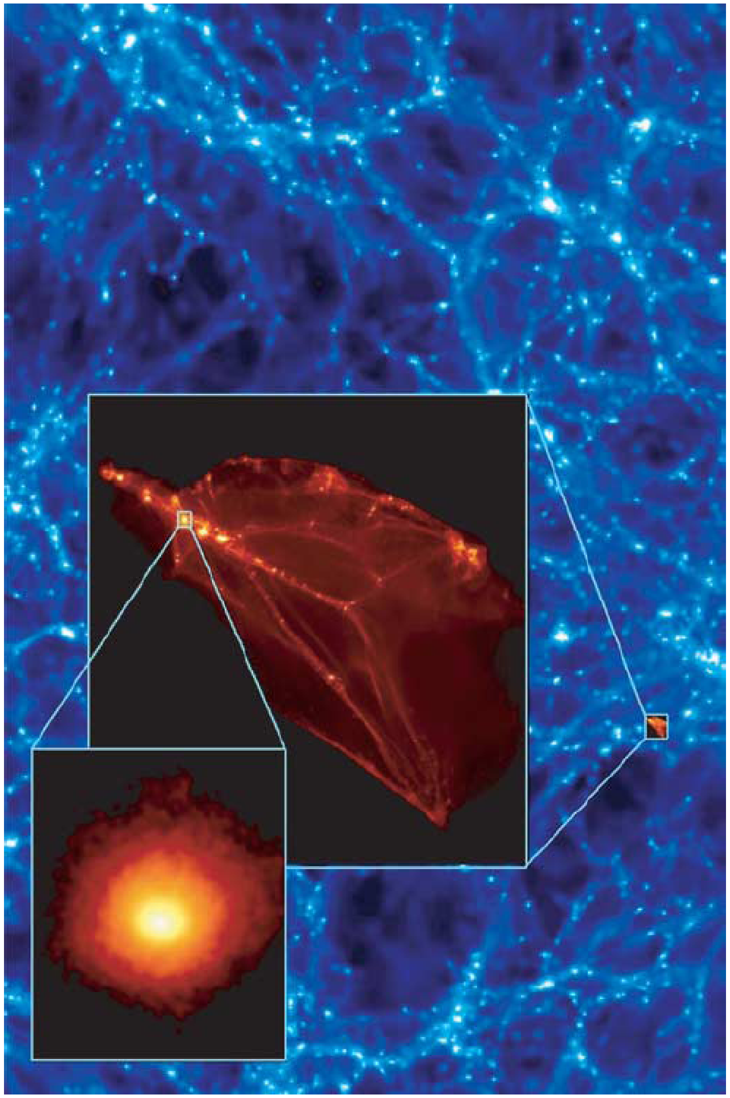
- **Direct detection: Large fluctuation in DM density**
- **Annihilation γ-ray: If survived, they dominate the flux.**

Processes which affect the survival

- Absorbed by larger halos
- Disrupted by potential of larger halo
- Disrupted by encounters with stars

Central structure is critical.

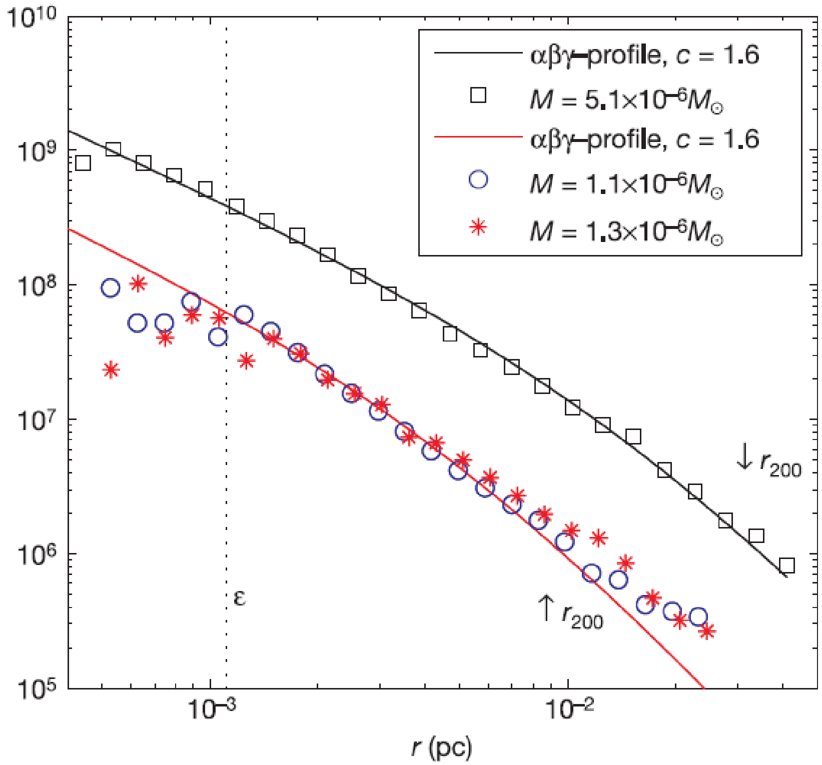
Previous work(s)



Diemand et al. 2005, Nature 433, 389

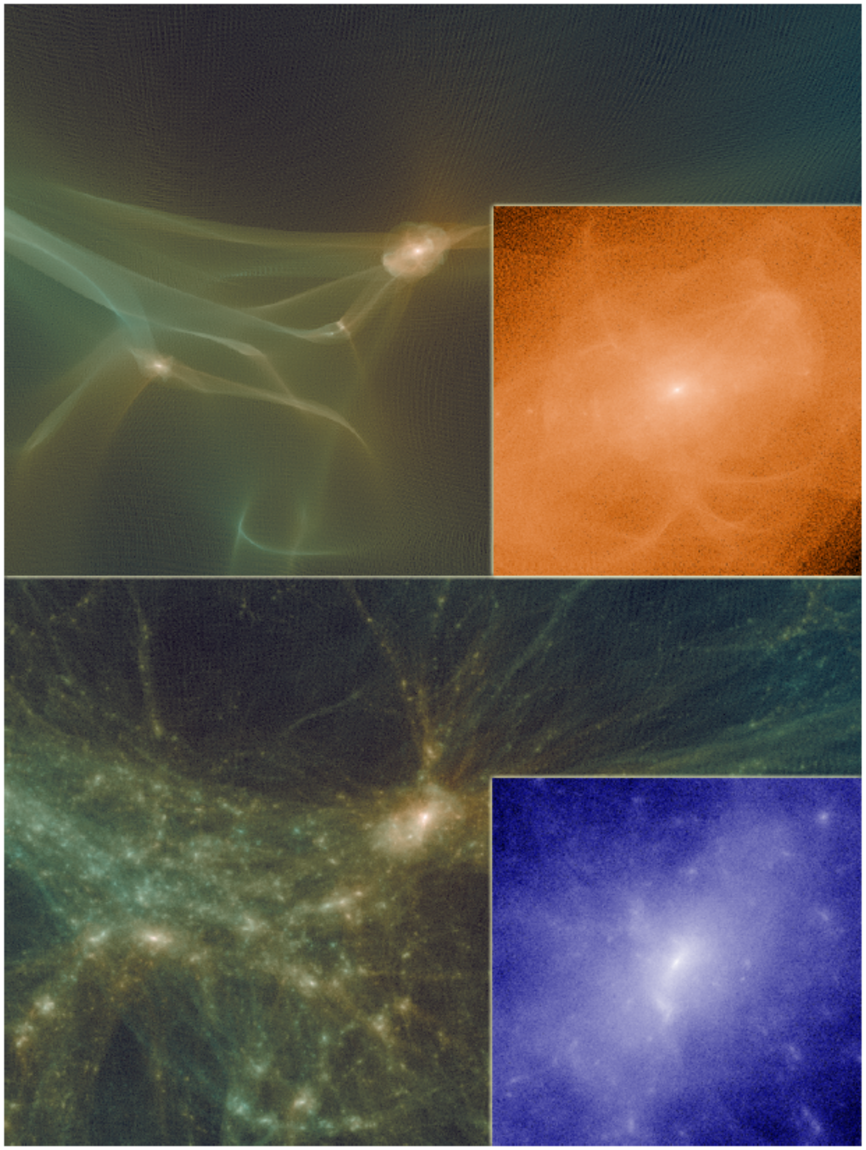
Usual Cosmological simulation
10^4 particles for Earth-mass halo

Density Profile



- **Essentially same result as NFW(1996)**
- **Quite natural because of low resolution**
- **Most likely completely wrong**

Ishiyama et al 2010



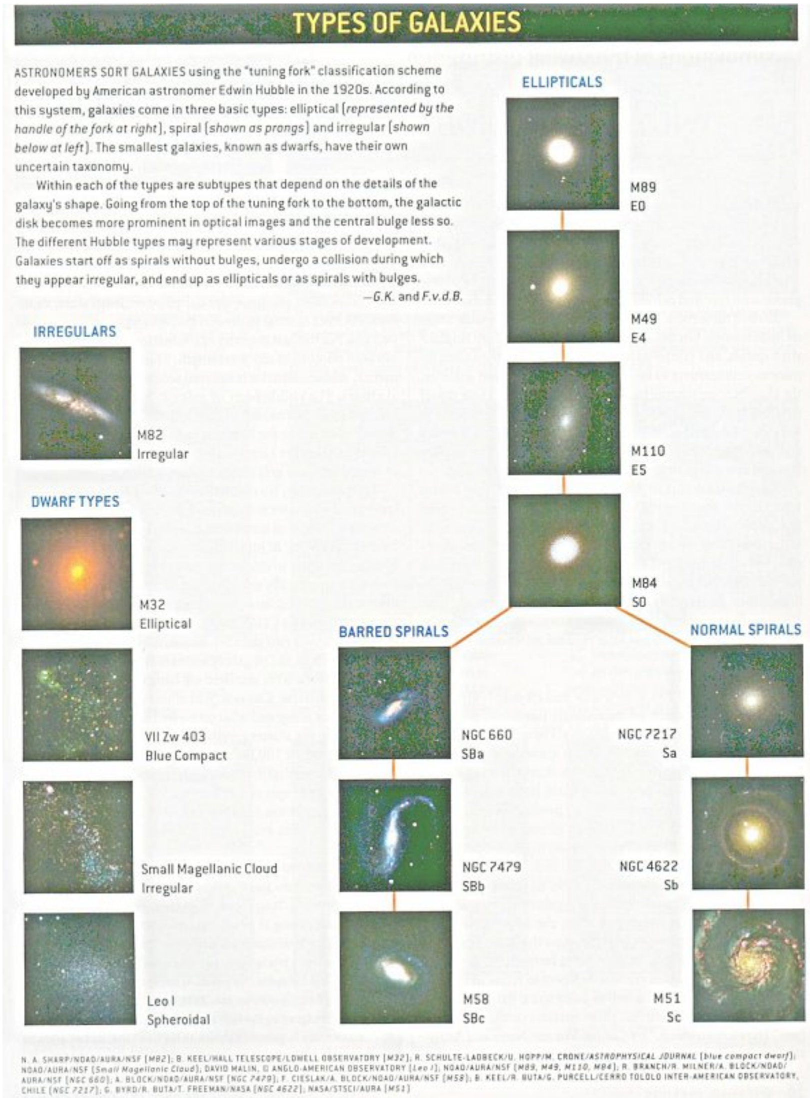
Ishiyama et al., 2010

100 times more particles compared to Diemand et al.

