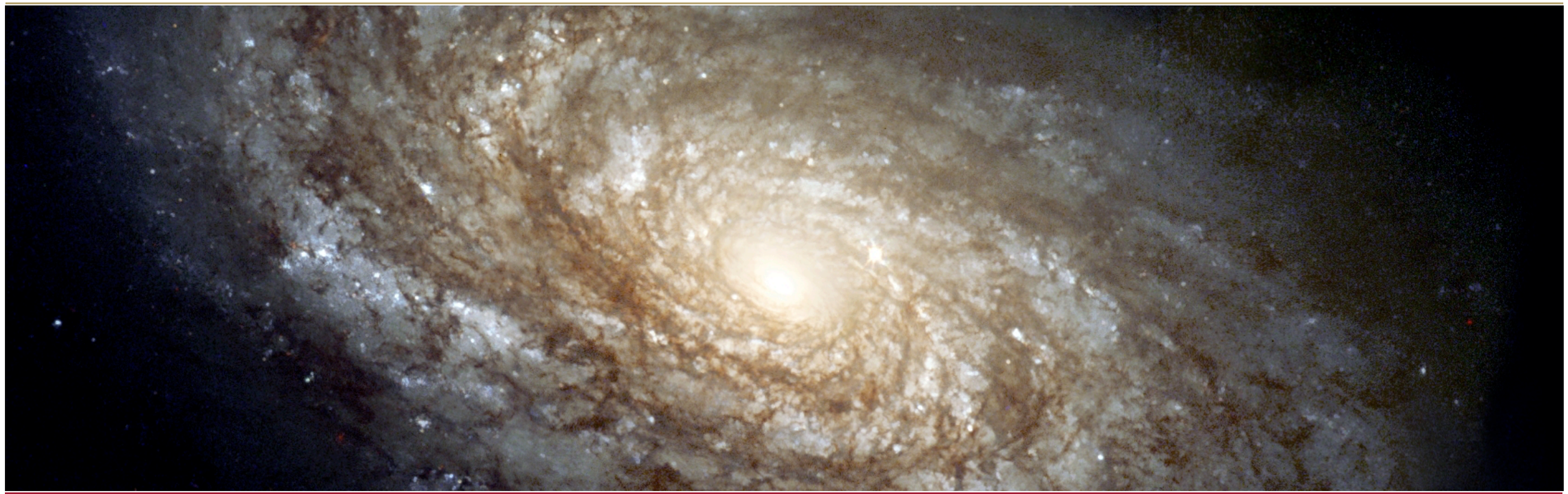
EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

FACULTY OF
SCIENCE
Institute of Astronomy and Astrophysics



N-Body modelling

Computational Physics Labwork Wintersemester 2022/23

24 Oktober 2022 · Christoph Schäfer

Topics

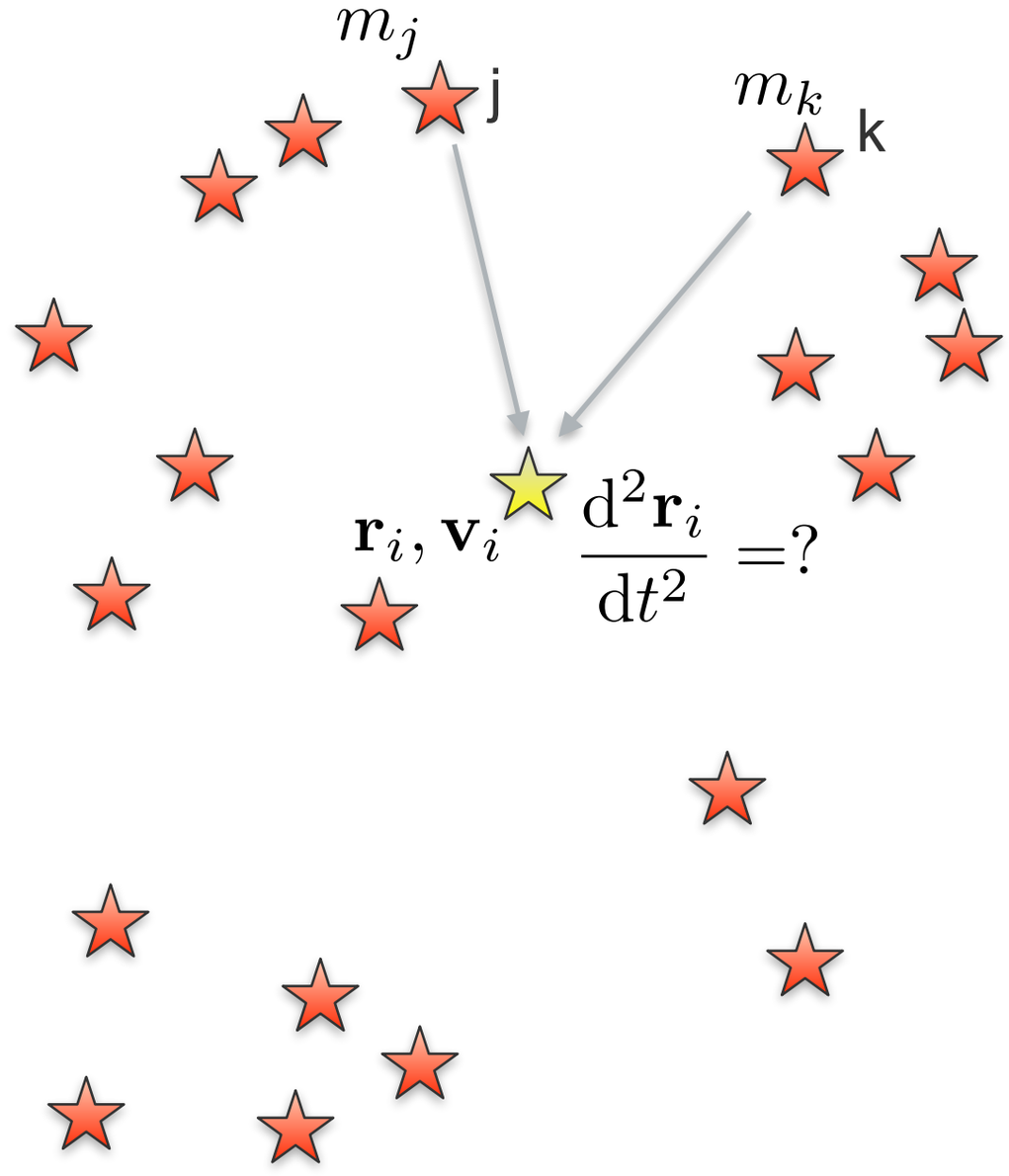
- Theory part
  - The classical astrophysical N-body problem
  - Exact N-body schemes and time integrators
  - Tree algorithms - Barnes & Hut Tree
- Hands-on exercises part
  - Implement some integrators
  - Test and validate with 2-body problem
  - Play around with 100 and 1000 bodies

Goals:

- Learn some basics about time integrators
- Implement your own ~~toy~~ nbody code

The classical astrophysical N-body problem

A number of N particles interact classically through Newton's Laws of Motion and Newton's Law of Gravitation.

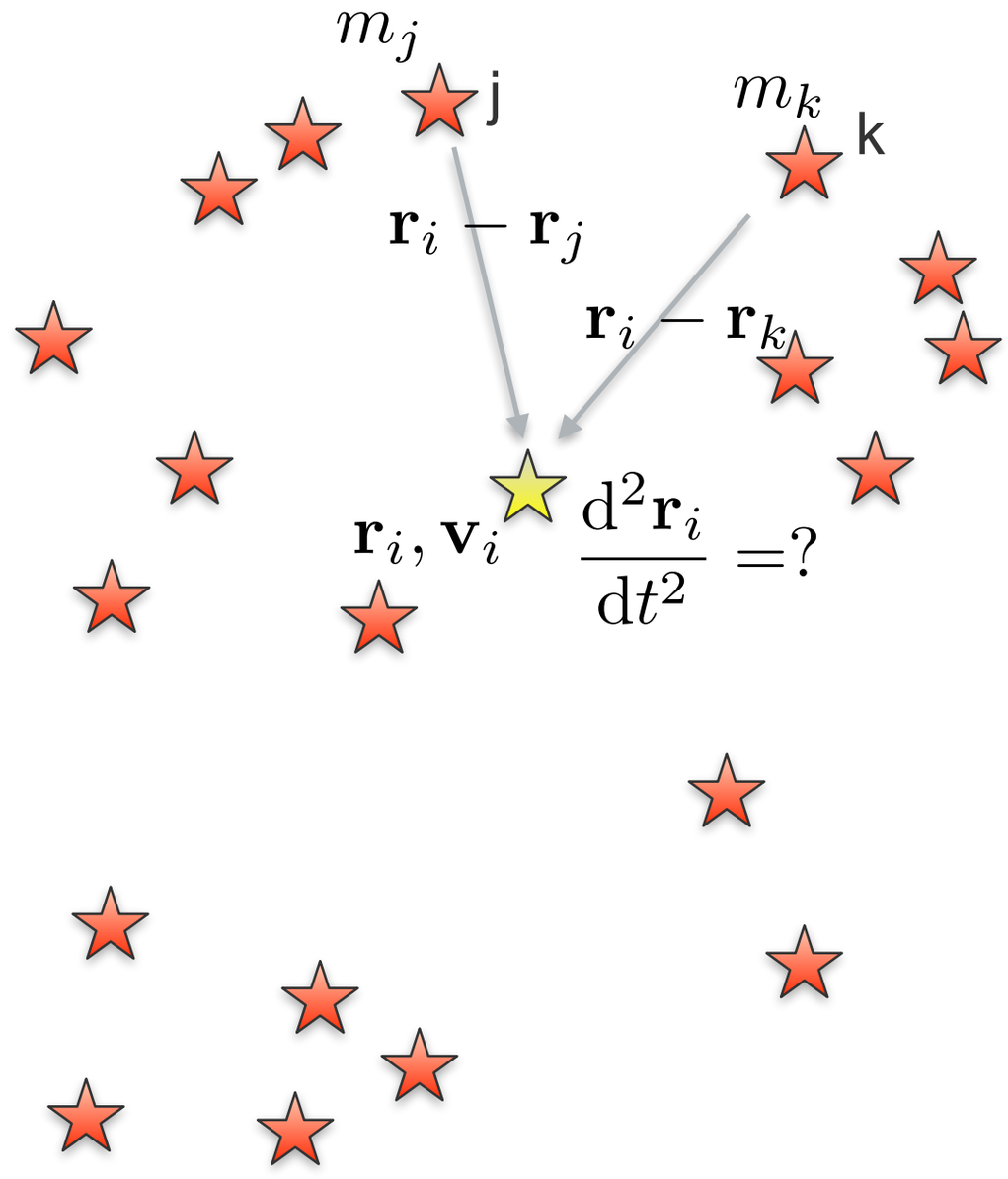


For N=1 and N=2, the equation of motion can be solved analytically.

The classical astrophysical N-body problem

A number of N particles interact classically through Newton's Laws of Motion and Newton's Law of Gravitation.

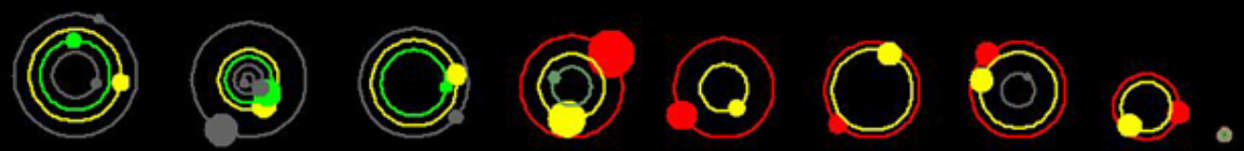
$$\frac{\mathrm{d}^2\mathbf{r}_i}{\mathrm{d}t^2} = -G\sum_{j\neq i}^{N} m_j \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3}$$

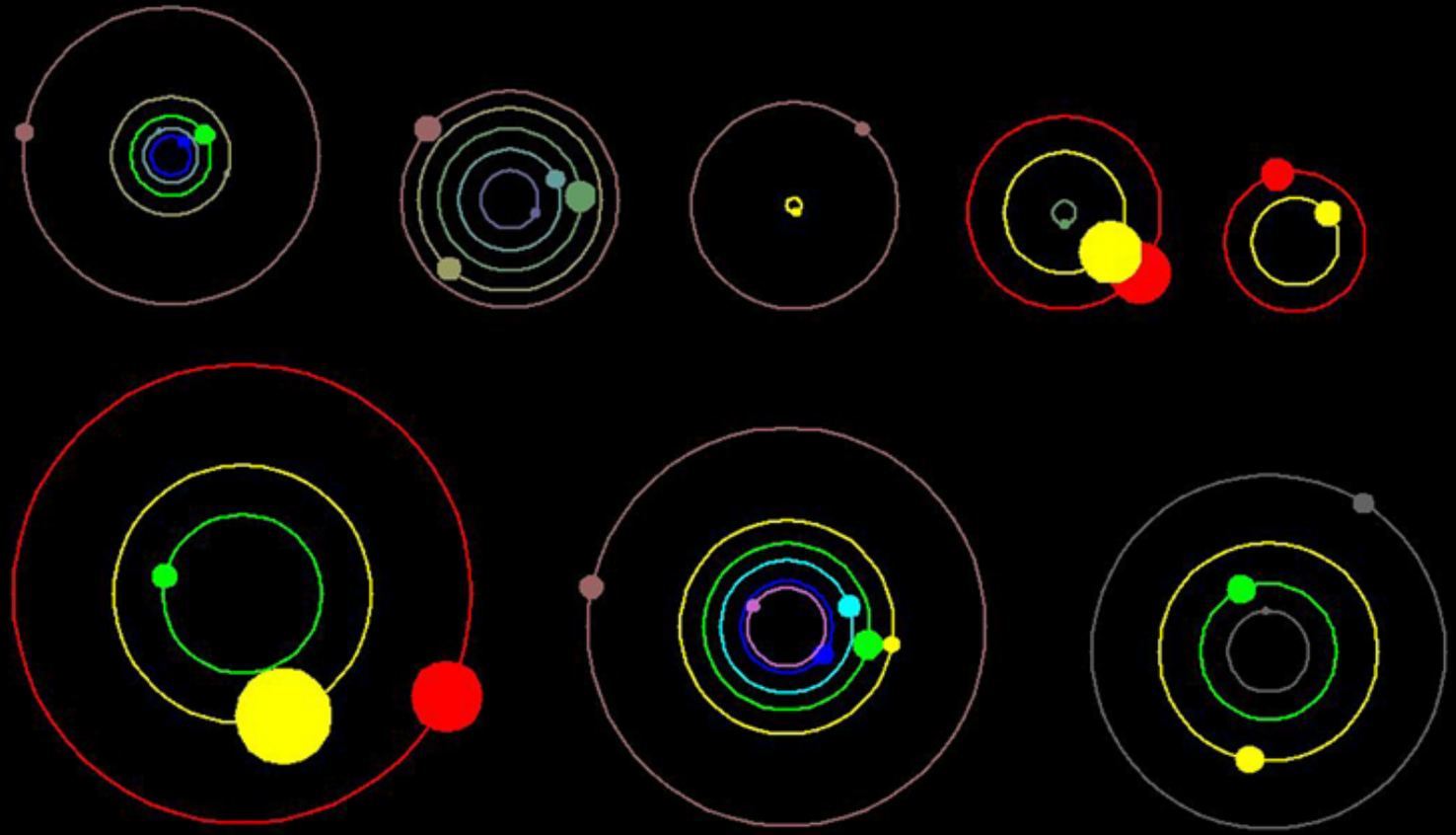


For N=1 and N=2, the equation of motion can be solved analytically.