- **Top: with cutoff**
- **Bottom: no cutoff**

You will hear more from Tomo Ishiyama (next talk)

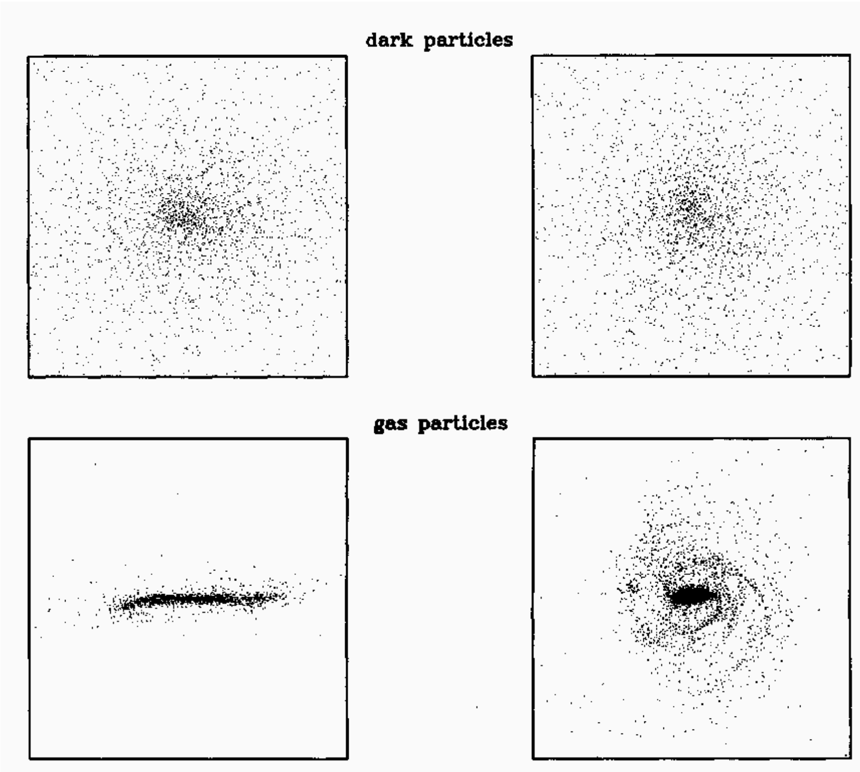
Simulation of galaxy formation



Basic Idea:

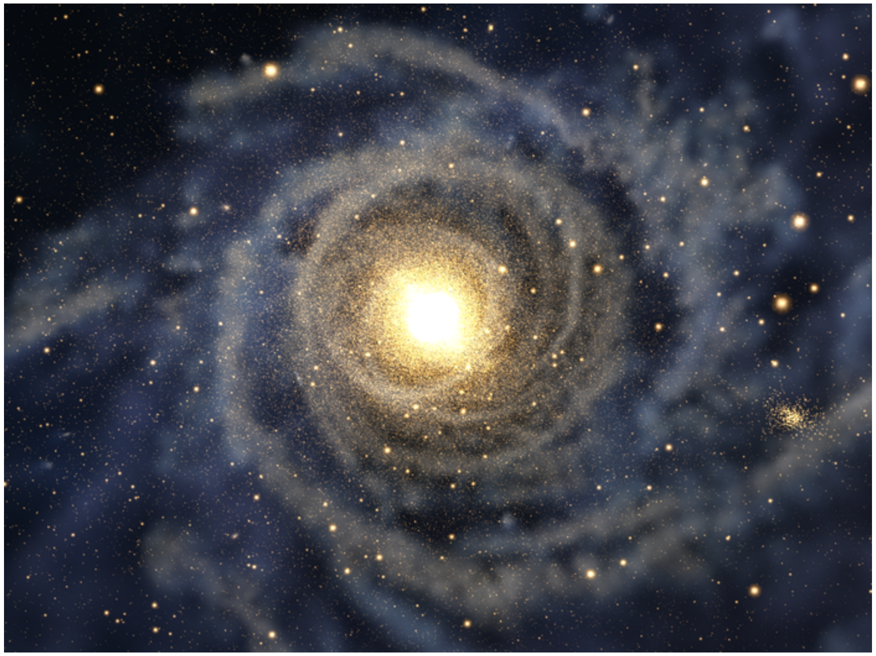
- "Holistic" simulation of galaxy, from initial density fluctuation
- To understand the origin of the variety of galaxies

Katz and Gunn 1992



- Dark Matter + gas + stars
- DM, star: particles
  gas :SPH particles
- 10^4 particles, Cray YMP 500-1000 hours
- mass resolution : 10^7 solar mass

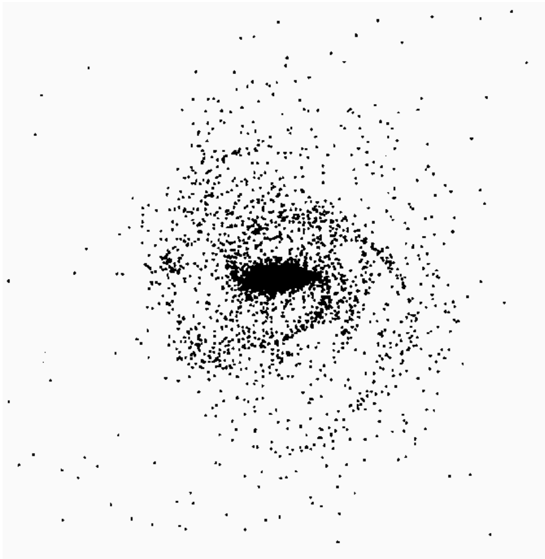
Saitoh et al. 2005

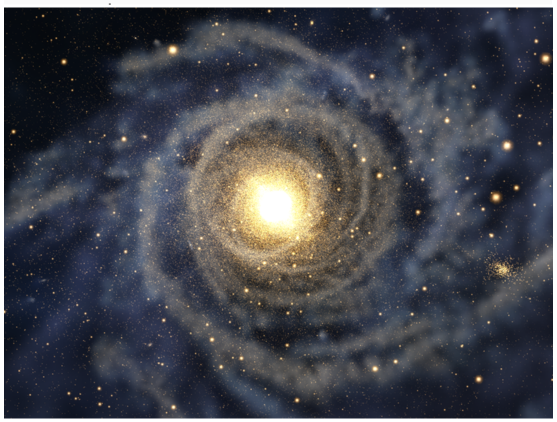


animation

- Dark Matter + gas + stars
- DM, stars: particles gas:SPH particles
- 2×10^6 particles, GRAPE-5 $\sim$ 1 year
- mass resolution : 10^4 solar mass

What gain from improved resolution?





- Not much?
- Important things: improved parametrization of "micro-physics", such as star formation mechanism, energy input from supernovae.

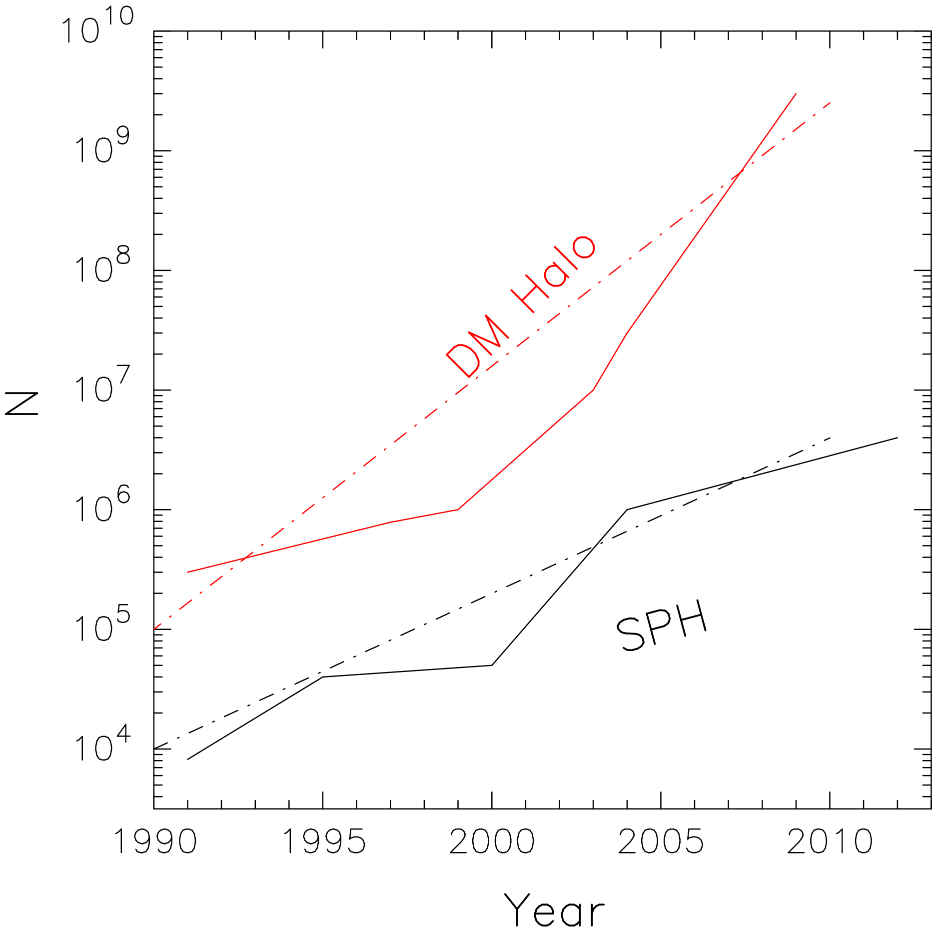
Modeling star formation

- Minimum need for star formation modeling: : 10^{-4} solar mass (probably we need much higher resolution...)
- What we can do now: : 10^{3} solar mass (10^{7} times too large)
- Need some way to form stars
  - Usual model: if interstellar gas is dense and cold enough, part of it will become stars in appropriate timescale.
  - three free parameters
  - The structure of the galaxy depends on these parameters
- Similar problem on supernovae.

What resolution do we need?

- We will know when we reach there....
- If mass of SPH particles is more than that of GMCs, clearly we are not doing things right.
- Theoretically, if we have sufficient resolution, we can just change all mass to stars (that is what the nature does).
- We are approaching there. -
- One or two orders of magnitude more?

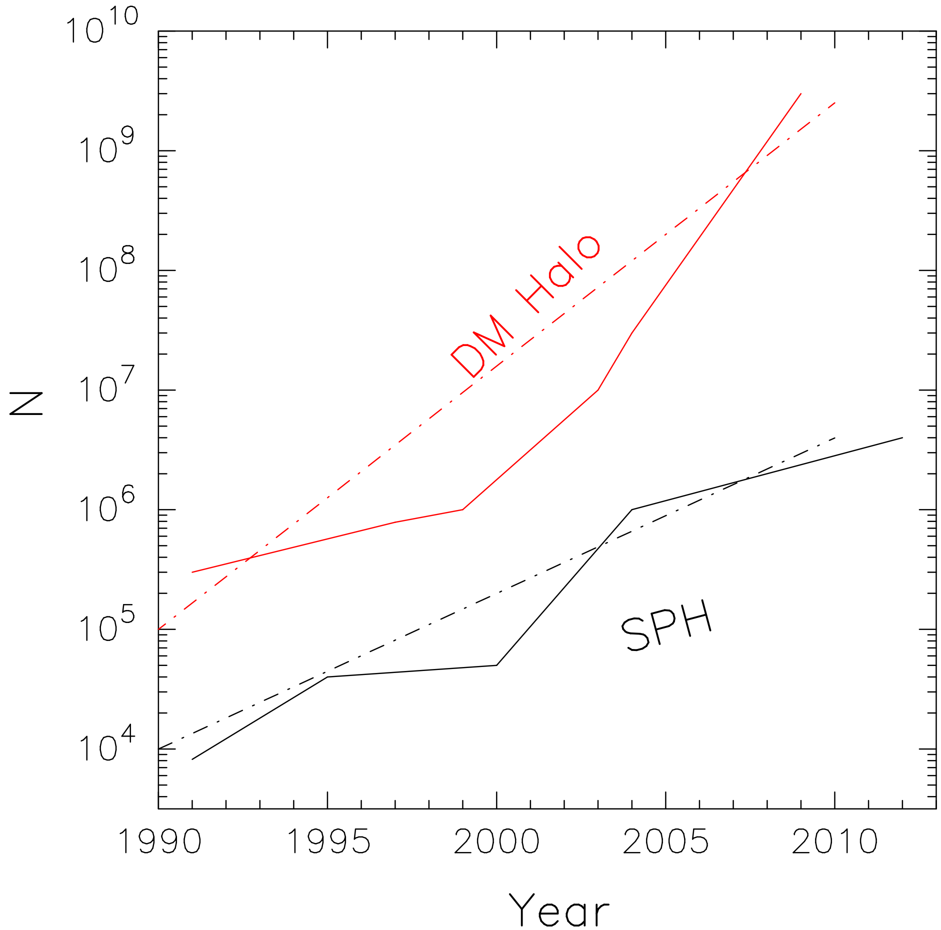
The History



Nbody DM halo:
10^4 increase in 20 years

SPH:
10^2 increase in 20 years

The History



Nbody DM halo:
10^4 increase in 20 years

SPH:
10^2 increase in 20 years

The Growing Nbody-SPH Gap

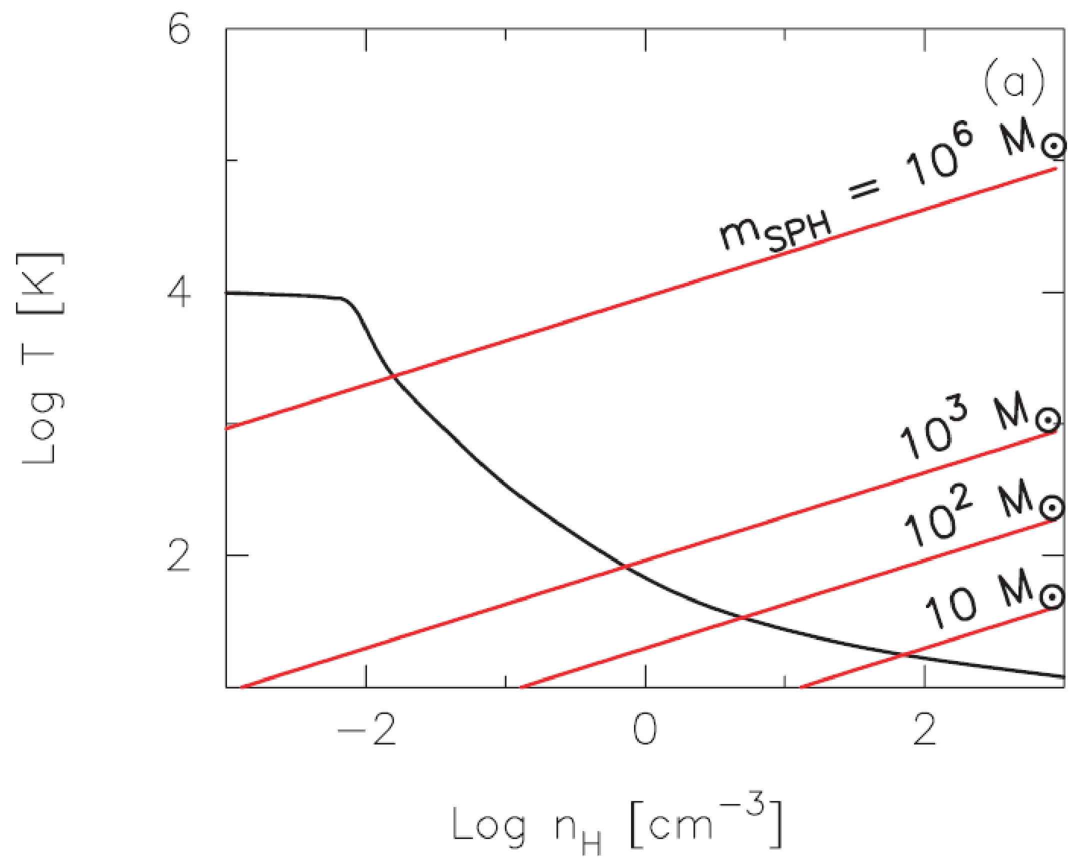
Why is N_{SPH} slow to increase?

Basic reason:

High resolution $\rightarrow$ small timescale

- **freefall timescale**
- **SNe feedback**

Freefall timescale



Not so severe:

1000 more particles → 10 times more steps.

SNe feedback

Timestep further decreases by a factor $\propto \sqrt{T_{SN}/T_{ISM}}$.

- **With higher resolution, T_{ISM} drops.**
- **When the mass of a star particle becomes less than $1000M_\odot$, one star particle might have less than one SN.**
- **When the mass of a gas particle reaches $100M_\odot$, timestep would reach ~ 10 years...**
- **Existing parallel Nbody+SPH codes become very inefficient with 10 year timesteps.**

Any fix for this timescale problem

Conventional “fix”: Prevent gas cooling below 10^4K.

- small structures are inhibited
- not much reason to go to higher resolution

A more “honest” way: develop new algorithms to tackle the timescale problem

What we have been trying for last several years.

What did we achieve

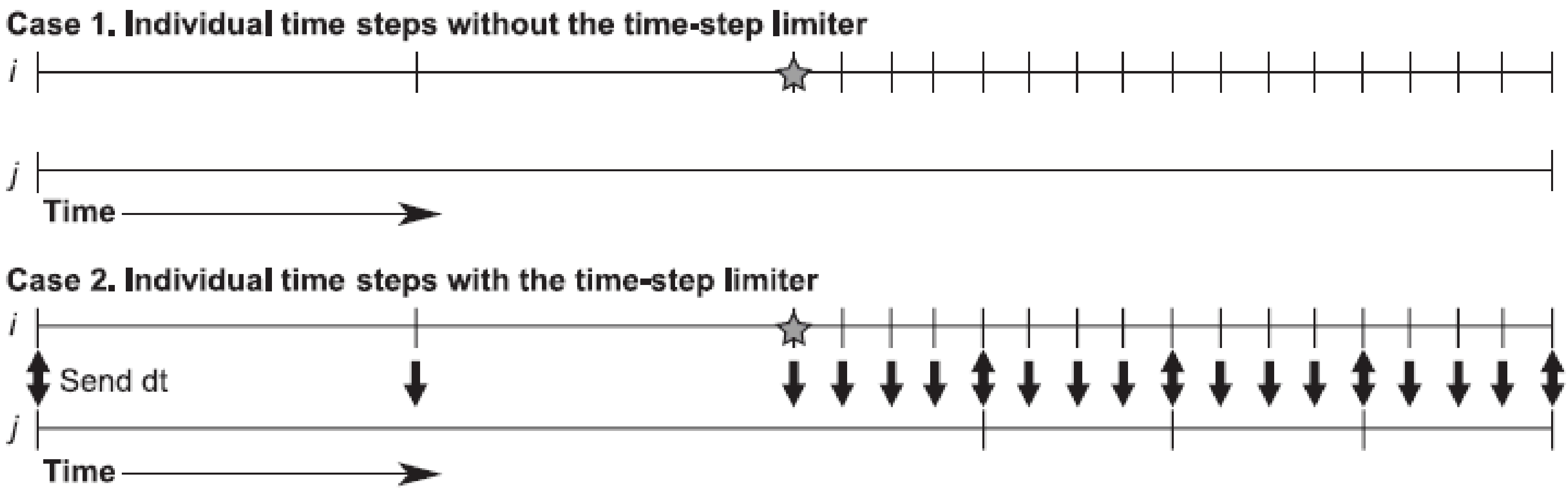
Makino and Saitoh 2012, PTEP 01A303

- Timestep limiter for individual-timestep SPH
- Asynchronous integration of SPH and gravity
- (New SPH to handle contact discontinuity)
- (New individual timestep with for treecode)

Timestep limiter

Saitoh and Makino 2009

One particle gets SNe energy (thermal or kinetic)



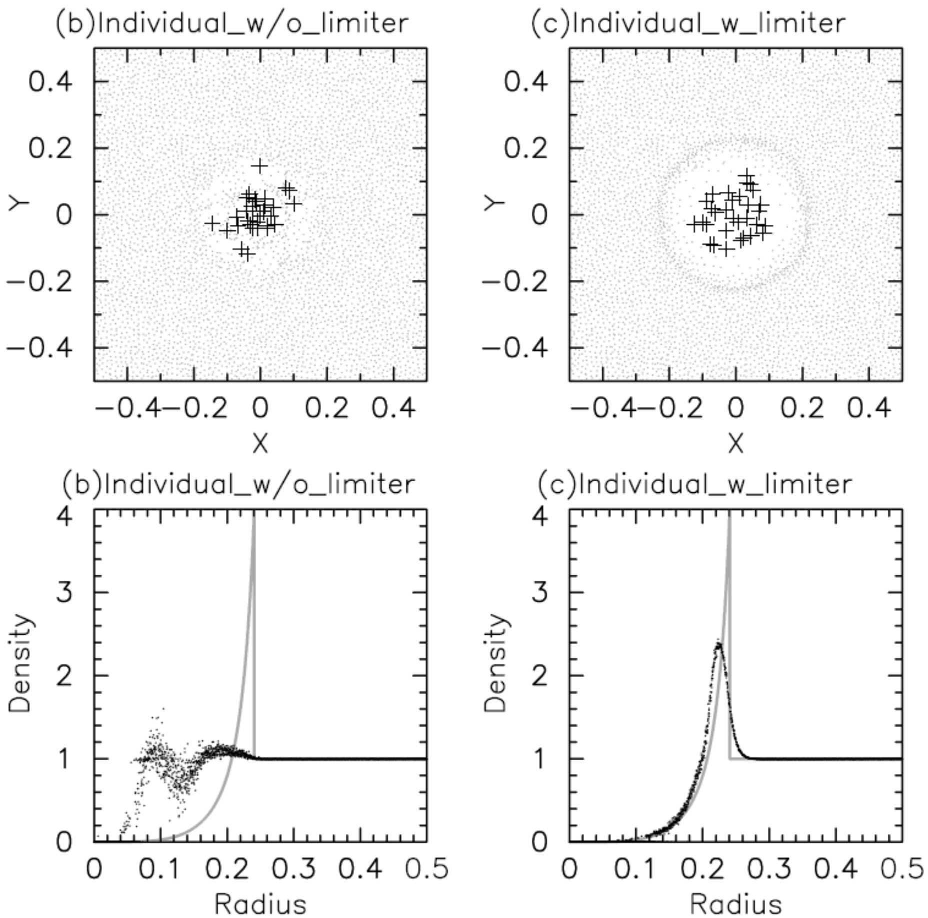
Neighbors should be pushed and heated up with the timescale of particle i.

Without limiter: neighbors cannot react for a very long time.

With limiter: neighbors can react.

Result of the limiter

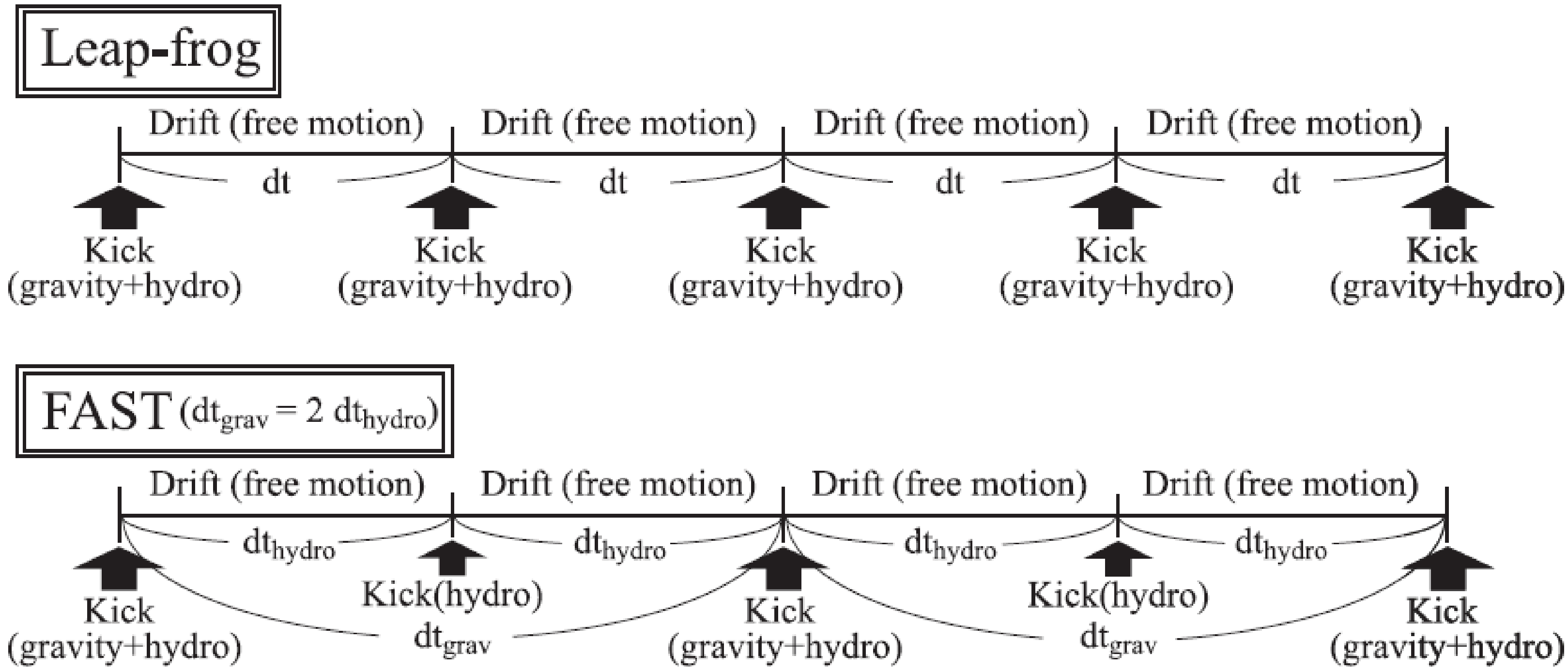
Point explosion experiment



Left: without limiter. Completely wrong result.