Applications

- Few-body systems N≅ 3-10
  - planetary systems

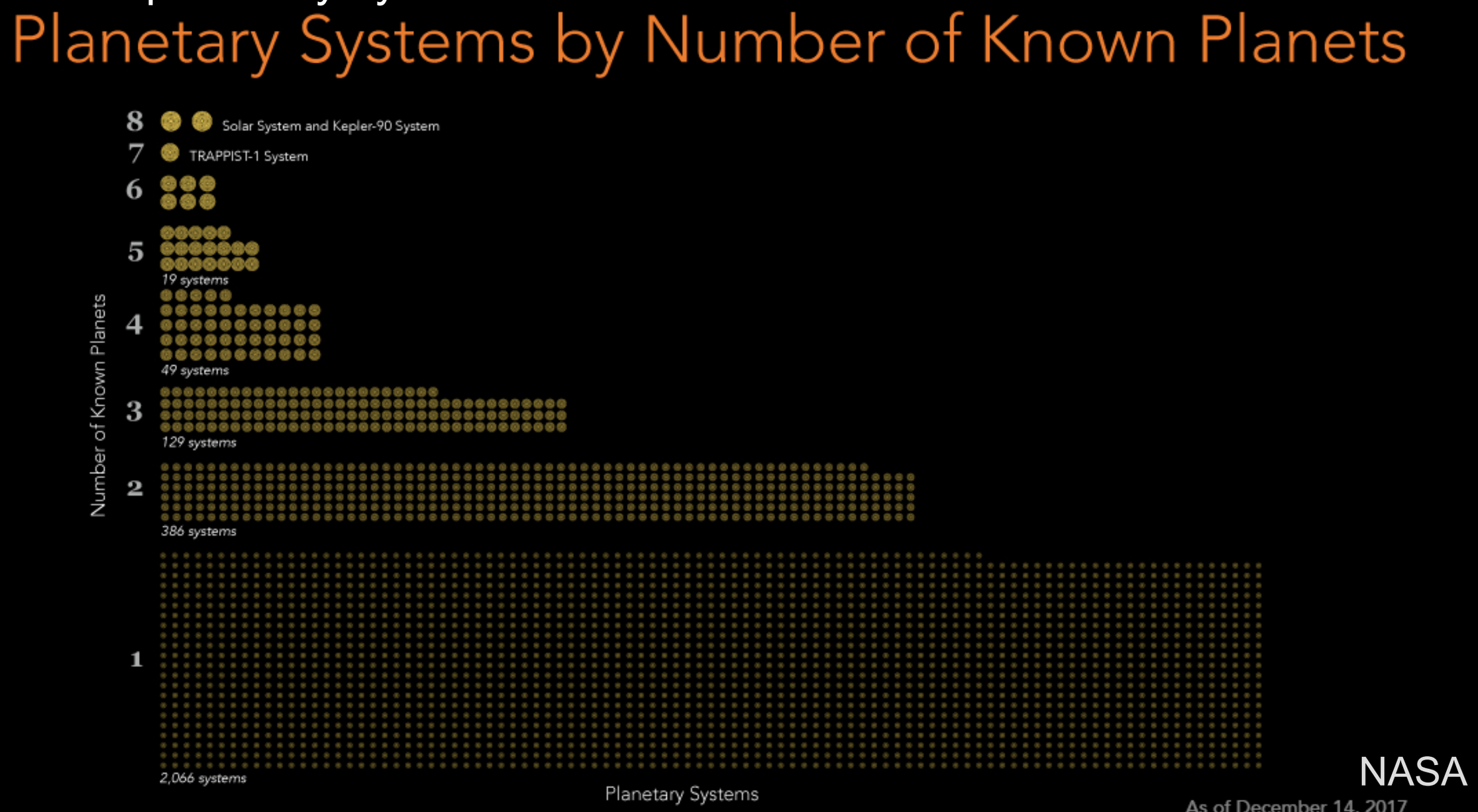




NASA/Ames/UC Santa Cruz 2012

Applications

- Few-body systems N≅ 3-10
  - planetary systems

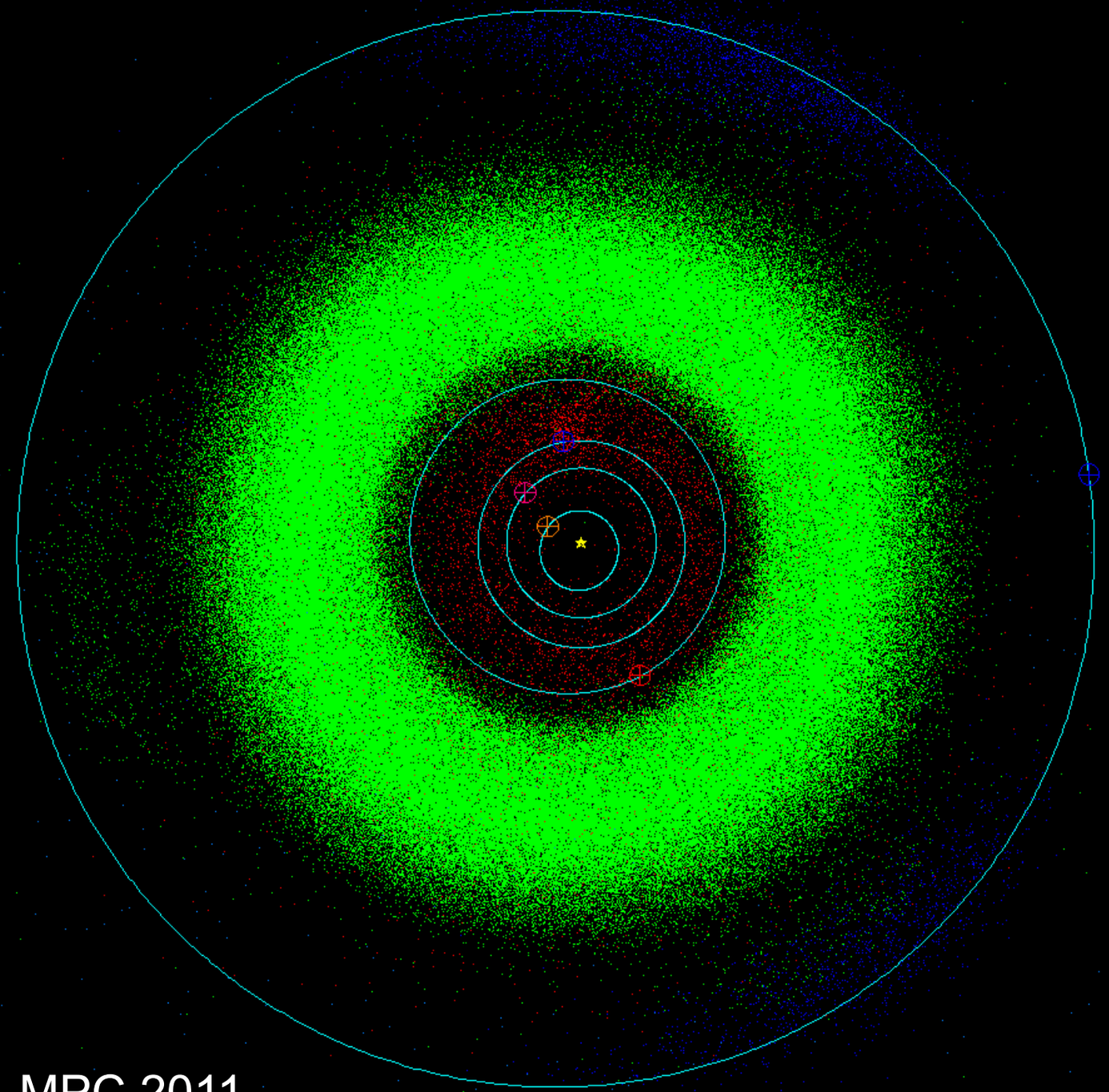


Applications

- Many-body systems
  N ≅ 10-400 and tracers

  planetary systems with minor bodies

  formation of asteroid families



MPC 2011

2011/01/01

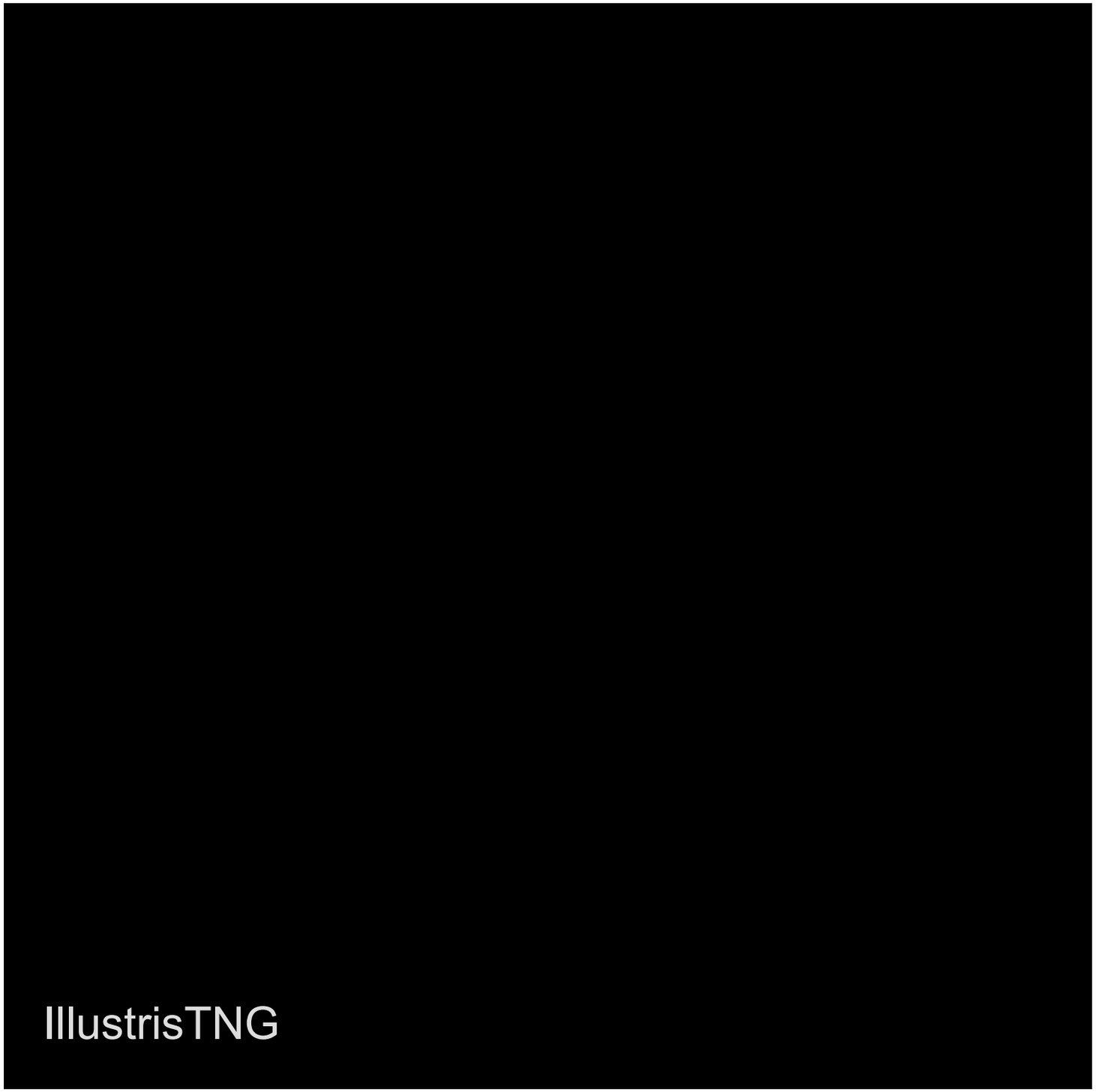
Applications

- Large N-body systems
  - globular star clusters
    $N > 10^3$
  - large-scale structure
    of the universe
  - galactic dynamics and
    cosmology $N > 10^6$

Applications

- Large N-body systems
  - globular star clusters
    $N > 10^3$
  - large-scale structure
    of the universe
  - galactic dynamics and
    cosmology $N > 10^6$

  - coupled hydro-
    nbody simulations,
    Tree-Particle-Mesh
    (Tree-PM)



Applications · Gravity is everywhere

- Few-body systems N≅ 3-10
  - planetary systems

celestial mechanics

- Many-body systems N≅10-400 and tracers
  - planetary systems with minor bodies - formation of asteroid families

- Large N-body systems
  - globular star clusters N > 10^3

dense stellar systems

  - large-scale structure of the universe
  - galactic dynamics and cosmology N > 10^6

galactic dynamics

Applications · today

- Few-body systems N≅ 3-10
  - planetary systems

celestial mechanics

- Many-body systems N≅10-400 and tracers
  - planetary systems with minor bodies - formation of asteroid families

- Large N-body systems
  - globular star clusters N > 10^3

dense stellar systems

  - large-scale structure of the universe
  - galactic dynamics and cosmology N > 10^6

galactic dynamics

Direct N-Body

set of N·dimension 2nd order differential equations

$$\frac{\mathrm{d}^2\mathbf{r}_i}{\mathrm{d}t^2} = -G\sum_{j\neq i}^{N} m_j \frac{\mathbf{r}_i - \mathbf{r}_j}{\left|\mathbf{r}_i - \mathbf{r}_j\right|^3}$$

⟶ 2N·dimension sets of 1st order differential equations

$$\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} = \mathbf{v}_i \qquad \frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} = \mathbf{a}_i = -G\sum_{j\neq i}^{N} m_j \frac{\mathbf{r}_i - \mathbf{r}_j}{\left|\mathbf{r}_i - \mathbf{r}_j\right|^3}$$

initial value problem: 2N·dimension additional conditions needed

$$\mathbf{r}_i(t_0), \mathbf{v}_i(t_0)$$

initial positions and velocities have to be known 🤔

Direct N-Body cont'd

THE GLOBAL SOLUTION OF THE *N*-BODY PROBLEM*

WANG QIU-DONG
Department of Mathematical Science, University of Cincinnati, Cincinnati, OH 45221-0025, U.S.A

(Received: 8 November, 1990; accepted: 26 February, 1991)