Right: with limiter. Good solution good agreement with non-individual timestep

Asynchronous integration of SPH and gravity



Decouple SPH interaction and gravitational interaction

Assign independent timesteps to them

Effect of asynchronous integration

- Can skip gravity calculation almost completely during early stage of SN remnant expansion
- In principle, (not in practice yet...) the calculation cost of SN feedback will become $O(N)$

New SPH to handle contact discontinuity

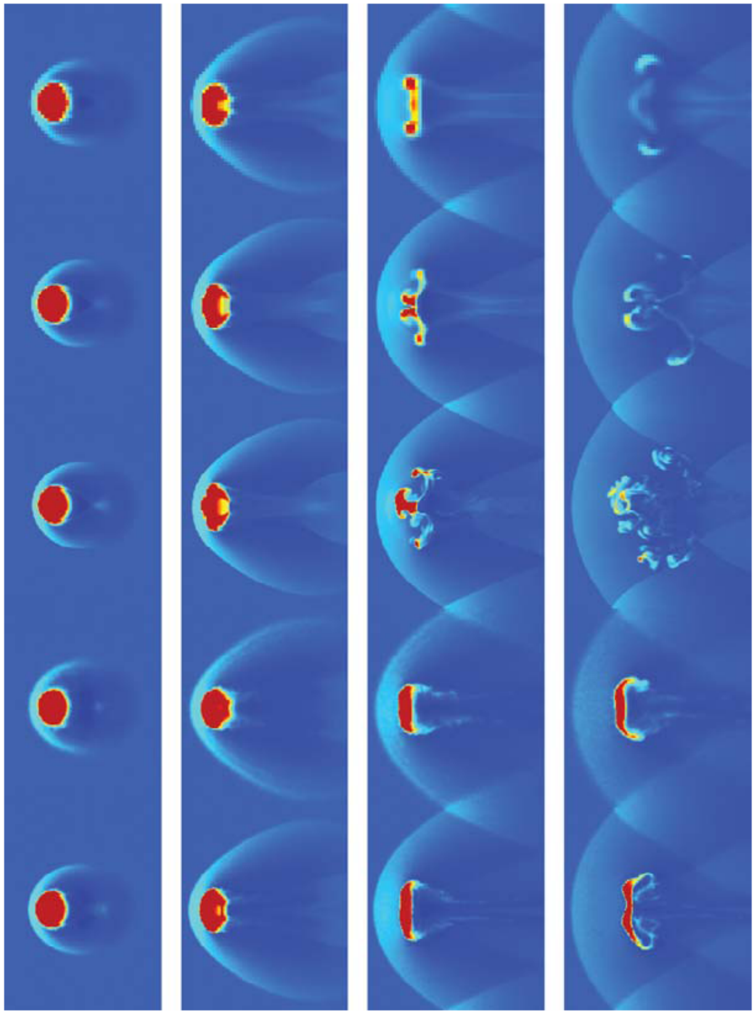
- SPH and Contact Discontinuity, KH instability
- Why does this happen?
- Problem with the standard SPH
- “New” SPH

SPH and Contact Discontinuity, KH instability

Agertz et al (MN 2007, 380, 963)

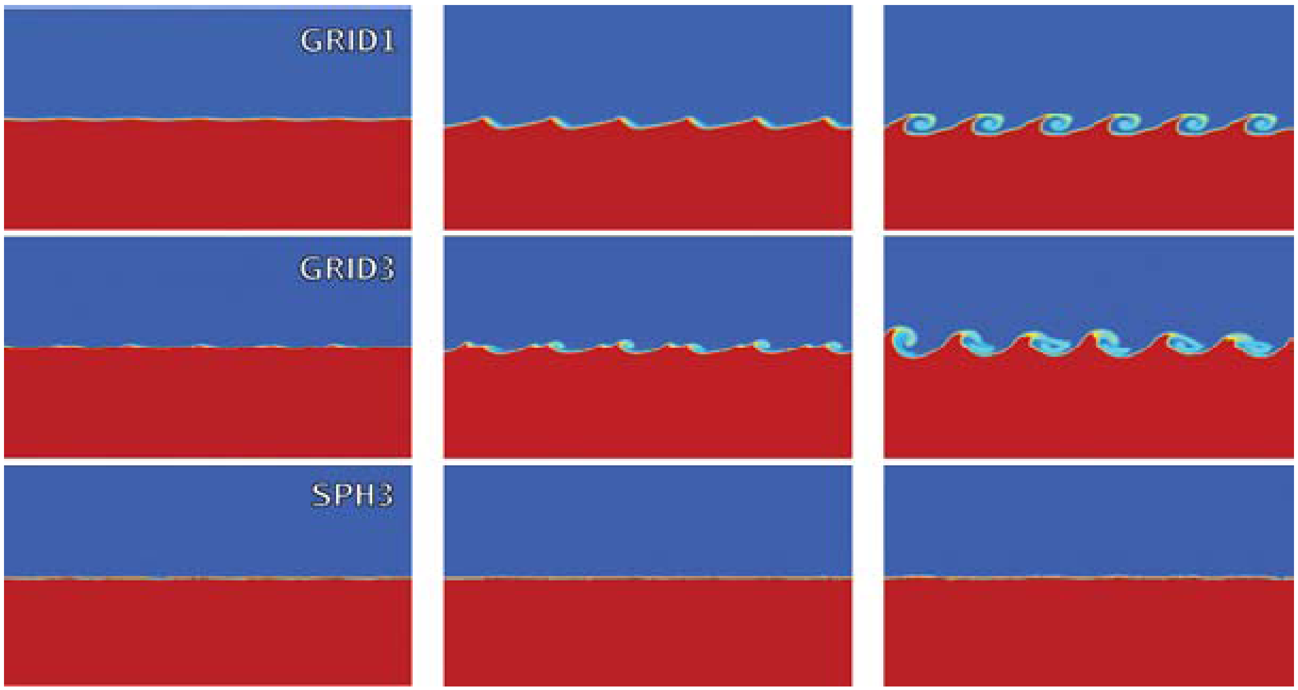
- **The result of a simple “Blob test” quite different on SPH Grid**
- **Kelvin-Helmholtz Instability is not correctly handled with SPH**
- **Is SPH usable?**

Difference (1)



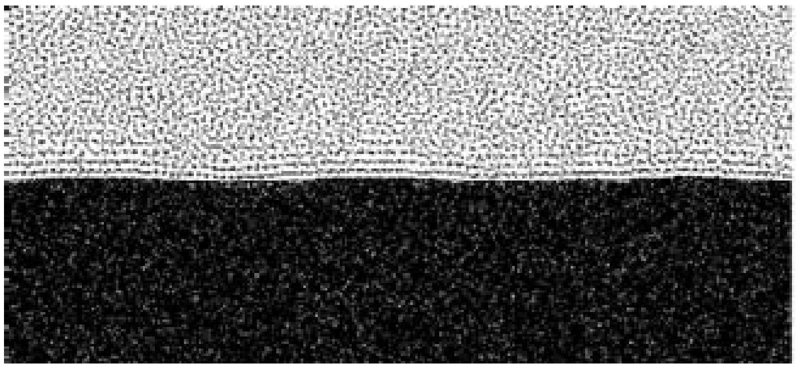
- Let a cold cloud (Temperature 1/10, density 10x) move with a supersonic velocity
- Upper three: Grid
- Lower two: SPH (1 and 10M particles)
- SPH suppresses the KHI at the fluid boundary

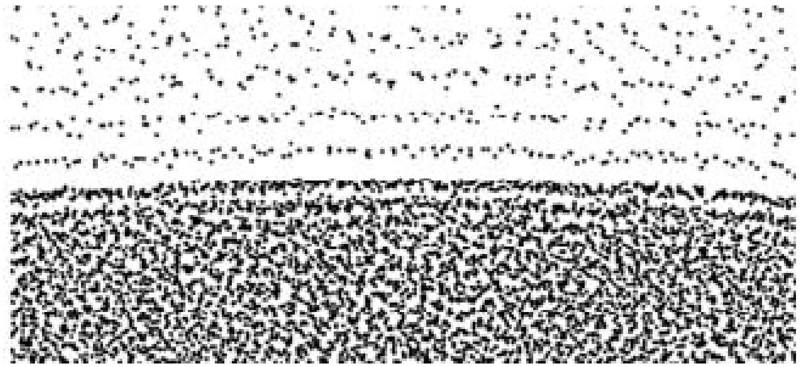
How different? (2)



SPH suppress KHI

How different? (3)





Strange-looking gap of particles at the two-fluid boundary.

Why does this happen?

Fundamental problem with SPH approximation

101 of SPH

Density estimate

$$\rho(x) = \sum_j m_j W(x - x_j), \tag{1}$$

Estimate of a quantity f

$$\langle f \rangle(x) = \int f(x') W(x - x') dx'. \tag{2}$$

101 of SPH continued(1)

grad of $\boldsymbol{f}$: $\langle \nabla f \rangle = \nabla \langle f \rangle$ use the following identity

$$1 = \sum_j m_j \frac{1}{\rho(x)} W(x - x_j). \tag{3}$$

and with a bit more approximation we have

$$\langle \nabla f \rangle (x) \sim \sum_j m_j \frac{f(x_j)}{\rho(x_j)} \nabla W(x - x_j). \tag{4}$$

101 of SPH continued(2)

Equation of motion evaluates $-\frac{1}{\rho}\nabla P$. Use the identity

$$\frac{1}{\rho}\nabla P = \frac{P}{\rho^2}\nabla \rho + \nabla \frac{P}{\rho^2}. \tag{5}$$

and symmetrize. The we have

$$\dot{v}_i = -\sum_j m_j \left(\frac{P_i}{\rho_i^2} + \frac{P_j}{\rho_j^2}\right) \frac{\partial}{\partial x_i} W(x_i - x_j), \tag{6}$$

Contact discontinuity

Standard SPH assumes the differentiability of ρ in the following two identities

$$1 = \sum_j m_j \frac{1}{\rho(x)} W(x - x_j). \tag{7}$$

$$\frac{1}{\rho}\nabla P = \frac{P}{\rho^2}\nabla\rho + \nabla\frac{P}{\rho^2}. \tag{8}$$

Density estimated with SPH is smoothed

- Density in the low- (high-) density side (near CD) is over- (under-)estimated,
- Therefore, pressure and its derivatives have $O(1)$ errors, and particles are redistributed.

Solution?