(2) Some comments: (i) Although the conclusion given here provides a way to integrate the n-body problem, one does not obtain a useful solution in series expansion. The reason for this is because the speed of convergence of the resulting solution is terribly slow. One has to sum, for example, an incredible number of terms, even for an approximate solution of first order in q, p, t. Because we know almost nothing about the complex singular point in the 7-plane, it seems hopeless to try to improve the convergence of such a series expansion.

see the paper by Florin Diacu about the mathematical history

Direct N-Body cont'd

Numerical integration of the equations of motion

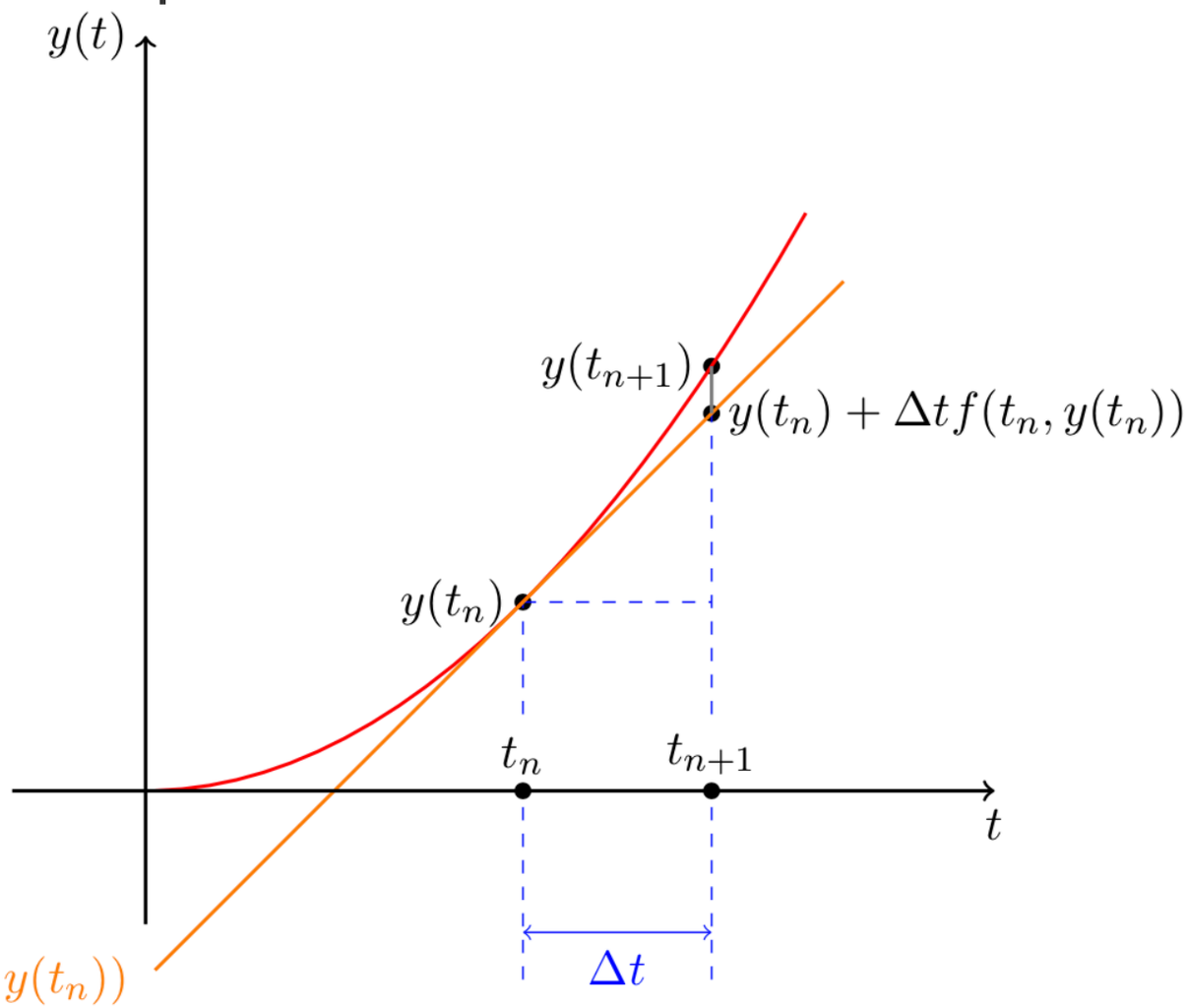
challenge: high accuracy vs. low computational cost

e.g., simple Euler integrator

$$\mathbf{v}_i(t_{n+1}) = \mathbf{v}_i(t_n) + \mathbf{a}_i(t_n)\Delta t$$

$$\mathbf{r}_i(t_{n+1}) = \mathbf{r}_i(t_n) + \mathbf{v}_i(t_n)\Delta t$$

$$\mathbf{a}_i = -G\sum_{j\neq i}^{N} m_j \frac{\mathbf{r}_i - \mathbf{r}_j}{|\mathbf{r}_i - \mathbf{r}_j|^3}$$

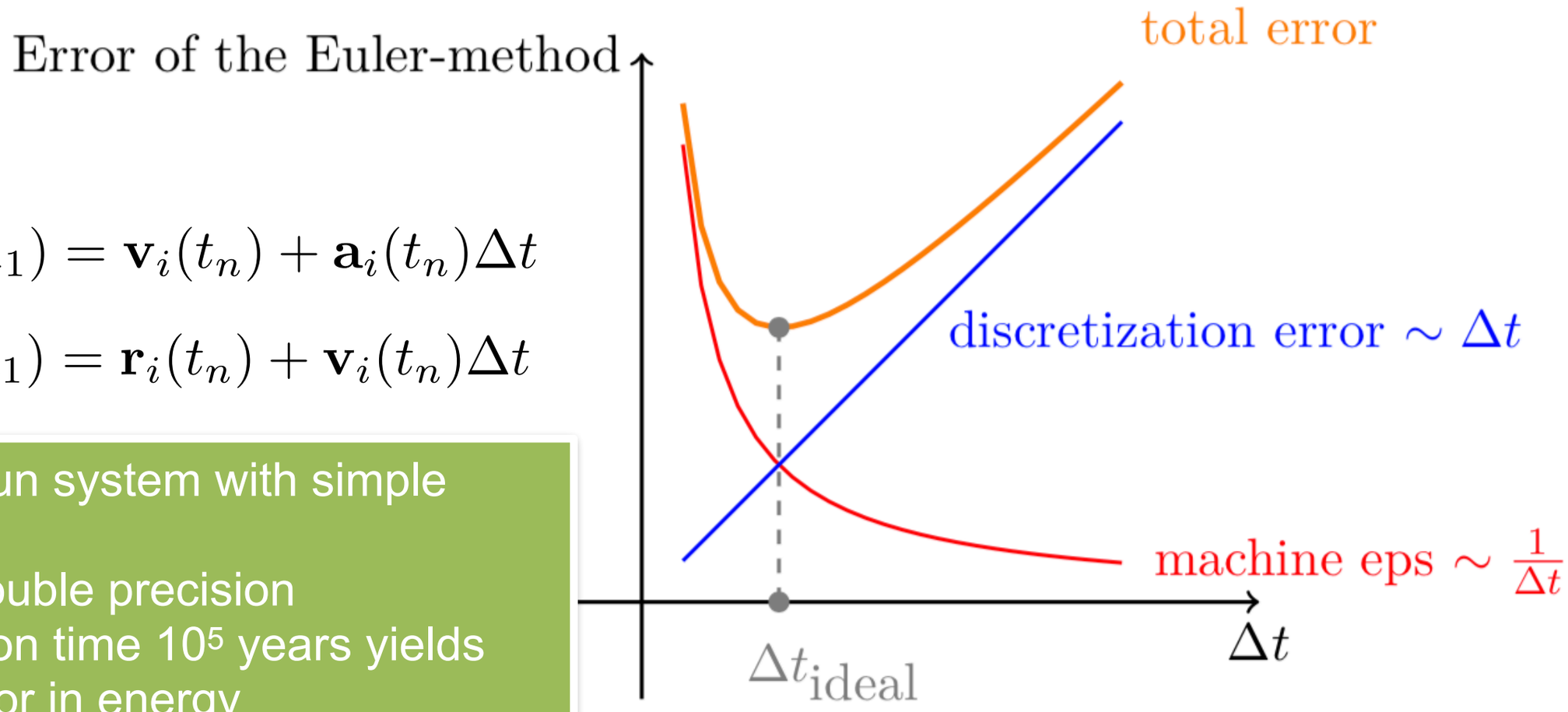


Direct N-Body cont'd

Numerical integration of the equations of motion
Different types of errors
e.g., simple Euler integrator

$$\mathbf{v}_i(t_{n+1}) = \mathbf{v}_i(t_n) + \mathbf{a}_i(t_n)\Delta t$$

$$\mathbf{r}_i(t_{n+1}) = \mathbf{r}_i(t_n) + \mathbf{v}_i(t_n)\Delta t$$



Earth-Sun system with simple Euler:
using double precision
simulation time 10^5 years yields
10% error in energy

Direct N-Body cont'd

Numerical integration of the equations of motion

A vast number of integrators have been developed

one step methods: Euler, Runge-Kutta,…

$$y_{n+1} = y_n + \Delta t \Phi(y, t, \Delta t)$$

e.g., Euler

$$\mathbf{v}_i(t_{n+1}) = \mathbf{v}_i(t_n) + \mathbf{a}_i(t_n)\Delta t$$

$$\mathbf{r}_i(t_{n+1}) = \mathbf{r}_i(t_n) + \mathbf{v}_i(t_n)\Delta t$$

multi step methods: Verlet, Adams–Bashforth, Nyström,…

$$y_{n+1} = y_n + \Delta t \sum_{i=0}^{q-1} \beta_i f(t_{n-i}, y_{n-i})$$

e.g., Verlet

$$\mathbf{r}_{n+1} = 2\mathbf{r}_n - \mathbf{r}_{n-1} + \mathbf{a}_n \Delta t^2$$

symplectic integrators: Leap-Frog, Wisdom-Holman, symplectic RKs,…

Hamiltonian systems

energy conservation

e.g., Leap-Frog for N-Body

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \mathbf{a}_{n+1/2}\Delta t$$

$$\mathbf{r}_{n+3/2} = \mathbf{r}_{n+1/2} + \mathbf{v}_{n+1}\Delta t$$

Time Integrators

- Consider the following problem

$$y'(x) = \frac{\mathrm{d}y}{\mathrm{d}x} = f(x, (y(x)), \quad a \leq x \leq b,$$
$$y(a) = y_0 .$$

- Finite interval [a,b] and function f(x,y) are given.

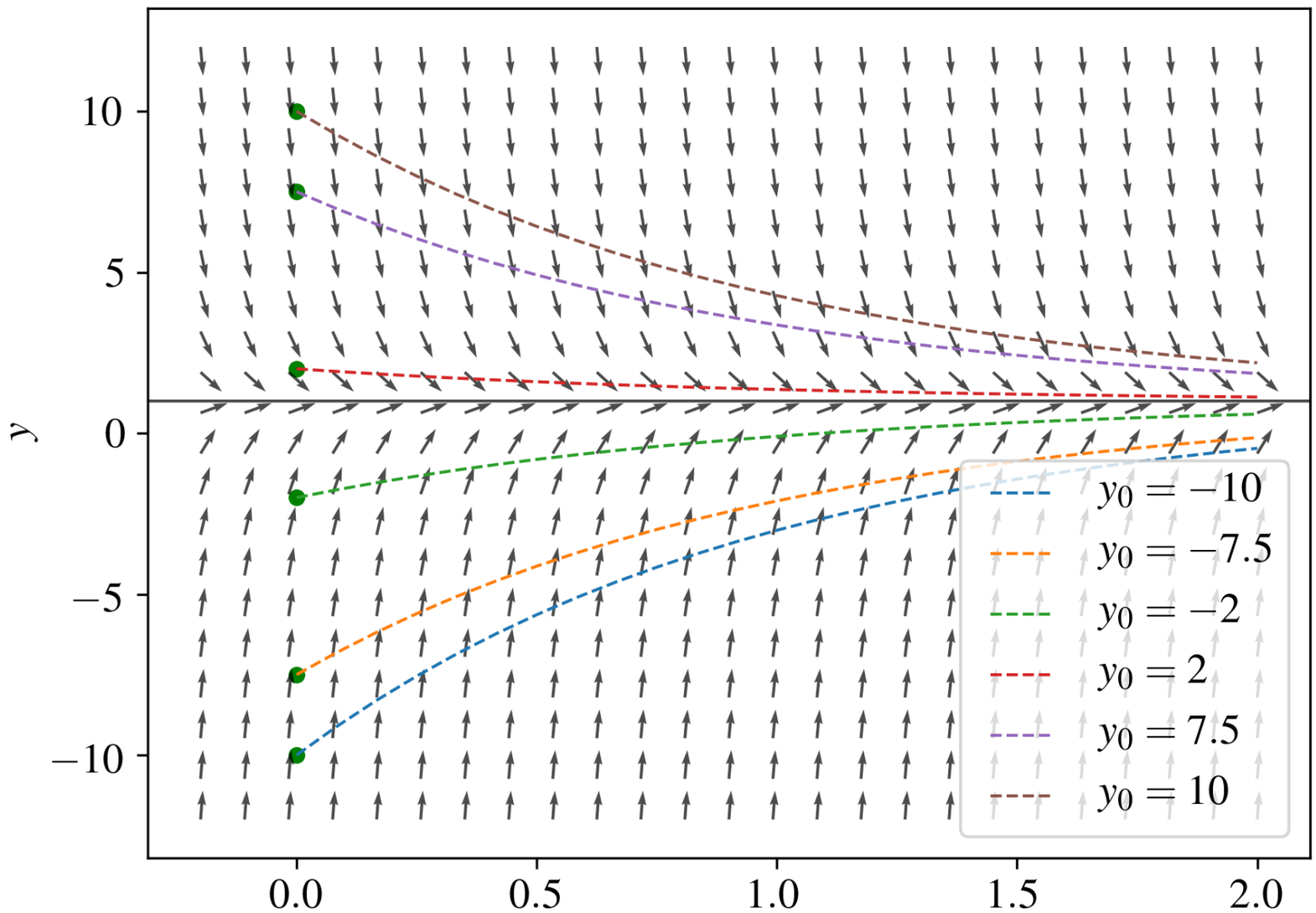
- We seek a function y(x) defined on [a,b], such that $y'(x) = f(x,y(x))$ for $a \leq x \leq b$, and such that y(x) satisfies the initial condition $y(a) = y_0$, where y_0 is some prescribed value.

- This is called an **initial value problem**, cf. N-Body problem.

Time Integrators

- The IVP consists of two parts: the differential equation which gives the relationship between y(x) and y′(x) and the initial condition $y(a) = y_0$.

- Slope field, vector plot at points (x, y) showing the slope [1, f (x, y(x))]. It gives an impression on the type of possible solutions. The choice of y_0 selects the specific, corresponding solution.



Time Integrators

- Basic idea: write dx and dy as finite intervals Δx and Δy. Discretise interval [a,b] with n points, and walk from point to point with a step size Δx=h.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\Delta y}{\Delta x} = f(x,y) \quad \longrightarrow \quad \Delta y \equiv y_{k+1} - y_k = \Delta x f(x,y)\,.$$

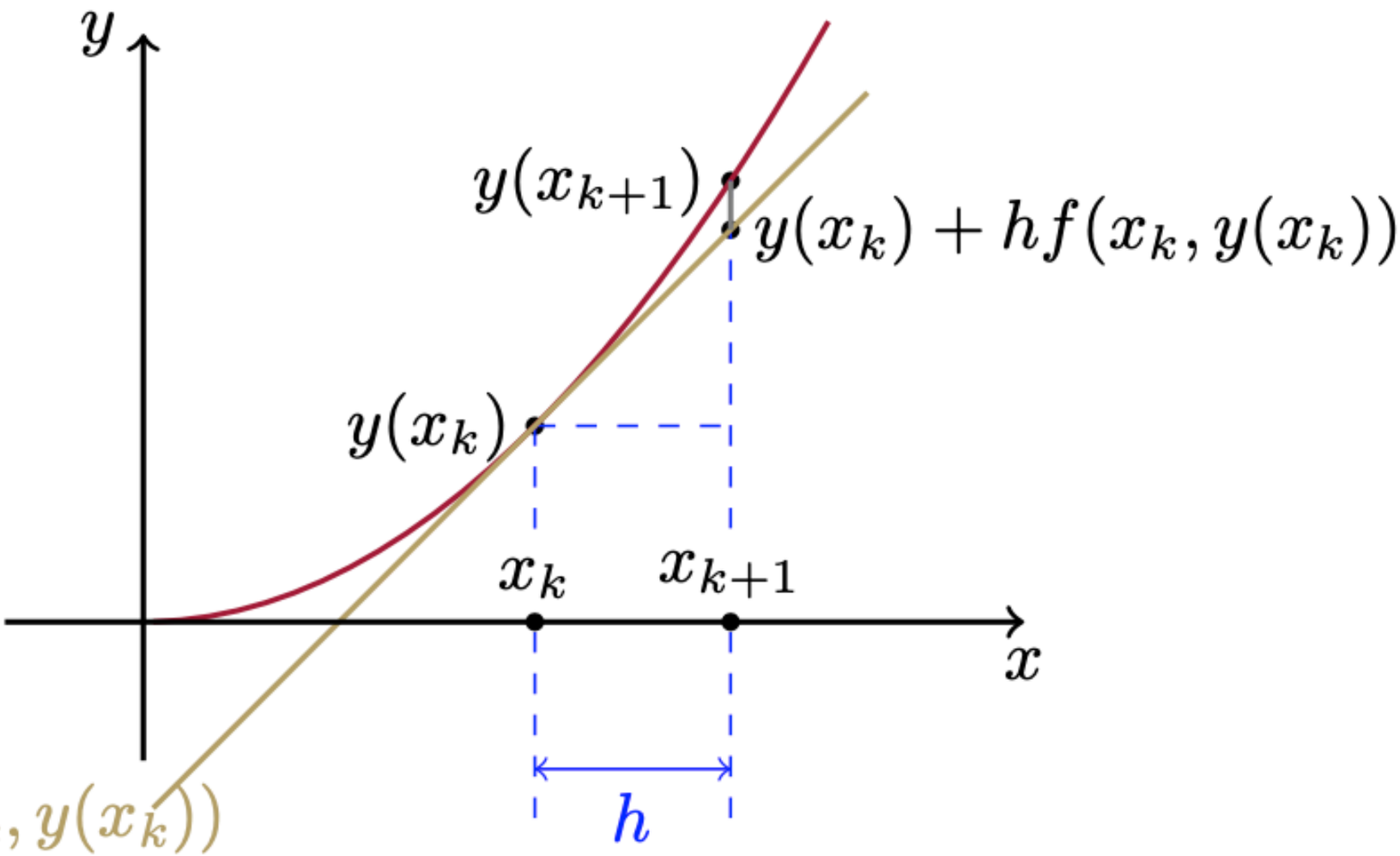
- In general, **one step methods** can be written in the following form

$$y_{k+1} = y_k + h\Phi(x_k, y_k; y_{k+1}, h)\,.$$

  - Start from x_k to x_{k+1} and y_k to y_{k+1}, only use previous values.
  - If the right hand side does not depend on y_{k+1}, i.e. $\Phi = \Phi(x_k, y_k; h)$, **explicit**, otherwise **implicit** method.
  - The function Φ is called **increment function**.

Time Integrators - Euler Integrator

- Increment function is f(x,y): $y_{k+1} = y_k + hy'_k = y_k + hf(x_k, y_k)$.



Time Integrators - Runge-Kutta Integrators

- The Runge-Kutta methods are based on integration by Lagrange polynomials.
- We rewrite the ODE as an integral equation

$$y' = f(x, y(x)) \quad \Leftrightarrow \quad \int_{x_k}^{x_{k+1}} \frac{\mathrm{d}y}{\mathrm{d}x} \mathrm{d}x = \int_{x_k}^{x_{k+1}} f(x, y(x)) \mathrm{d}x,$$

  and obtain

$$y(x_{k+1}) = y(x_k) + \int_{x_k}^{x_{k+1}} f(x, y(x)) \mathrm{d}x$$

- Now, the integral on the right hand side is solved by the help of polynomial expansion. One possibility is to apply the trapezoidal rule with $h = x_{k+1} - x_k$ and use

$$\int_{x_k}^{x_{k+1}} y' \mathrm{d}x \approx \frac{h}{2} \left[y'(x_k) + y'(x_{k+1}) \right].$$

Time Integrators - Runge-Kutta Integrators

- Since $y'(x_{k+1})$ is not known yet, we have to approximate it, e.g., with an Euler integrator step

$$y'(x_k) = K_1 = f(x_k, y_k),$$
$$y'(x_{k+1}) = K_2 = f(x_{k+1}, y_{k+1}) = f(x_k + h, y_k + hK_1)\,.$$

- This gives us a 2nd order Runge-Kutta integrator, knows as **Heun's method**

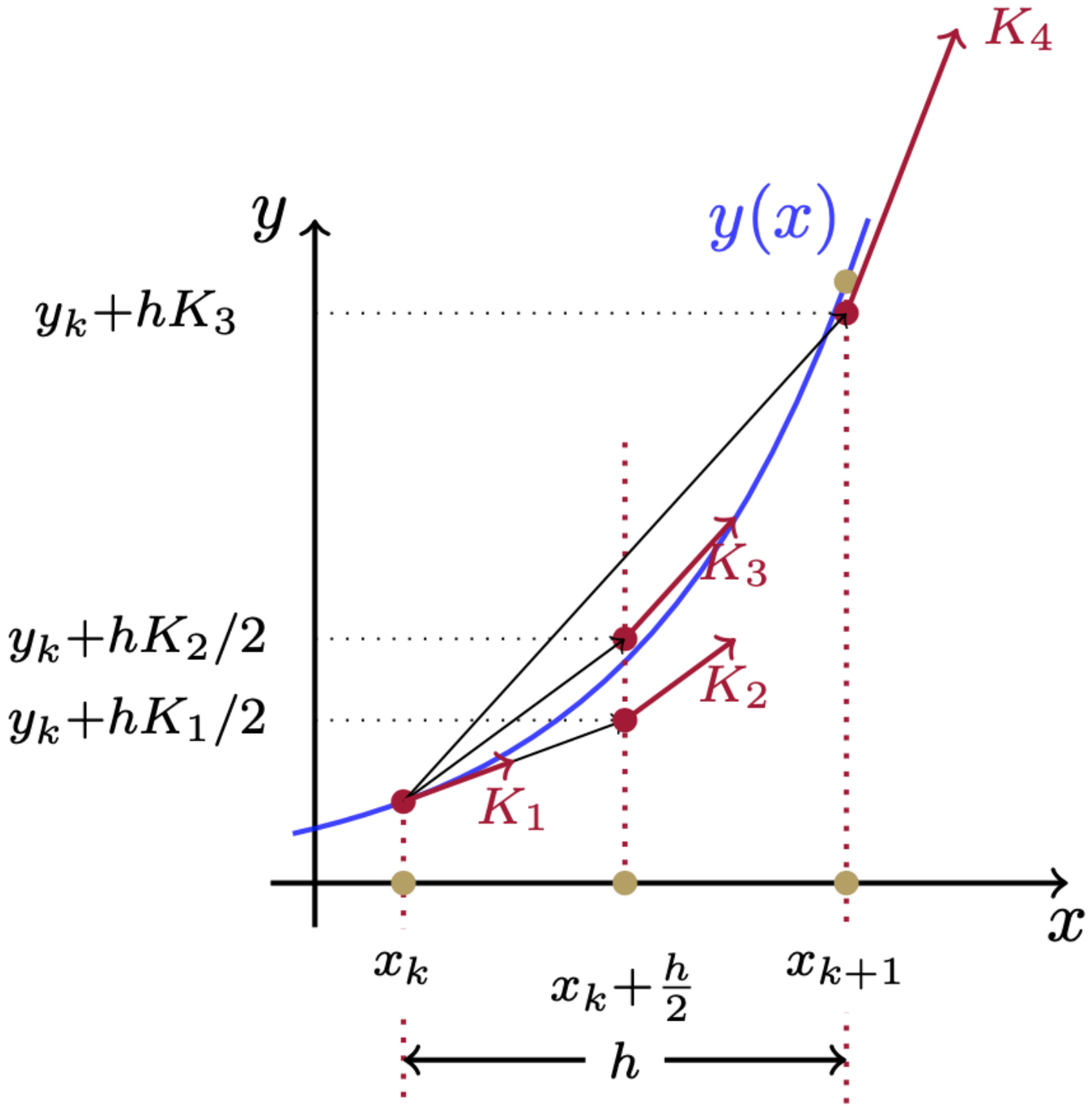
$$K_1 = f(x_k, y_k),$$
$$K_2 = f(x_k + h, y_k + hK_1),$$
$$y_{k+1} = y_k + \frac{h}{2}(K_1 + K_2)\,.$$

- Heun's method is also a vivid example for the diversity in integrators: It is both a **2nd order RK** and a simple **predictor-corrector method**.

Time Integrators - Runge-Kutta Integrators - RK4

- One step from x_k to x_{k+1}

$$K_1 = f(x_k, y_k)$$
$$K_2 = f(x_k + \frac{h}{2}, y_k + \frac{h}{2}K_1)$$
$$K_3 = f(x_k + \frac{h}{2}, y_k + \frac{h}{2}K_2)$$
$$K_4 = f(x_k + h, y_k + hK_3)$$
$$y_{k+1} = y_k + \frac{h}{6}(K_1 + 2K_2 + 2K_3 + K_4)\,.$$



Time Integrators - Runge-Kutta Integrators - RK4

```
# rk4 integrator, fixed step size, f is function which returns derivative
# integration in interval [x0,x1], y0 is initial value
def rk4(f, x0, y0, x1, h):
    N = (x1-x0)//h
    N = int(N)
    x = np.zeros(N+1)
    y = np.zeros(N+1)
    y[0] = yi = y0
    x[0] = xi = x0
    for i in range(1,N+1):
        k1 = h * f(yi,xi)
        k2 = h * f(yi+0.5*k1, xi+0.5*h)
        k3 = h * f(yi+0.5*k2, xi+0.5*h)
        k4 = h * f(yi+k3,xi+h)
        x[i] = xi = x0 + i*h
        y[i] = yi = yi + (k1+k2+k2+k3+k3+k4)/6
    return x,y
```

Time Integrators - multi-step methods

- So far, we have only used the information from the system during one step from k to k+1 and discarded all information from the steps before.

- The multi-step methods also use the information from these former steps.

- Advantage: higher accuracy possible, more efficient.
  Disadvantages: more memory required, more complicated algorithms.

- A **linear** multistep method uses a linear combination of the y_k and $f(x_k,y_k)$ of the previous s steps, its general form is

$$y_{k+s} + a_{s-1}y_{k+s-1} + a_{s-2}y_{k+s-2} + \ldots + a_0 + y_k = h(b_s f(x_{k+s}, y_{k+s}) + b_{s-1}f(x_{k+s-1}, y_{k+s-1}) + \ldots + b_0 f(x_k, y_k)\,.$$

Time Integrators - multi-step methods

- Or

$$\sum_{j=0}^{s} a_j y_{k+j} = h \sum_{j=0}^{s} b_j f(x_{k+j}, y_{k+j}) \,.$$

with $a_s = 1$. The coefficients $a_0, \ldots, a_{s-1}$ and $b_0, \ldots, b_s$ determine the method.

- For $b_s = 0$, one obtains an explicit scheme, since y_{k+j} can be directly computed.
- For $b_s \neq 0$, the scheme is implicit and has to be solved iteratively.
- Simple example: Two-step Adams-Bashforth method (cf. to Euler)

$$y_{k+2} = y_{k+1} + \frac{3}{2} h f(x_{k+1}, y_{k+1}) + \frac{1}{2} h f(x_k, y_k) \,.$$

Time Integrators - Symplectic Integrators

- Symplectic integrators are designed for the numerical solution of a Hamiltonian system

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}$$

- Main idea: The transformation of the coordinates and momenta from t_0 to a later time t_1 is canonical and hence phase space volume preserving
- In the following, we focus on separable Hamiltonians

$$H(p, q) = T(p) + V(q)$$

  with kinetic energy T and potential energy V.

- Introducing $z = (q, p)$, we can write the Hamiltonian eqs using the Poisson bracket {} ($\{f, g\} = \frac{\partial f}{\partial q}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial q}$):

$$\dot{z} = \{z, H\}$$

Time Integrators - Symplectic Integrators

- With the operator $D_H = \{\,\cdot\,, H\}$, the formal solution of this equation is

$$z(t) = \exp(tD_H)z(0).$$

- Or for a separable Hamiltonian

$$z(t) = \exp(t(D_T + D_V))z(0).$$

- We now approximate the time-evolution operator by a product of operators

$$\exp(t(D_T + D_V)) = \Pi_{i=1}^{k} \exp(c_i t D_T) \exp(d_i t D_V) + \mathcal{O}(t^{k+1})$$

  where $\sum_i c_i = 1 = \sum_i d_i$

- Both operators $\exp(c_i t D_T)$ and $\exp(d_i t D_V)$ provide a symplectic map (phase-space volume conserved mappings), and since $D_T^2 z = \{\{z, T\}, T\} = \{(\dot{q},0), T\} = 0$, we find

$$\exp(c_i D_T) = \sum_{n=0}^{\infty} \frac{(c_i D_T)^n}{n!} = 1 + c_i D_T$$

  and likewise for D_V.