"Fundamental" reason

$\boxed{\rho \text{ is smooth but } u \text{ contains jump}}$

We could solve the problem by smoothing u. Several proposals

- Use kernel-estimated u
- Let u diffuse (artificial conductivity)
- Use density which is continuous at CD.

Sort of working, but not a "true" solution.

Our proposal: Basic idea

At CD, there is not jump in the pressure or internal energy. Only the density jumps. Why SPH approximation breaks down?

Because we use density to calculate other quantities.

$$\langle f \rangle(x) = \sum_j \frac{m_j f(x_j)}{\rho(x_j)} W(x - x_j). \tag{9}$$

We replace volume element dx by $m_j/\rho(x_j)$

In principle, ANY quantity should by okay as far as it gives correct estimate for the volume element.

Our proposal: Principle

What should we use instead of the mass density?

An ideal gas is described by the equation of state $PV = nRT$. Here, mass density does not appear. The RHS is the thermal energy.

Can't we use the pressure itself, which is equivalent to the energy density?

Each SPH particle has energy (or entropy). So we can evaluate pressure distribution without using mass density.

Pressure is continuous at CD. So there can be no large error.

Formulation (1)

Define internal energy per particle as

$$\boldsymbol{U_j = m_j u_j,} \tag{10}$$

($\boldsymbol{u}$ is per unit mass). Define the energy density as

$$\boldsymbol{q = \sum_j U_j W(x - x_j).} \tag{11}$$

Other quantities can be calculated as

$$\boldsymbol{\langle f \rangle(x) = \sum_j \frac{U_j f(x_j)}{q(x_j)} W(x - x_j),} \tag{12}$$

Spacial derivatives are given by

$$\boldsymbol{\langle \nabla f \rangle(x) = \sum_j \frac{U_j f(x_j)}{q(x_j)} \nabla W(x - x_j).} \tag{13}$$

Formulation (2)—Energy Equation

$$\frac{d\boldsymbol{u}}{d\boldsymbol{t}} = -\frac{\boldsymbol{P}}{\boldsymbol{\rho}}\boldsymbol{\nabla}\cdot\boldsymbol{v}. \tag{14}$$

The divergence of the velocity is given by

$$\boldsymbol{\nabla}\cdot\boldsymbol{v} = \sum_j(\boldsymbol{v}_i - \boldsymbol{v}_j)\frac{\boldsymbol{U}_j}{\boldsymbol{q}_j}\boldsymbol{\nabla}\boldsymbol{W}(\boldsymbol{x} - \boldsymbol{x}_j). \tag{15}$$

$\boldsymbol{P/\rho}$ **is calculated as follows. Using EOS**

$$\boldsymbol{P}_i = (\boldsymbol{\gamma} - 1)\boldsymbol{q}_i. \tag{16}$$

Formulation (3)—Energy Equation

The density appears since the LHS is per unit mass.

To rewrite this to per-particle form, use

$$\boldsymbol{\rho_i = \frac{m_i q_i}{U_i}}. \tag{17}$$

Then we have

$$\boldsymbol{\dot{U}_i = \sum_j (\gamma - 1) \frac{U_i U_j}{q_j} (v_i - v_j) \nabla W (x_i - x_j)}. \tag{18}$$

Formulation (4)—Equation of Motion

From Energy equation we derive EoM using energy conservation. Energy change of two particles, due to the interaction between them are

$$\dot{\boldsymbol{U}}_{ij} + \dot{\boldsymbol{U}}_{ji} = (\gamma - 1)\boldsymbol{U}_i\boldsymbol{U}_j\left(\frac{1}{\boldsymbol{q}_i} + \frac{1}{\boldsymbol{q}_j}\right)(\boldsymbol{v}_i - \boldsymbol{v}_j)\nabla\boldsymbol{W}(\boldsymbol{x}_i - \boldsymbol{x}_j). \tag{19}$$

This should be equal to the change of the kinetic energy

$$\frac{\boldsymbol{m}_i\boldsymbol{m}_j}{\boldsymbol{m}_i + \boldsymbol{m}_j}(\boldsymbol{v}_i - \boldsymbol{v}_j)(\dot{\boldsymbol{v}}_i - \dot{\boldsymbol{v}}_j). \tag{20}$$

Therefore, velocity change is

$$(\dot{\boldsymbol{v}}_i - \dot{\boldsymbol{v}}_j) = -(\gamma - 1)\frac{\boldsymbol{m}_i + \boldsymbol{m}_j}{\boldsymbol{m}_i\boldsymbol{m}_j}\boldsymbol{U}_i\boldsymbol{U}_j\left(\frac{1}{\boldsymbol{q}_i} + \frac{1}{\boldsymbol{q}_j}\right)\nabla\boldsymbol{W}(\boldsymbol{x}_i - \boldsymbol{x}_j), \tag{21}$$

Formulation (5)—Equation of Motion

Using the conservation of the center of mass we have

$$\boldsymbol{m_i \dot{v}_i = -\sum_j (\gamma - 1) U_i U_j \left(\frac{1}{q_i} + \frac{1}{q_j} \right) \nabla W(x_i - x_j).} \quad (22)$$

- **RHS does not depend on mass**
- **This form is symmetric (between i and j particles)**

(The resulted formulation is same as that in Ritchie and Thomas 2001)

Examples

Standard SPH1 New SPH1

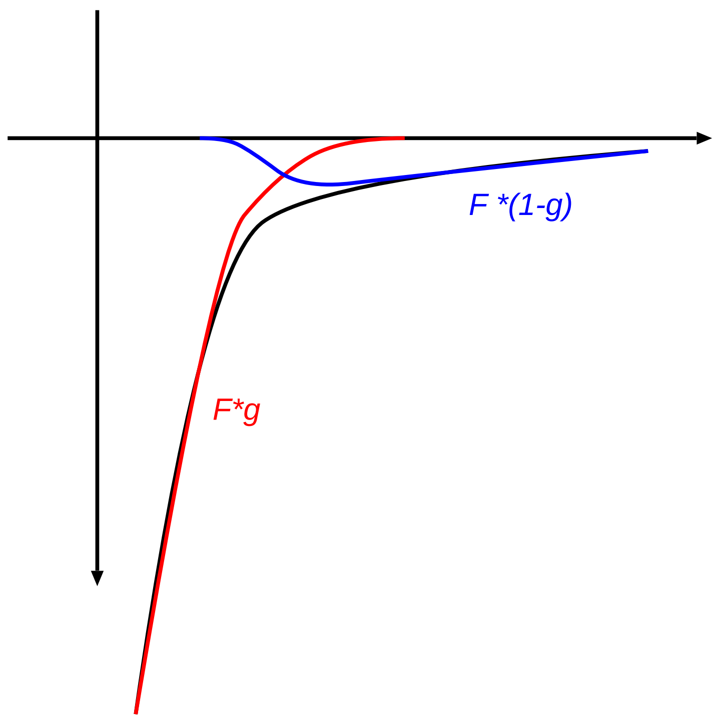
Standard SPH2 New SPH2

- **Seems to work fine**
- **Not ideal for strong shock or free surface (large pressure gap)**
-

New individual timestep with for treecode

Divide a pairwize interaction to near and far terms

$$\boldsymbol{F}_{ij} = -\boldsymbol{G}\boldsymbol{m}_i\boldsymbol{m}_j\frac{\boldsymbol{r}_{ij}}{|\boldsymbol{r}_{ij}|^3} = \boldsymbol{F}_{ij}(1 - \boldsymbol{g}(|\boldsymbol{r}_{ij}|) + \boldsymbol{F}_{ij}\boldsymbol{g}(|\boldsymbol{r}_{ij}|)$$



- **Integrate $\boldsymbol{F} * \boldsymbol{g} +$ kinetic energy with high-order individual timestep scheme**
- **Integrate $\boldsymbol{F} * (1 - \boldsymbol{g})$ with tree+leap frog**
- **$\boldsymbol{g}$ should have compact support, and $\boldsymbol{C}^4$**

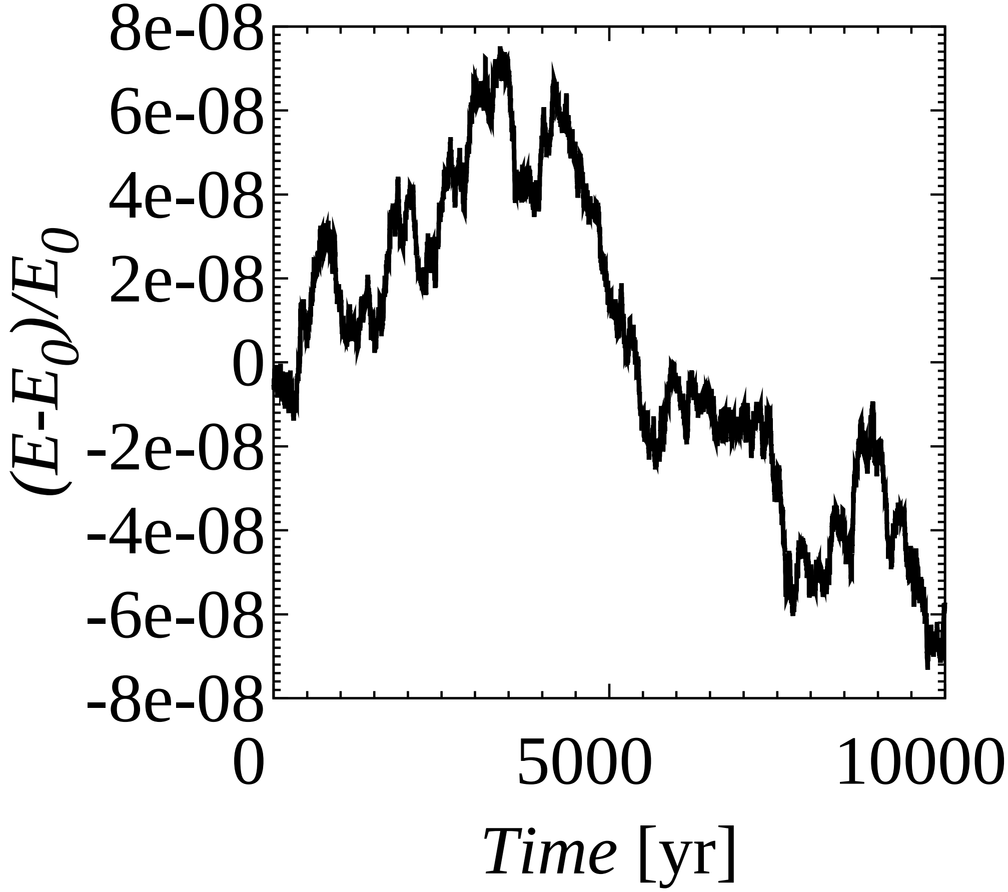
Implementation

Oshino et al. 2011, PASJ.