Time Integrators - Symplectic Integrators

- We finally end up with a symplectic mapping that conserves both energy and momentum

$$\begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} q + tc_i \partial_p T(p) \\ p \end{pmatrix} \text{ from } \exp(tc_i D_T)$$

$$\begin{pmatrix} q \\ p \end{pmatrix} \mapsto \begin{pmatrix} q \\ p - td_i \partial_q V(q) \end{pmatrix} \text{ from } \exp(td_i D_V)$$

- or in coordinates for space and velocity

$$x_{i+1} = x_i + c_i v_{i+1} \Delta t$$
$$v_{i+1} = v_i + d_i \frac{F(x_i)}{m} \Delta t.$$

- Simple example, symplectic Euler with c=d=1

$$x_{i+1} = x_i + v_{i+1} \Delta t$$
$$v_{i+1} = v_i + \frac{F(x_i)}{m} \Delta t.$$

- Further reading: Kinoshita, H., Yoshida, H., & Nakai, H., Symplectic integrators and their application to dynamical astronomy, Celest. Mech. Dyn. Astron., 1991

Time Integrators - Symplectic Integrators

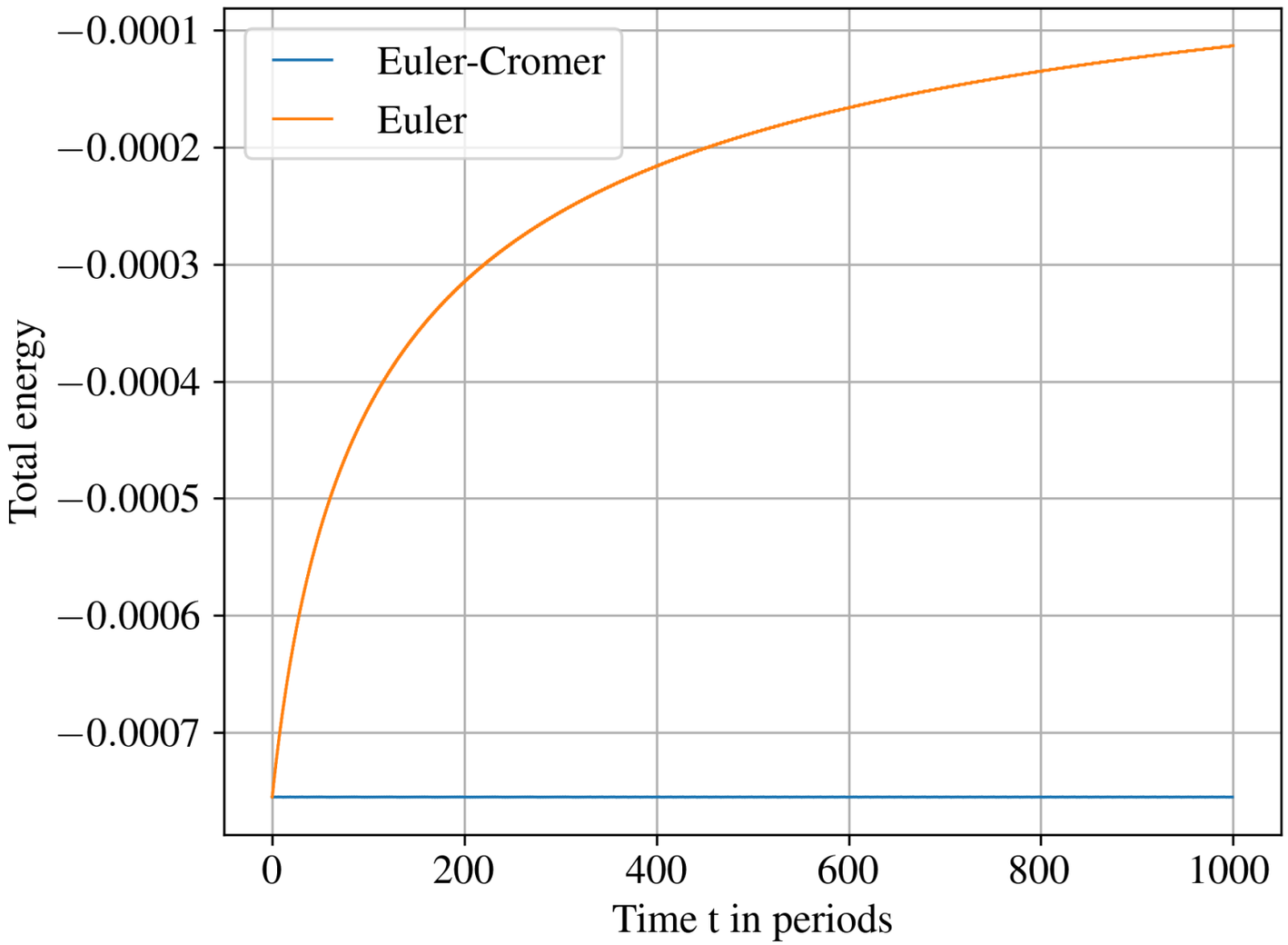
Euler

$$x_{i+1} = x_i + v_i \Delta t$$
$$v_{i+1} = v_i + a_i \Delta t.$$

Euler-Cromer aka symplectic Euler

$$x_{i+1} = x_i + v_{i+1} \Delta t$$
$$v_{i+1} = v_i + a_i \Delta t.$$

Energy,
2-body problem,
integrated for 1000
orbits



Time Integrators - Symplectic Integrators

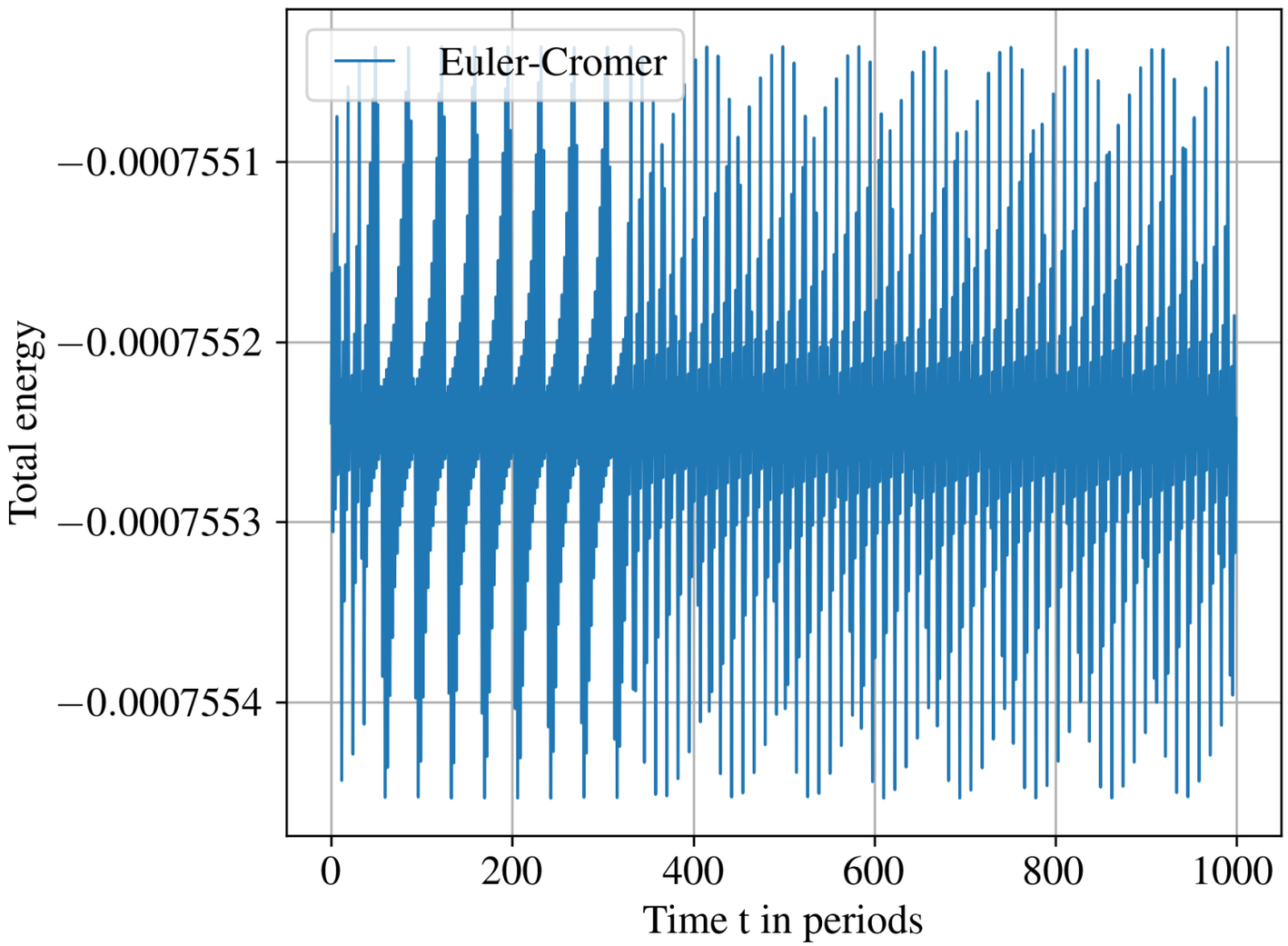
Euler

$$x_{i+1} = x_i + v_i \Delta t$$
$$v_{i+1} = v_i + a_i \Delta t.$$

Euler-Cromer aka symplectic Euler

$$x_{i+1} = x_i + v_{i+1} \Delta t$$
$$v_{i+1} = v_i + a_i \Delta t.$$

Energy,
2-body problem,
integrated for 1000
orbits



Direct N-Body cont'd

Gravitation is long range interaction
calculation of accelerations is N^2

some ideas to speed up simulations

Direct N-Body cont'd

Gravitation is long range interaction
calculation of accelerations is N^2

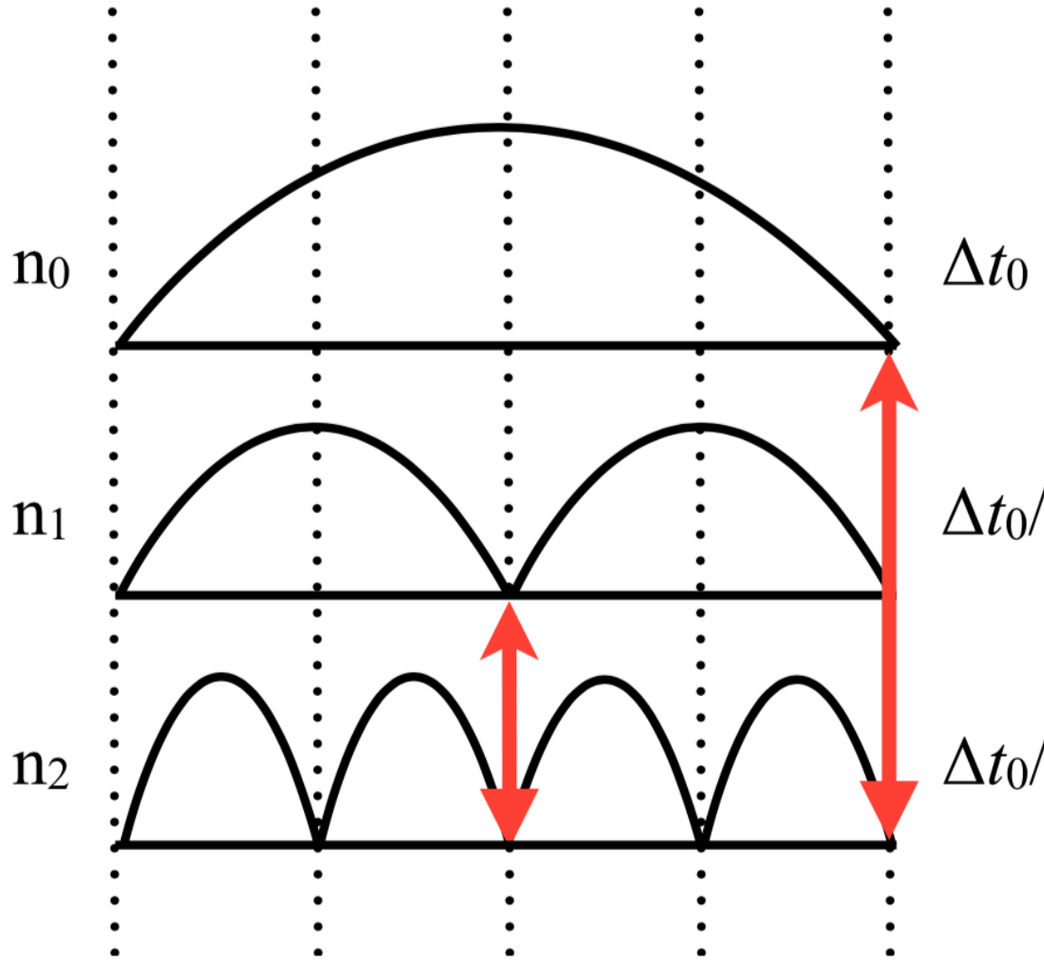
some ideas to speed up simulations

- individual timesteps

$$\Delta t_n = \Delta t_0 / 2^n$$

challenges:
- condition for timestep switch
- individual timesteps break symmetry

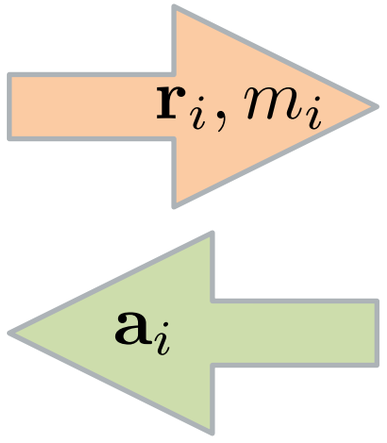
Dehnen & Read 2011

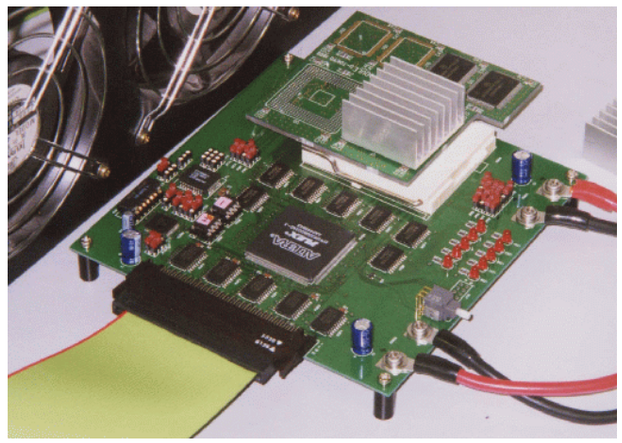


$$\sum_i n_i = N$$

Synchronisation points

Direct N-Body cont'd



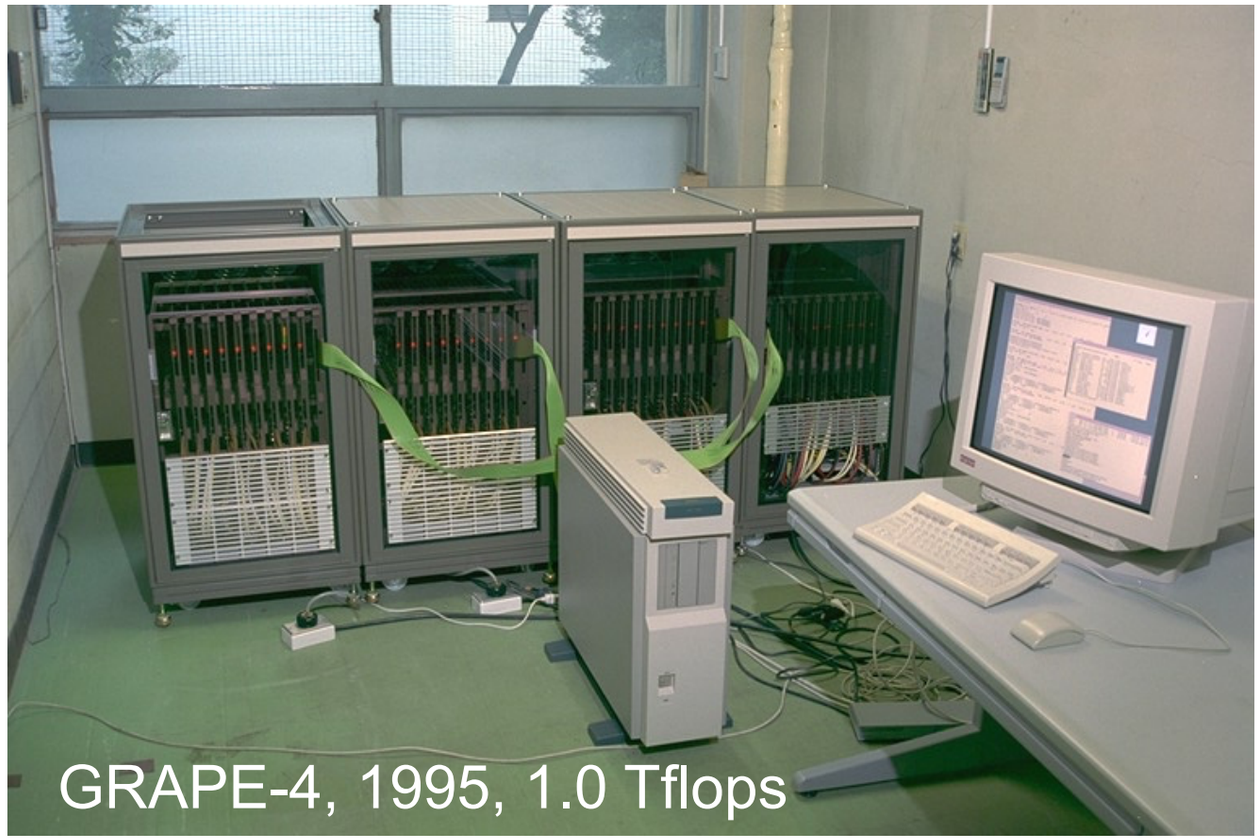


Gravitation is long range interaction
calculation of accelerations is N^2

some ideas to speed up simulations

- special purpose hardware
  GRAvity PipE

last release 200x,
last supercluster 2005:
gravitySimulator, 32 GRAPE boards, 4 Tflops
128'000 particles


GRAPE-4, 1995, 1.0 Tflops

Approximate N-Body

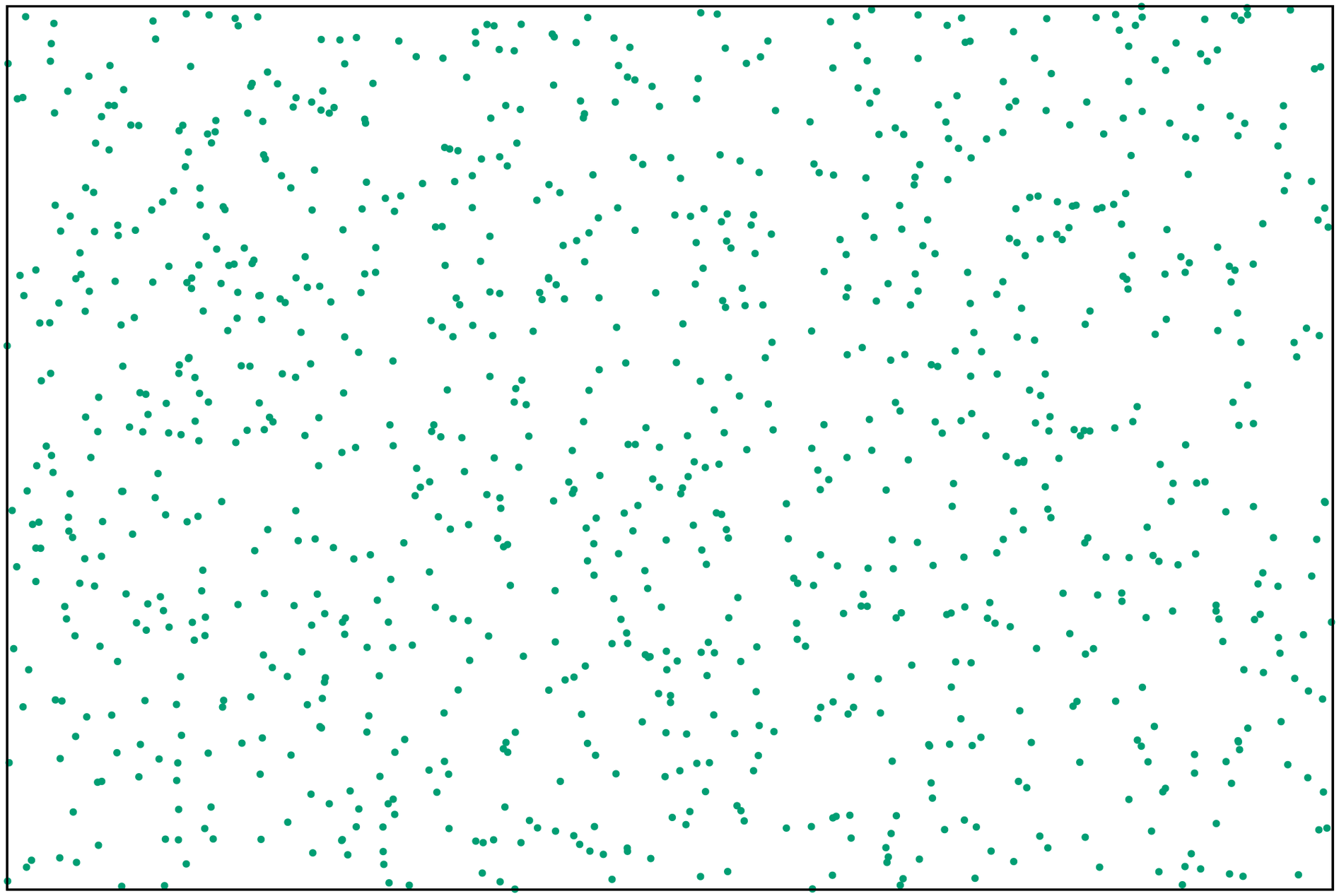
Gravitation is long range interaction
calculation of accelerations is N^2
can we decrease the computations for the cost of lower accuracy?

Barnes-Hut (1986) · hierarchical tree method
basic idea: **reduce** the number of terms in sum

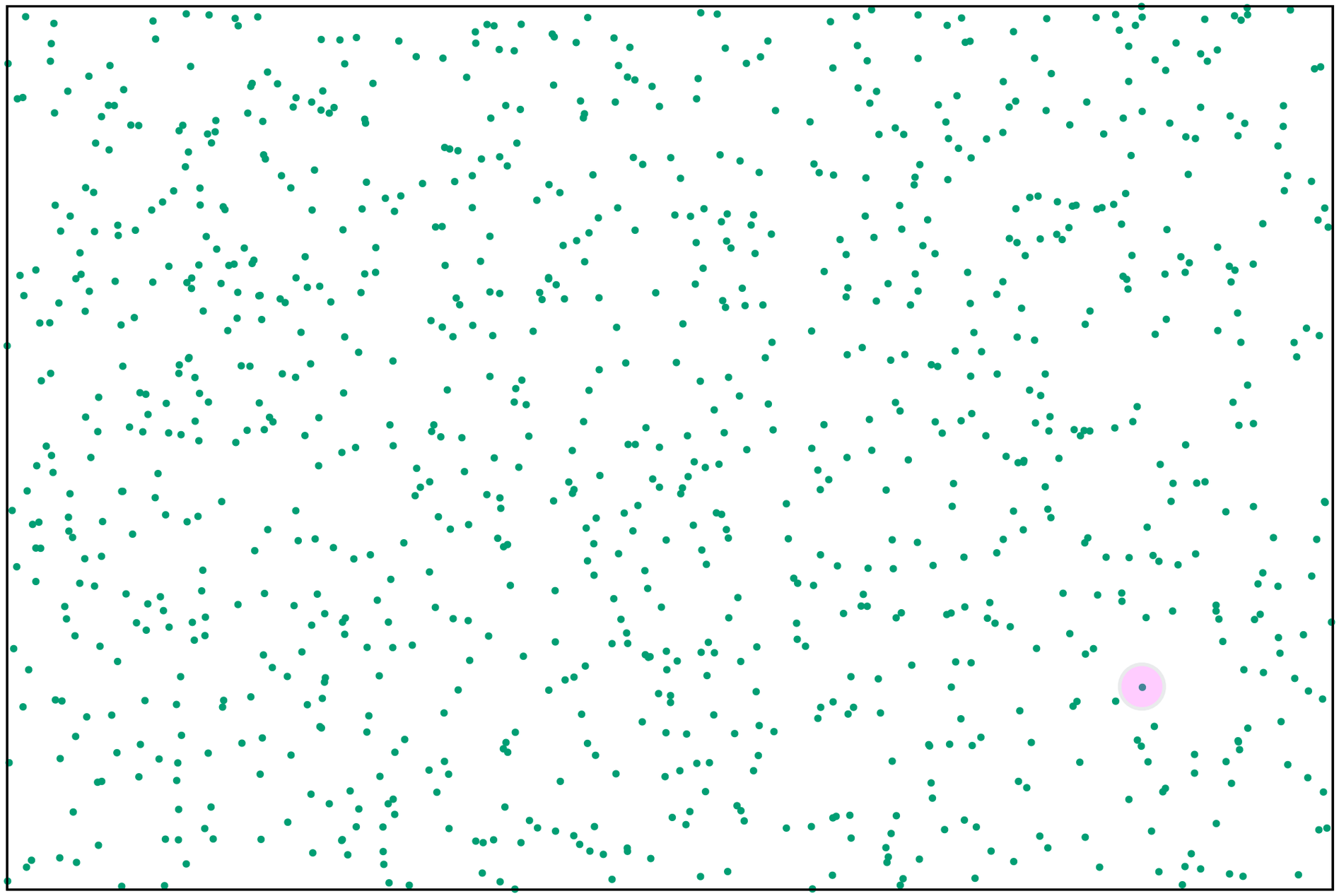
$$\mathbf{a}_i = -G \sum_{j \neq i}^{N} m_j \frac{\mathbf{r}_i - \mathbf{r}_j}{\left|\mathbf{r}_i - \mathbf{r}_j\right|^3}$$

How can we achieve this in a physically correct way?