Many practical issues

- **Scale for g**
- **Accuracy and timestep for tree**
- **Timestep criteria for individual timestep**

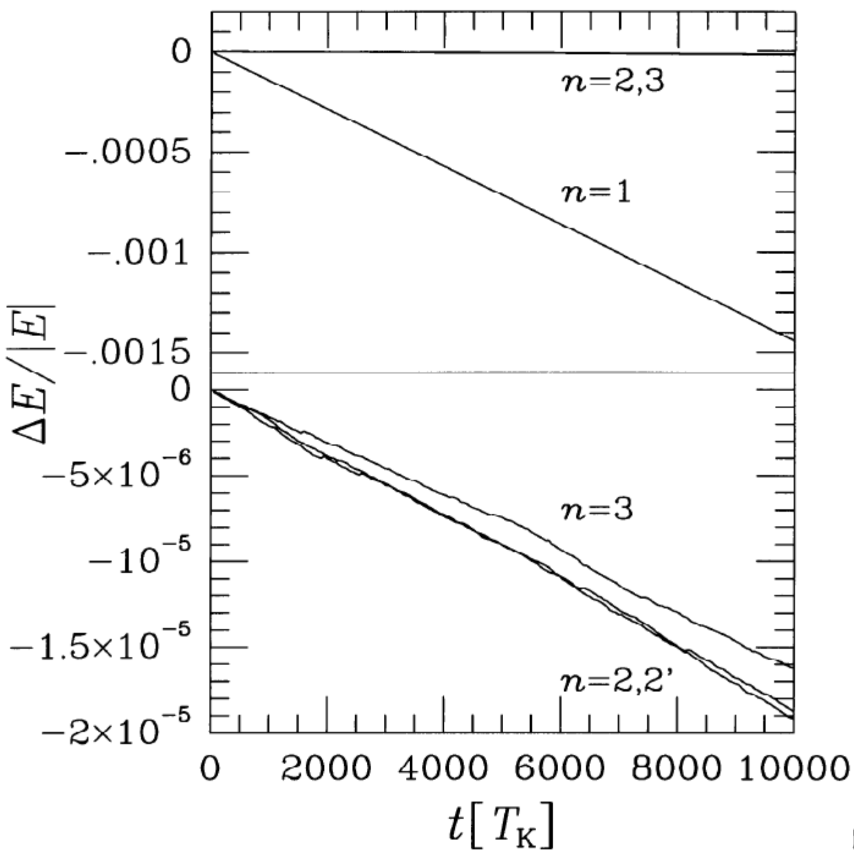
Energy Error



- **Planetary growth calculation similar to Kokubo and Ida 1995. 10,000 particles**
- **Collision and accretion implemented**
- **Energy very well conserved.**

How well?

Kokubo, Yoshinaga and JM, 1998



- **High-accuracy, semi-time-symmetric scheme (iteration applied to solar gravity only)**
- **energy relative error 2^{-5} after 10^4 years**
- **Oshino et al. gives error smaller by three orders of magnitude**
- **partly because larger N**

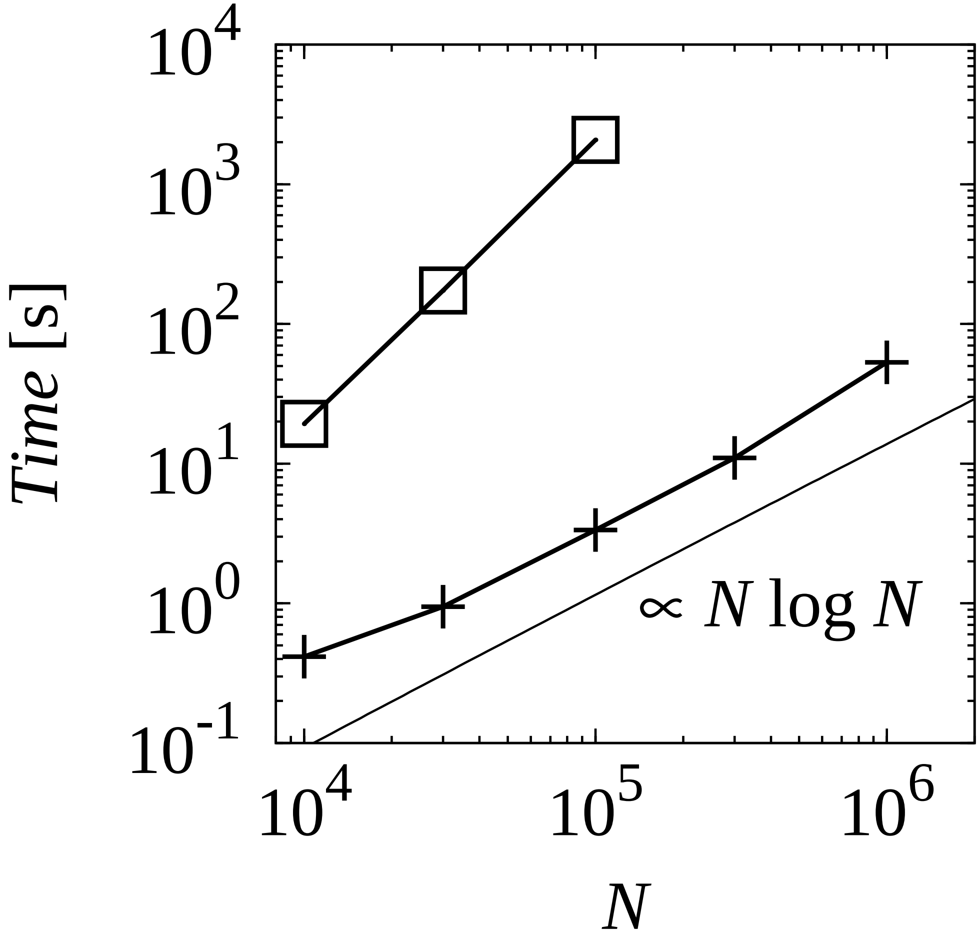
By the way, how many citations?

Kokubo and JM 2004: 8

Kokubo, Yoshinaga and JM 1998: 50

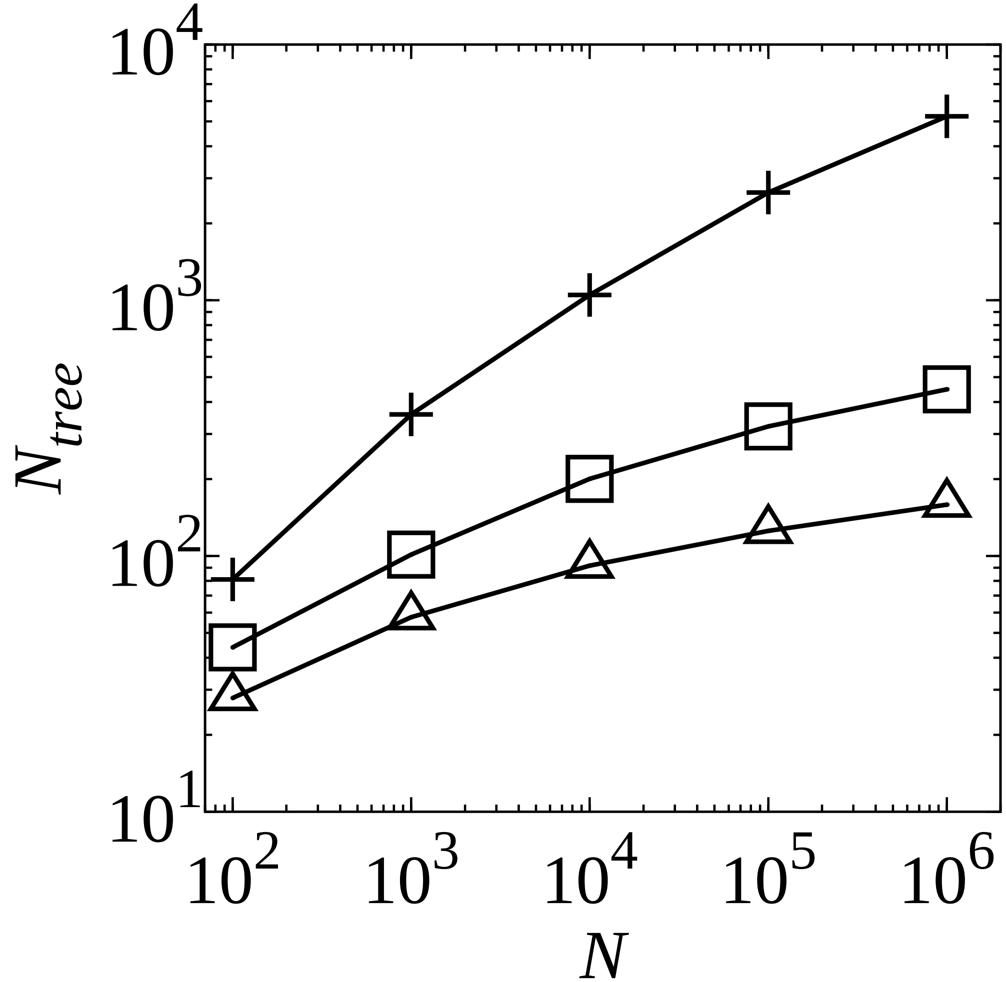
??????

CPU time



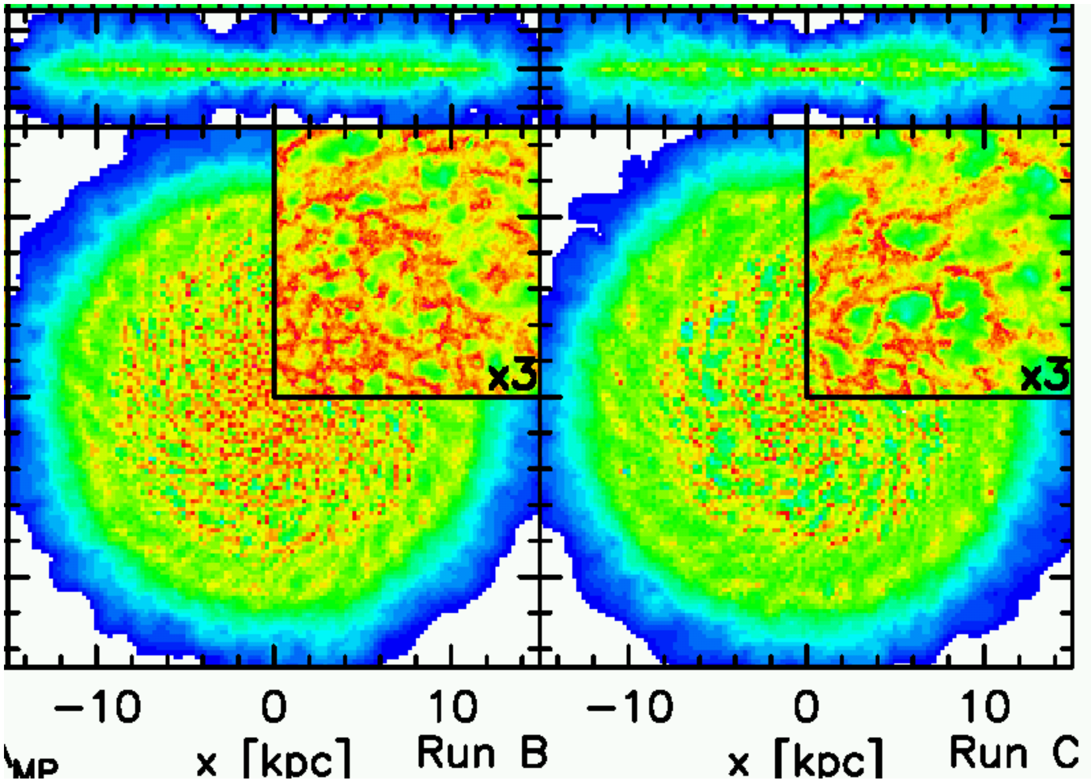
- **CPU time per one tree step**
- **Square: usual N^2 calculation**
- **Cross: New method**

Calculation cost



- Interactions per particle per tree step
- Opening angle: $\theta = 0.1, 0.5, 1.0$ from top to bottom
- Even with $\theta = 0.1$, calculation cost reduced by a factor of 200 for 1M particles