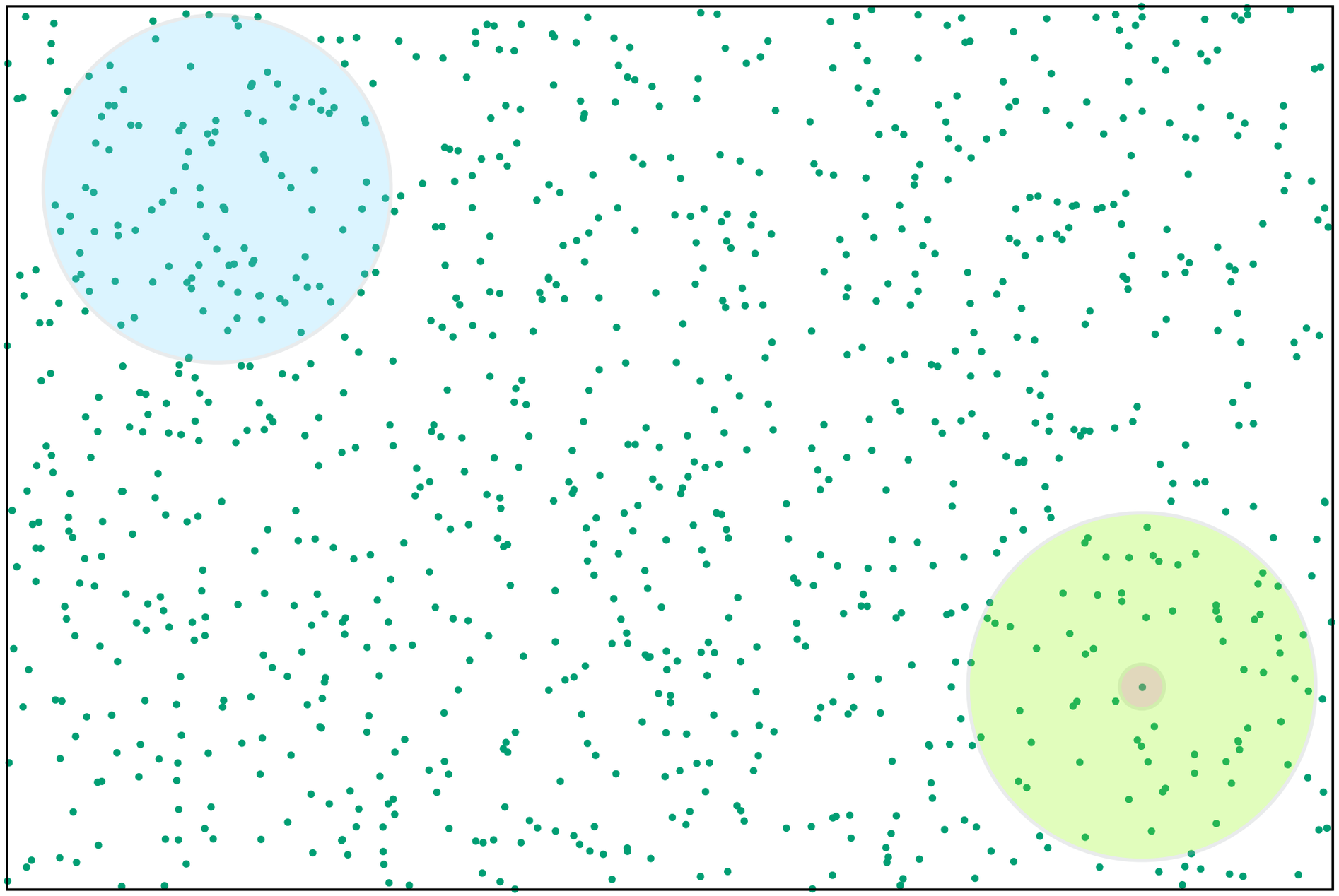
Approximate N-Body - Barnes-Hut Tree



Approximate N-Body - Barnes-Hut Tree

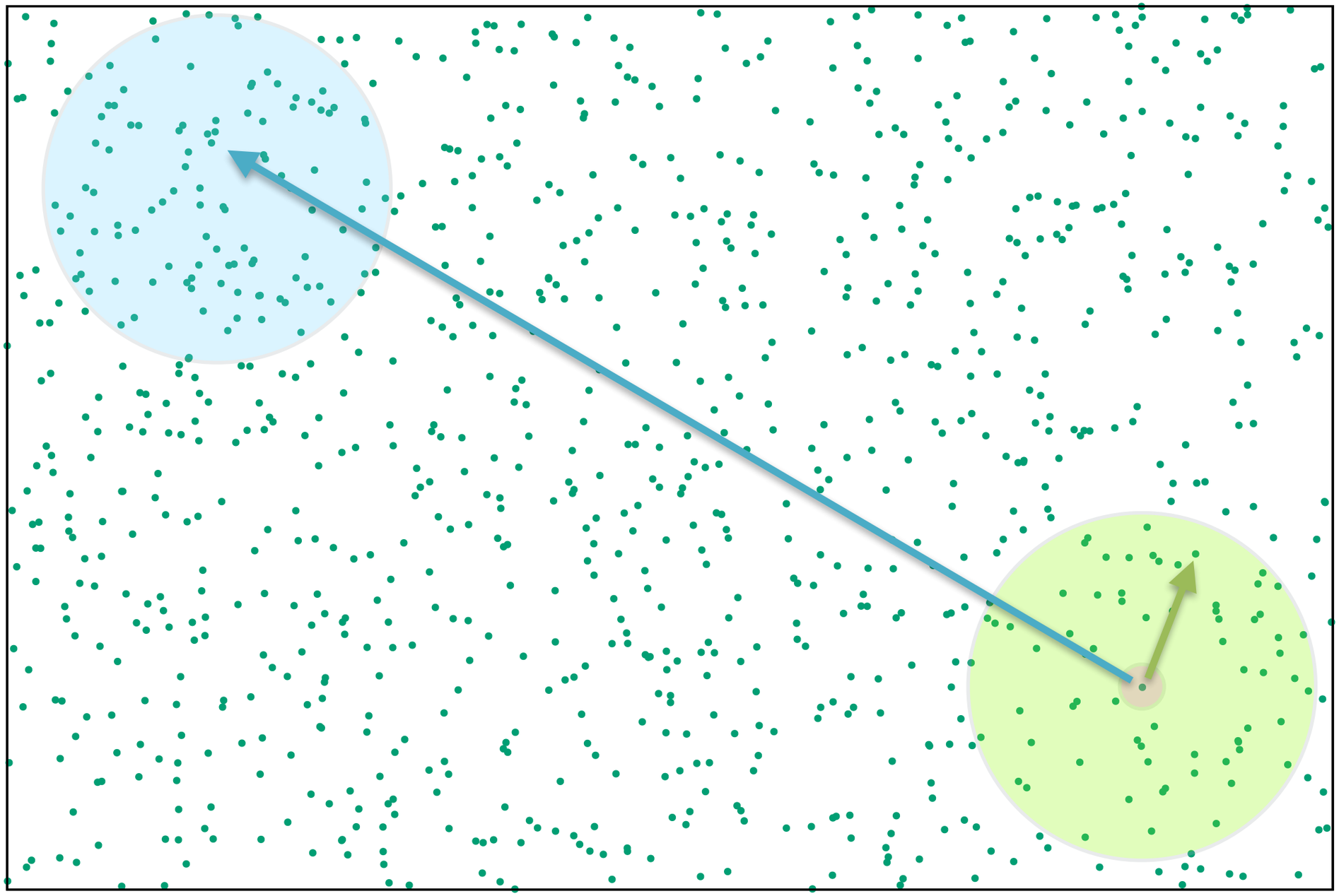


Approximate N-Body - Barnes-Hut Tree



$$\mathbf{a} \sim \frac{1}{r^2}$$

Approximate N-Body - Barnes-Hut Tree



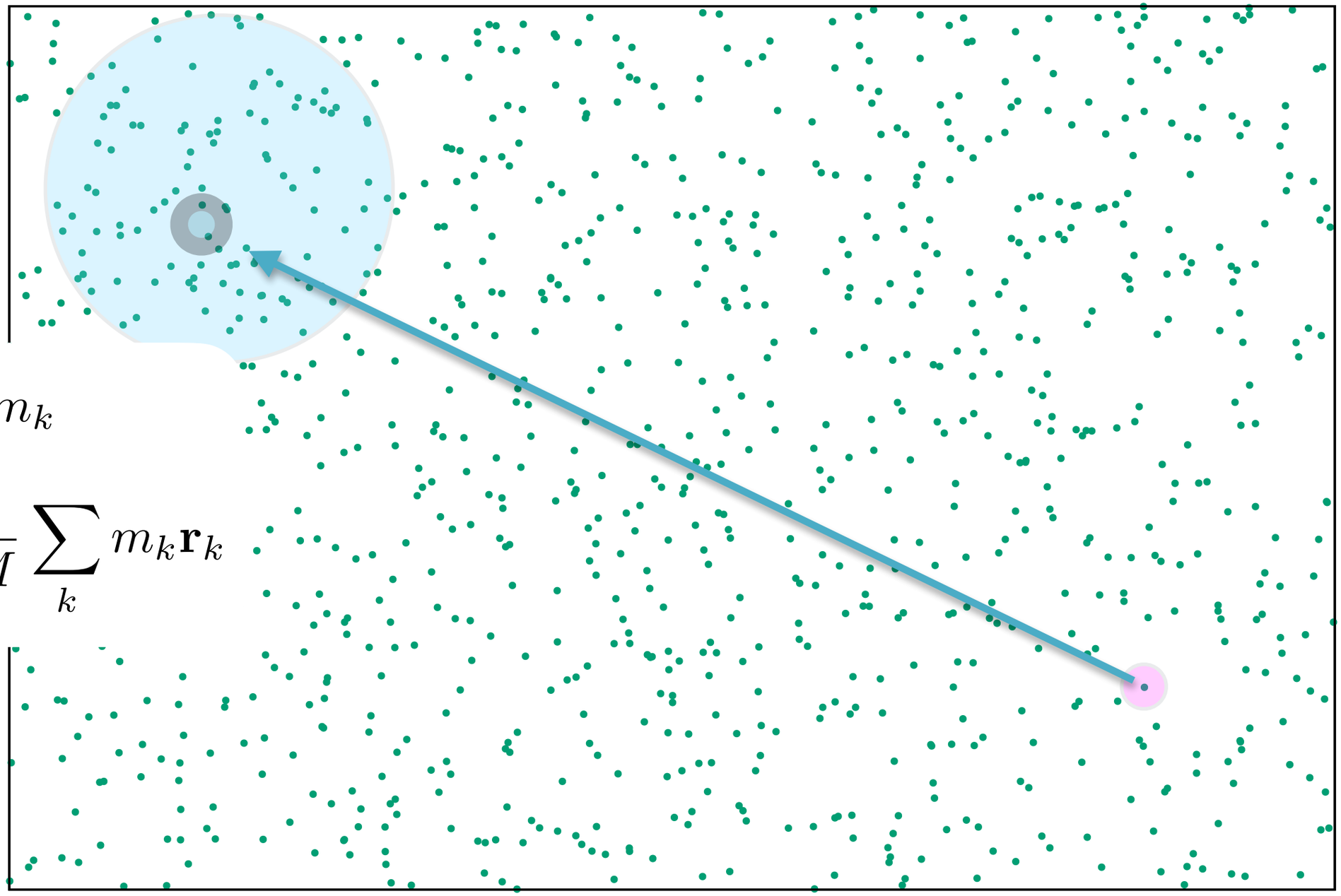
$$\mathbf{a} \sim \frac{1}{r^2}$$

Approximate N-Body - Barnes-Hut Tree

calculate center of mass and mass of blue region

$$M = \sum_k m_k$$

$$\mathbf{r}_{\text{com}} = \frac{1}{M} \sum_k m_k \mathbf{r}_k$$

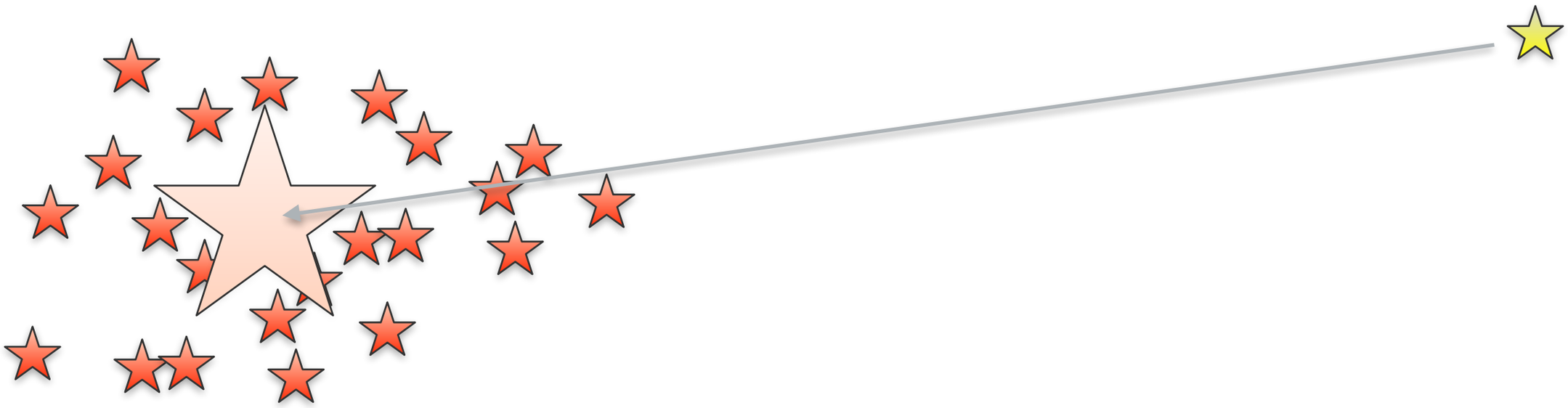


Approximate N-Body - Barnes-Hut Tree

- sort bodies into a hierarchical structure
- add condition for approximation of long range vs. short range

Approximate N-Body - Barnes-Hut Tree

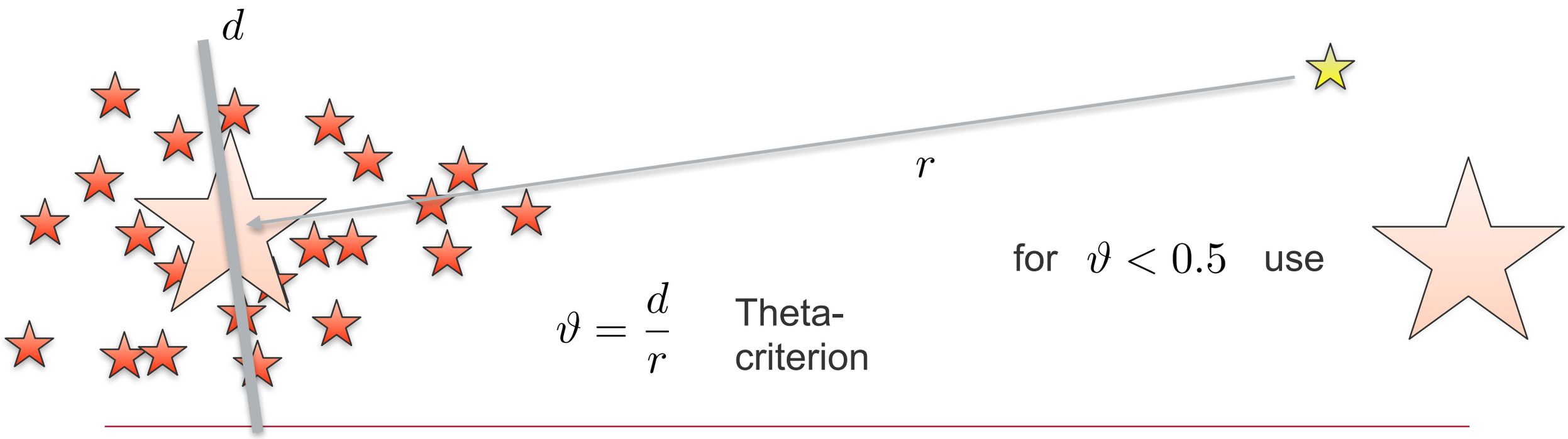
- sort bodies into a hierarchical structure
- add condition for approximation of long range vs. short range



Approximate N-Body - Barnes-Hut Tree

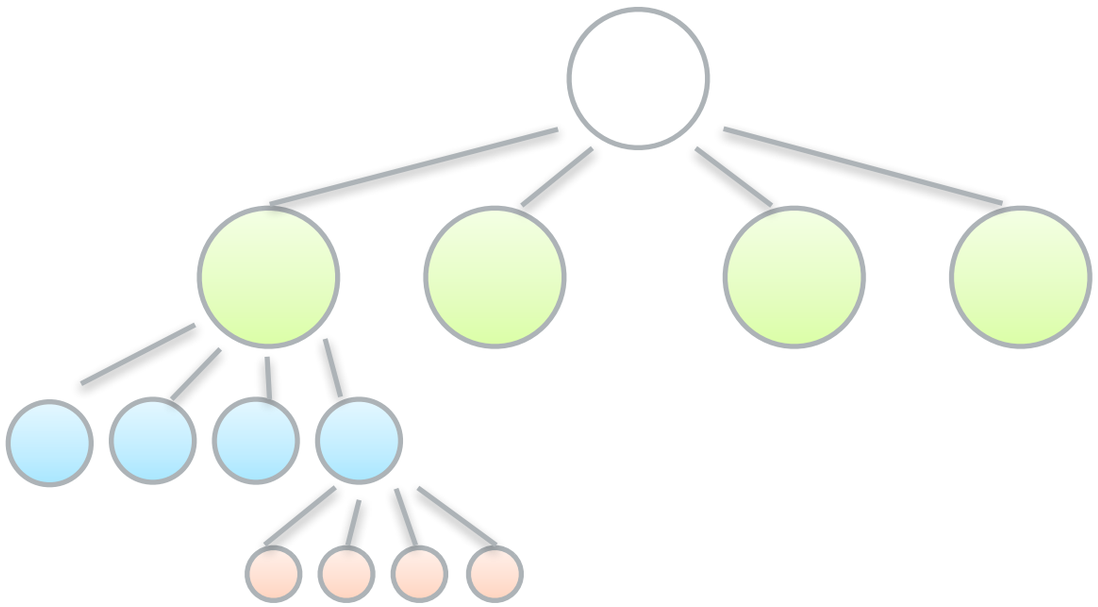
- sort bodies into a hierarchical structure