Saitoh et al. 2007



Changed the star formation timescale by a factor of 15

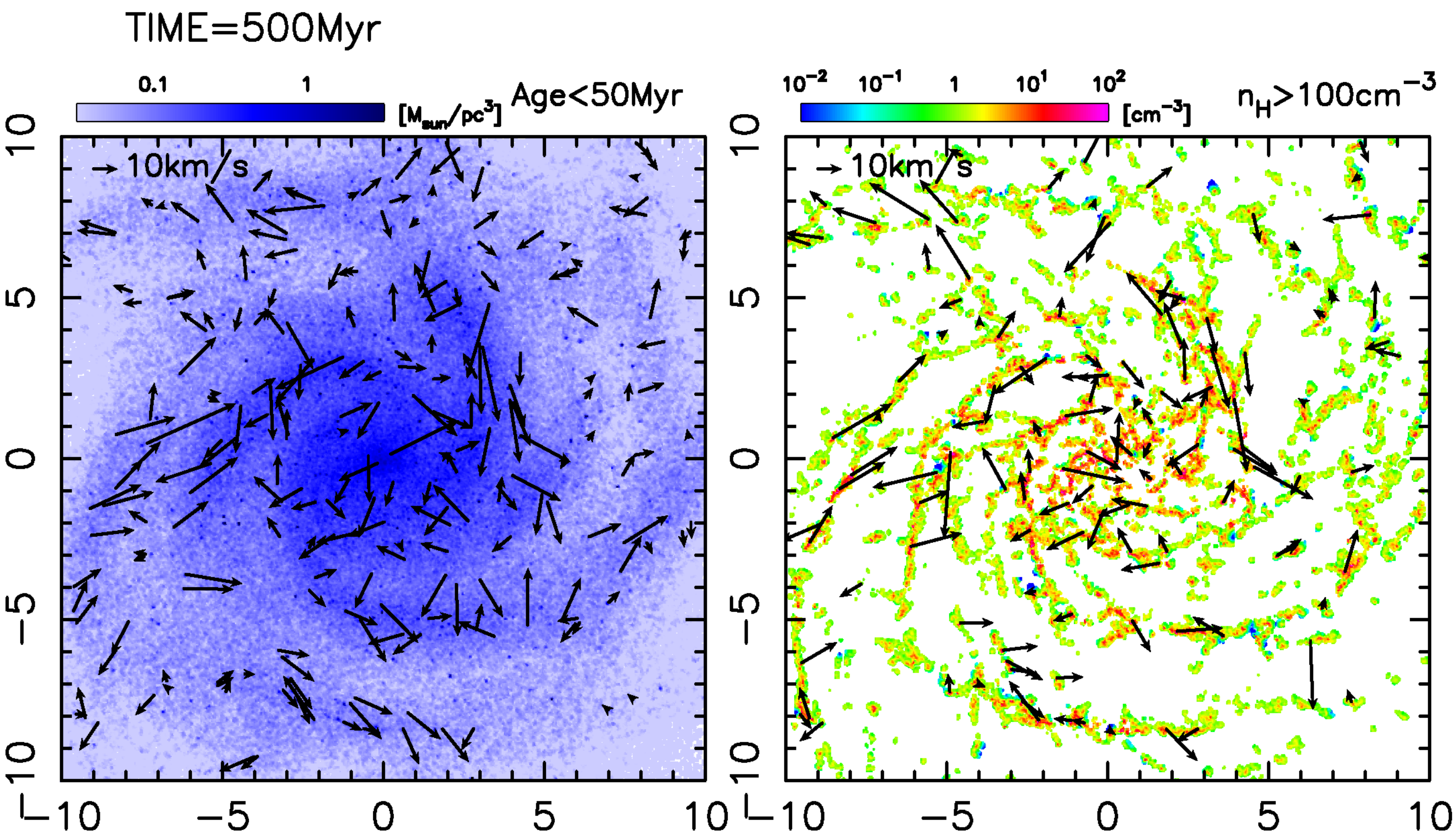
little difference in the result

(In low-resolution calculation, the galaxy would have exploded.)

Galactic disk

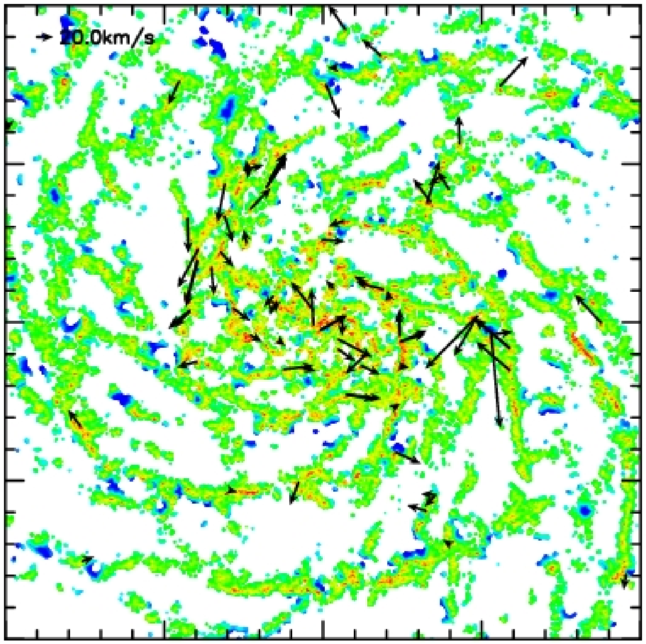
animation 1 2)

Spiral structure and deviation from the circular motion



Left: distribution of stars Right: cold gas

High-resolution model and observation



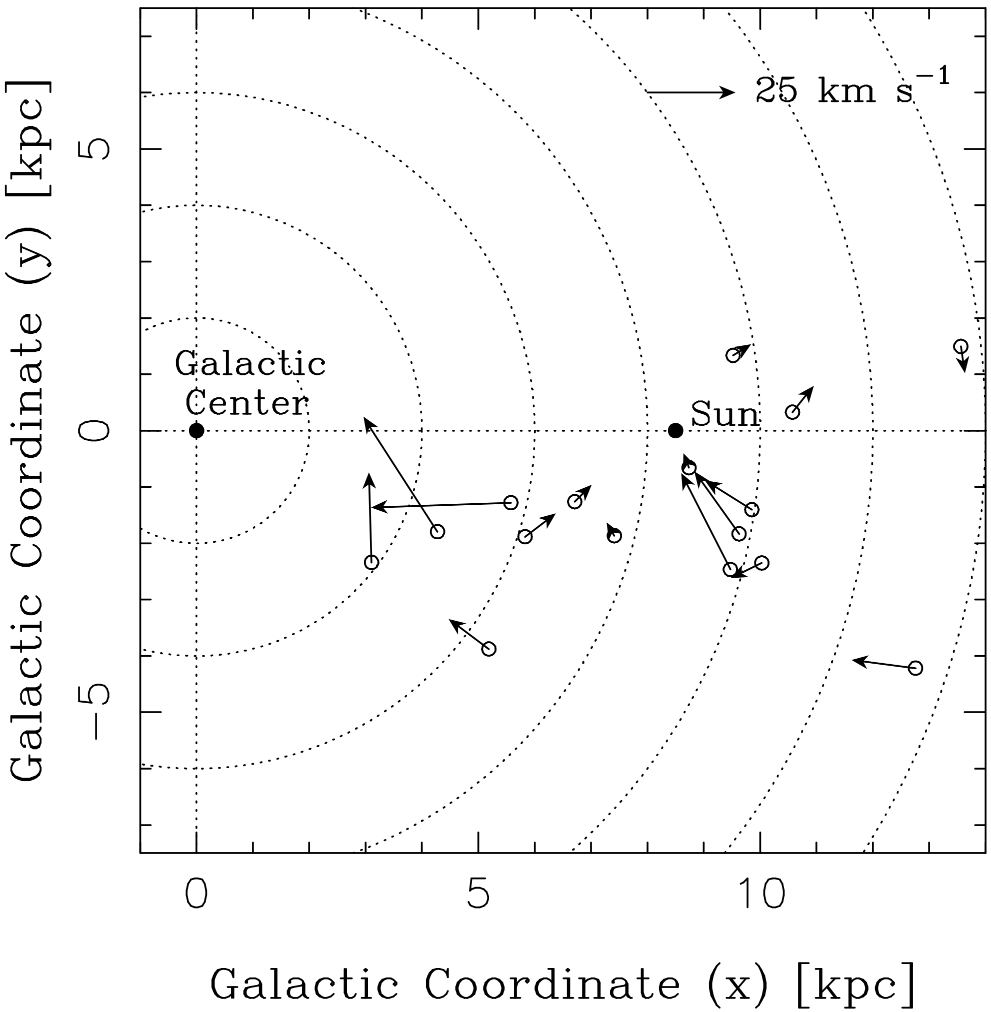
Results from high-resolution simulations

- Star-formation is regulated by large-scale dynamics.
- Observed (multi-arm) spirals can be explained by transient, but recurrent arms.
- These results are robust. Independent of assumption on microphysics such as star-formation timescale.

Observation of Milkyway spiral arms (VLBI)

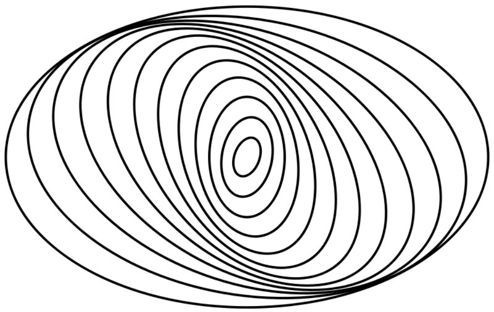
- **Large non-circular motion ($\sim$ 30km/s)**
- **Many data points shows inward motion and counter rotation**
- **Some signs of spacial correlation?**

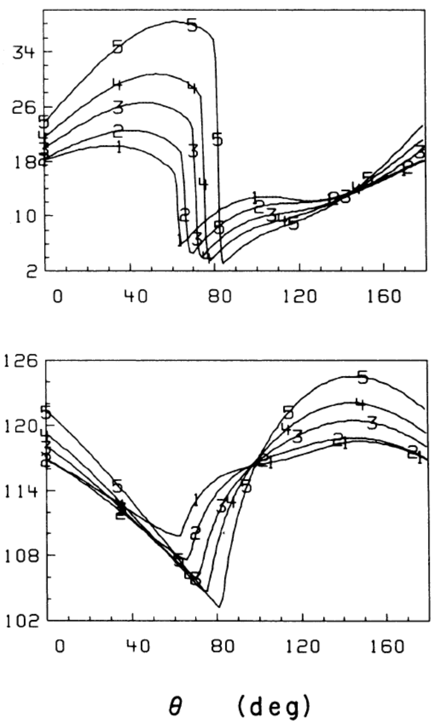
How these motions are induced?



What you learn from textbooks

Stationary density wave



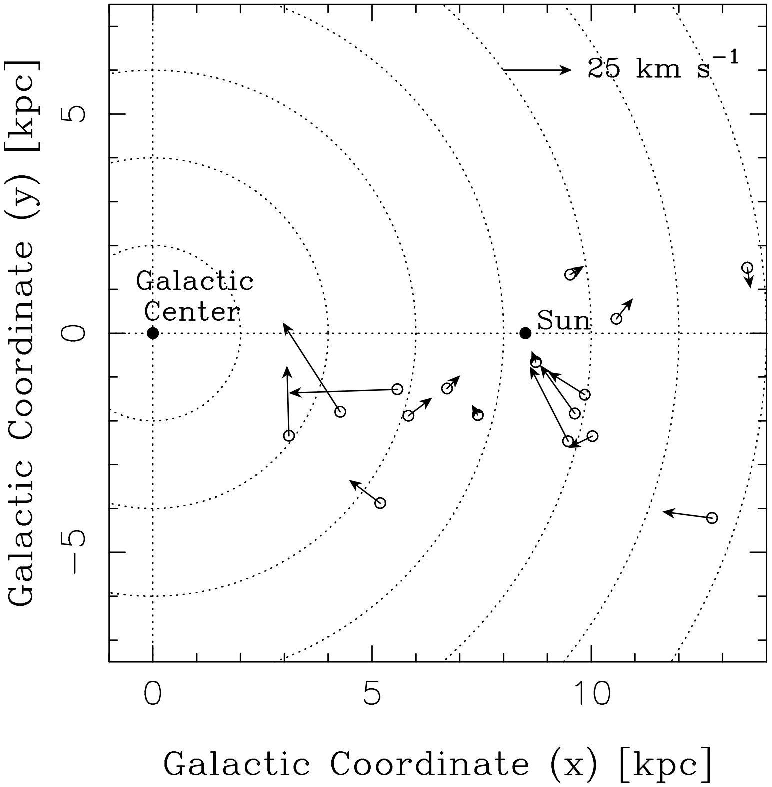


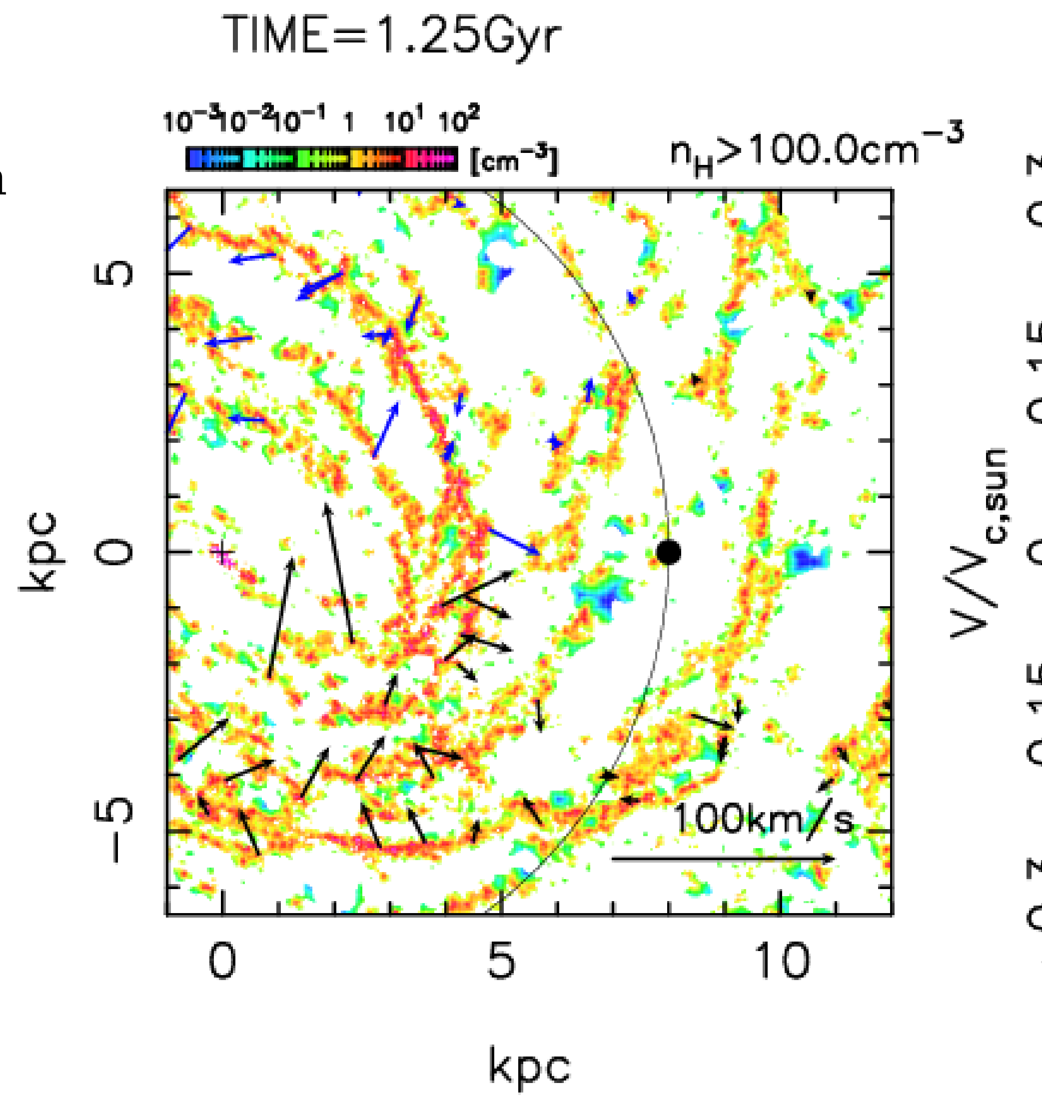
- Spiral arms are not material arms, but density waves
- gas is compressed when it passes through the bottom of the potential well, and form stars there
- It is very difficult to generate non-circular velocity > 10km/s

Quite different from both observation and simulation

Comparison

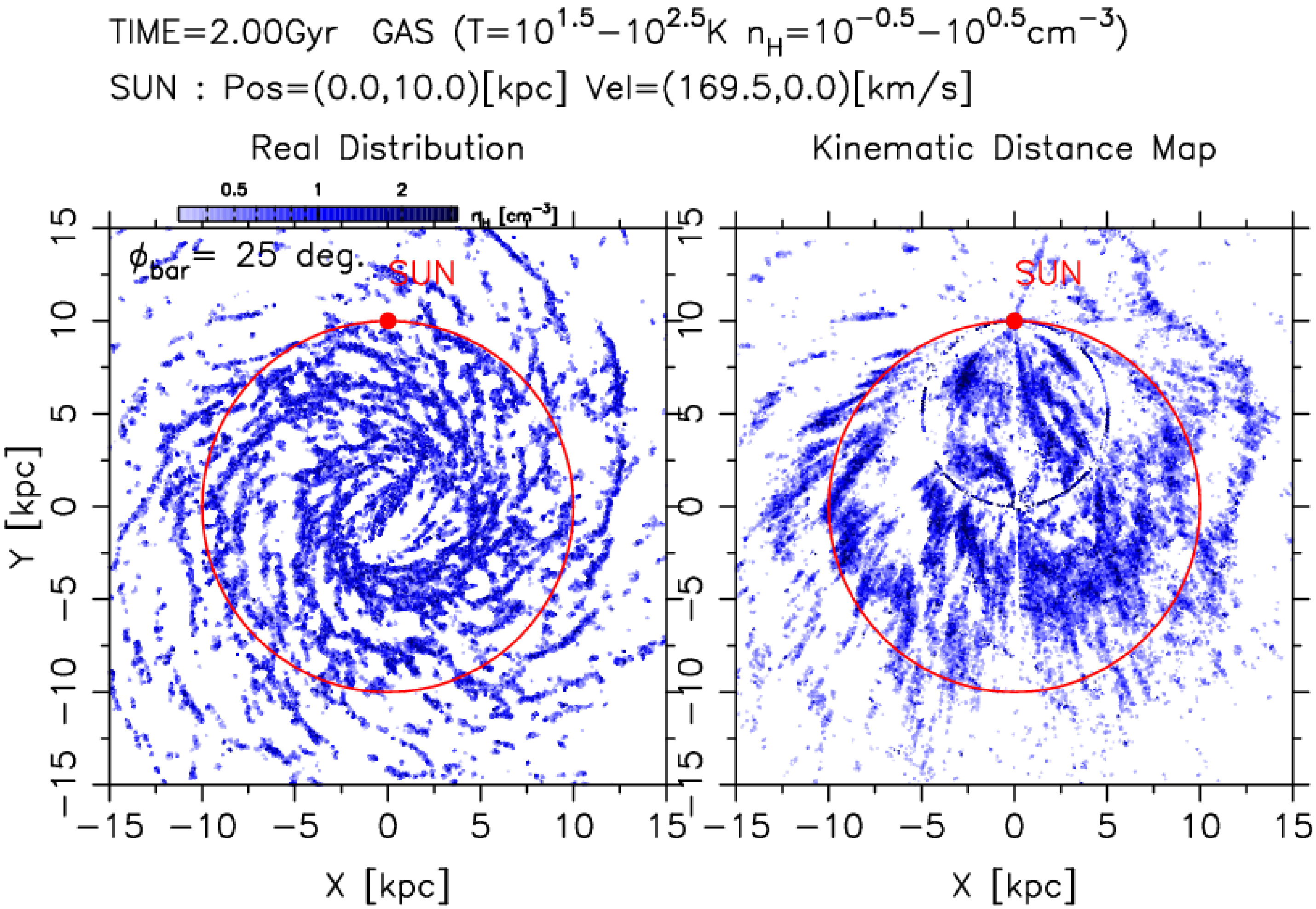
between observation and simulation





Look sort of similar?

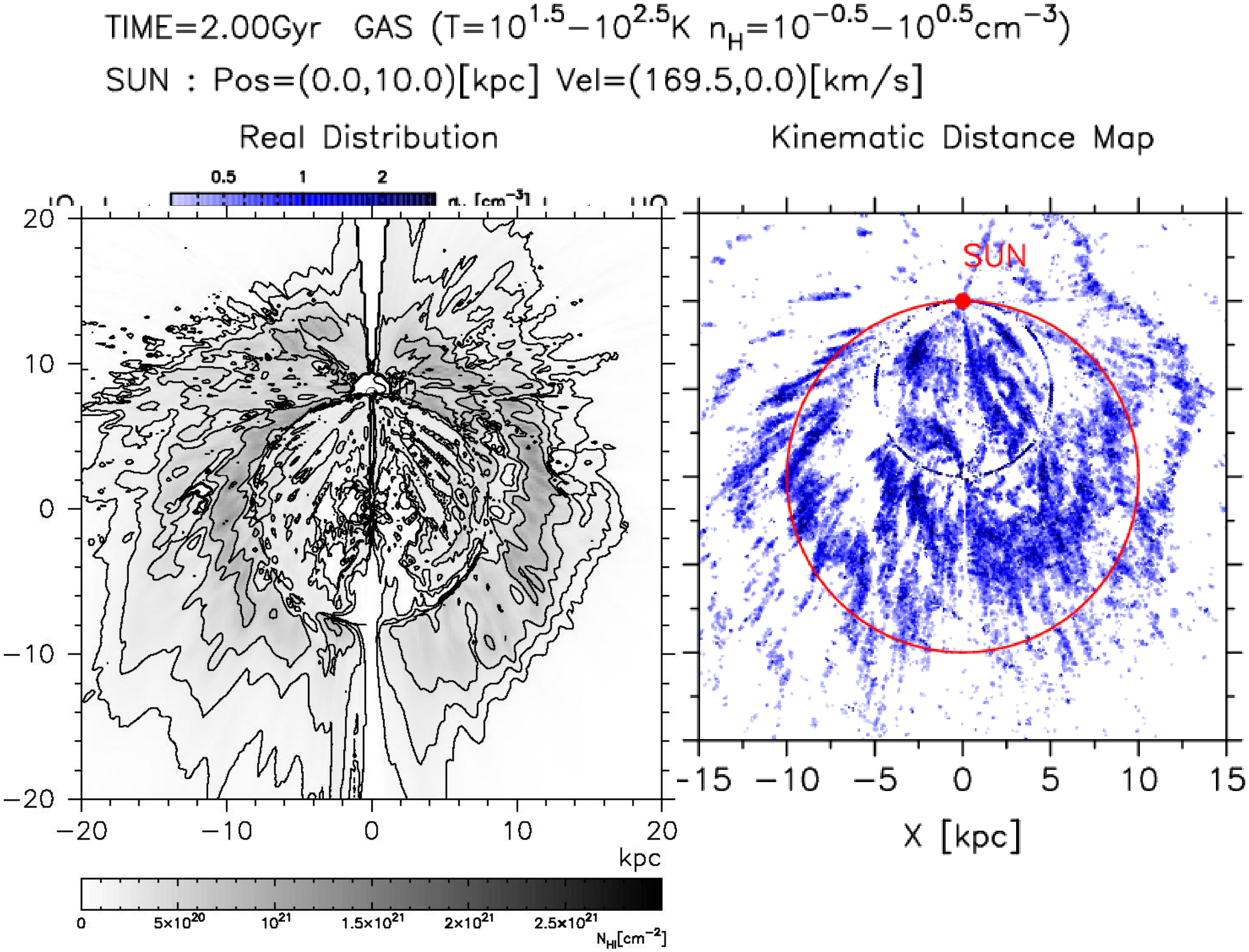
Kinematic distance



Left: Actual distribution Right: Kinematic distance

Quite different...

Kinematic distance



Left : HI observation (Nakanishi and Sofue 2003)

Lots of similar structures

Summary on SPH simulation of spiral arms

- In high-resolution SPH simulations, spiral arms naturally form
- Spiral arms are not stationary, but transient and recurrent
- “VLBI” and “HI” observations of simulation results look very similar to those of Milky way.

Summary

- **Dark matter simulations have advanced very much in the last two decades**
- **However, we have not yet understand the nature or formation mechanism of the central cusp of DM halos**
- **Galaxy formation simulation with Nbody+SPH method have not that advanced**
- **We have devised several new algorithms which might resolve bottlenecks.**