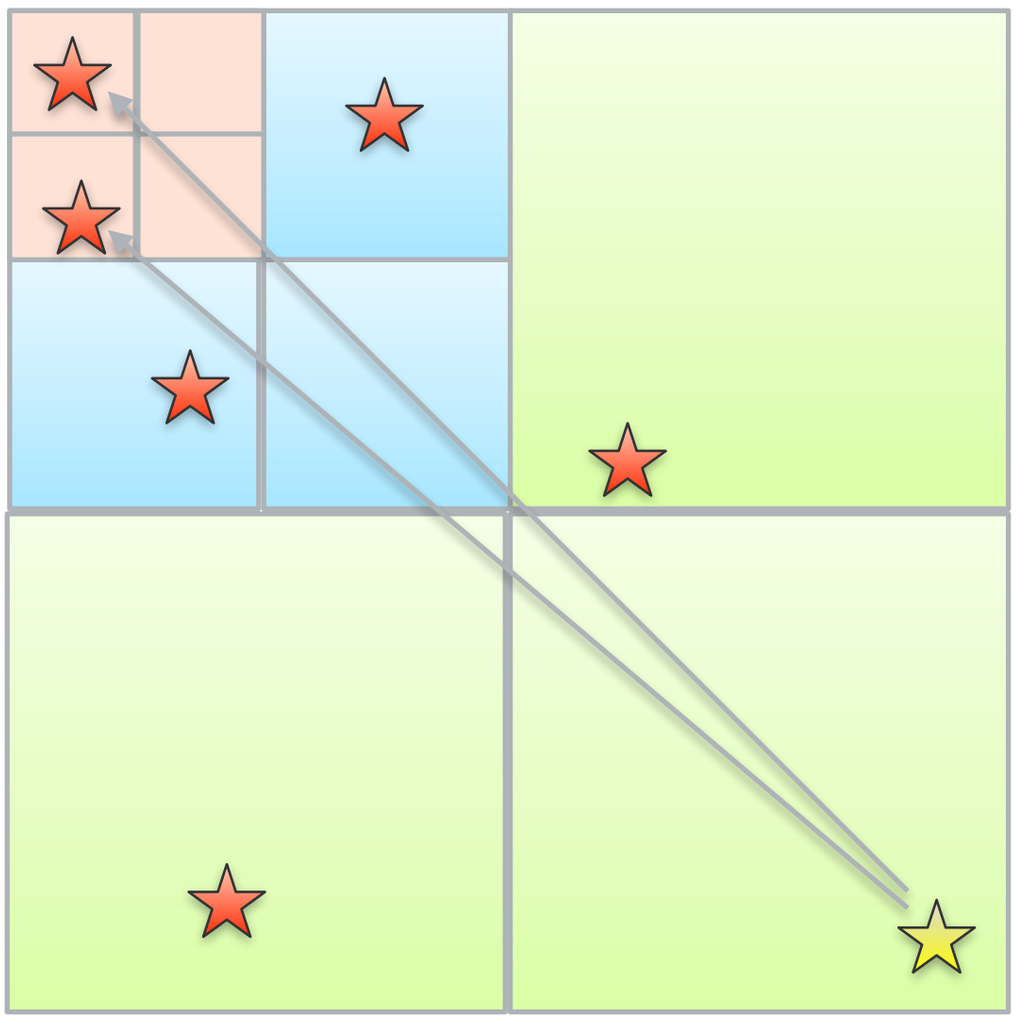
- add condition for approximation of long range vs. short range
  theta criterion



Approximate N-Body - Barnes-Hut Tree

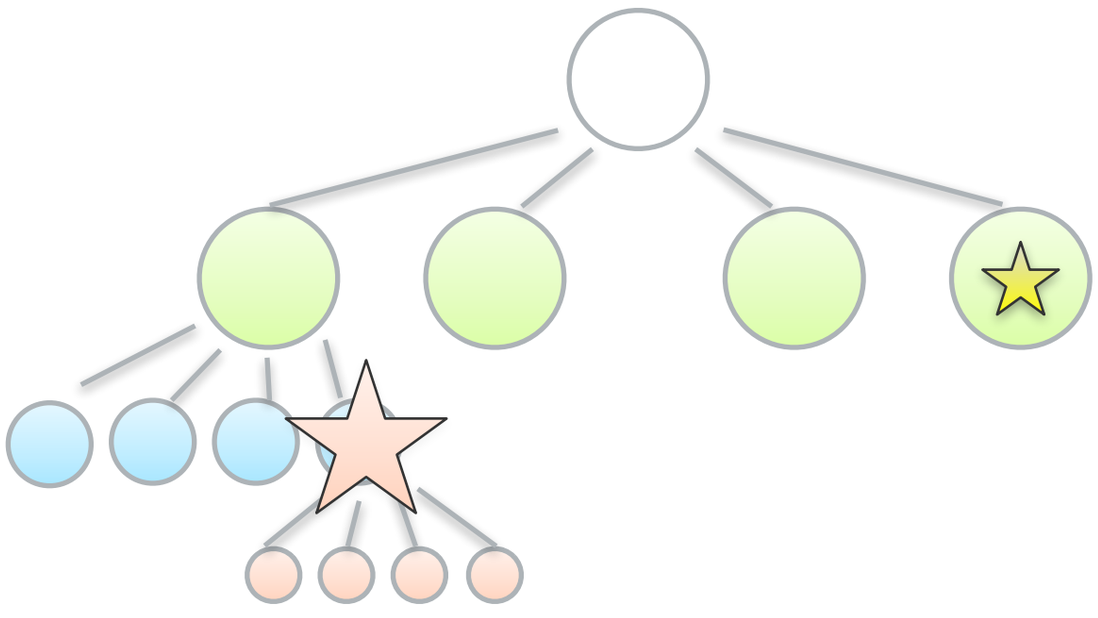
each node has $2^{dimension}$ children:
2D quadtree
3D octree

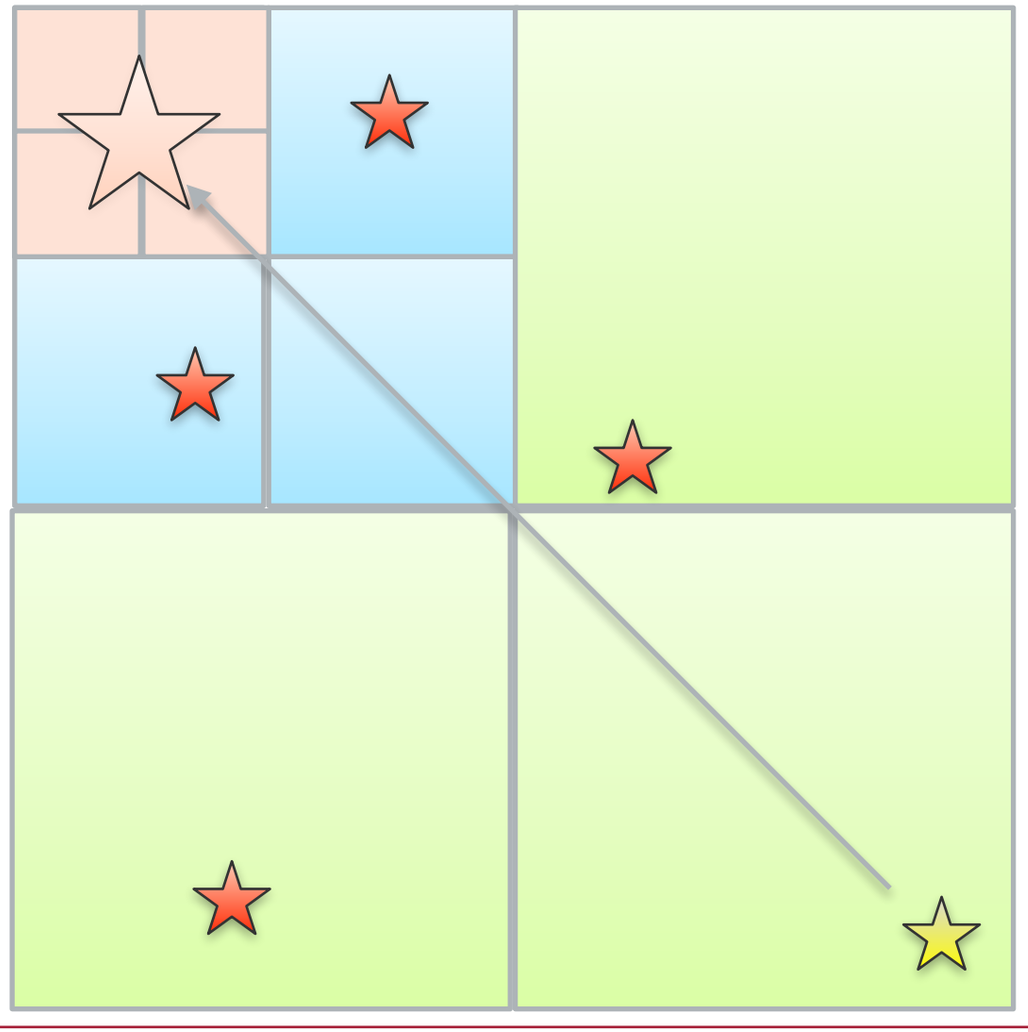




Approximate N-Body - Barnes-Hut Tree

each node has $2^{dimension}$ children:
2D quadtree
3D octree

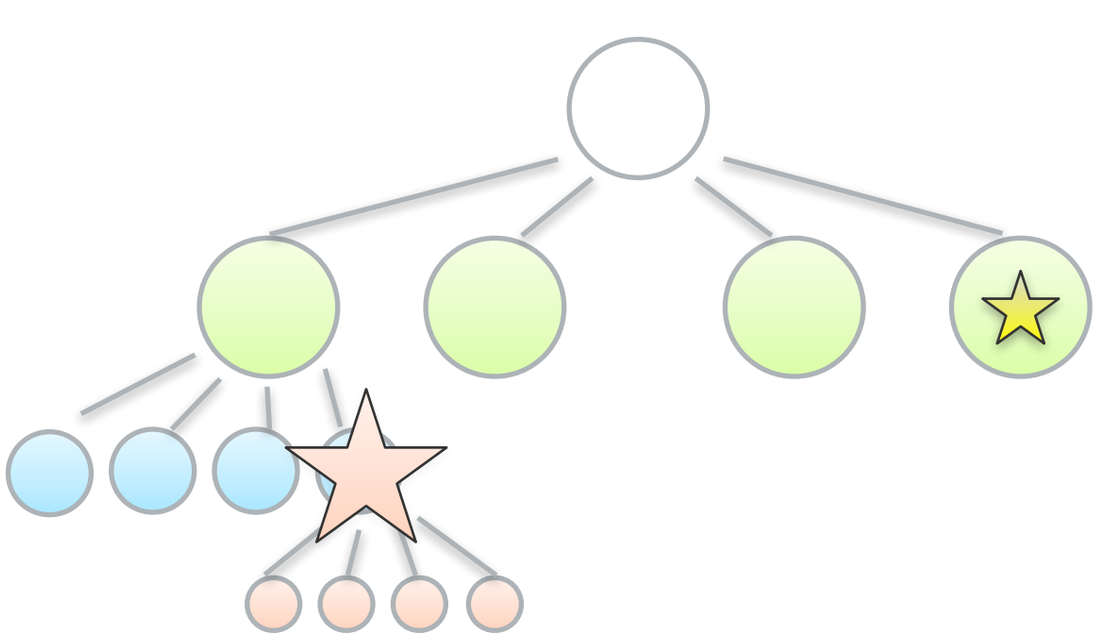


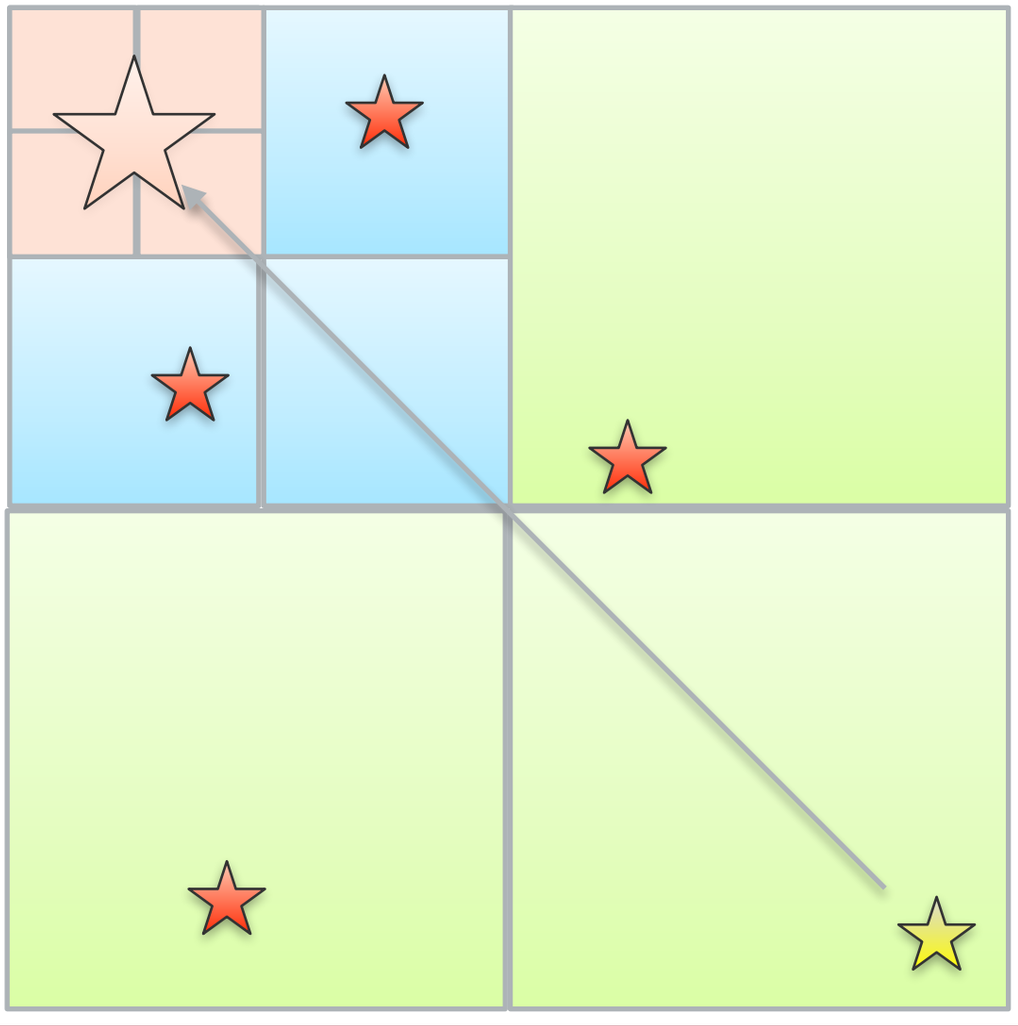


Approximate N-Body - Barnes-Hut Tree

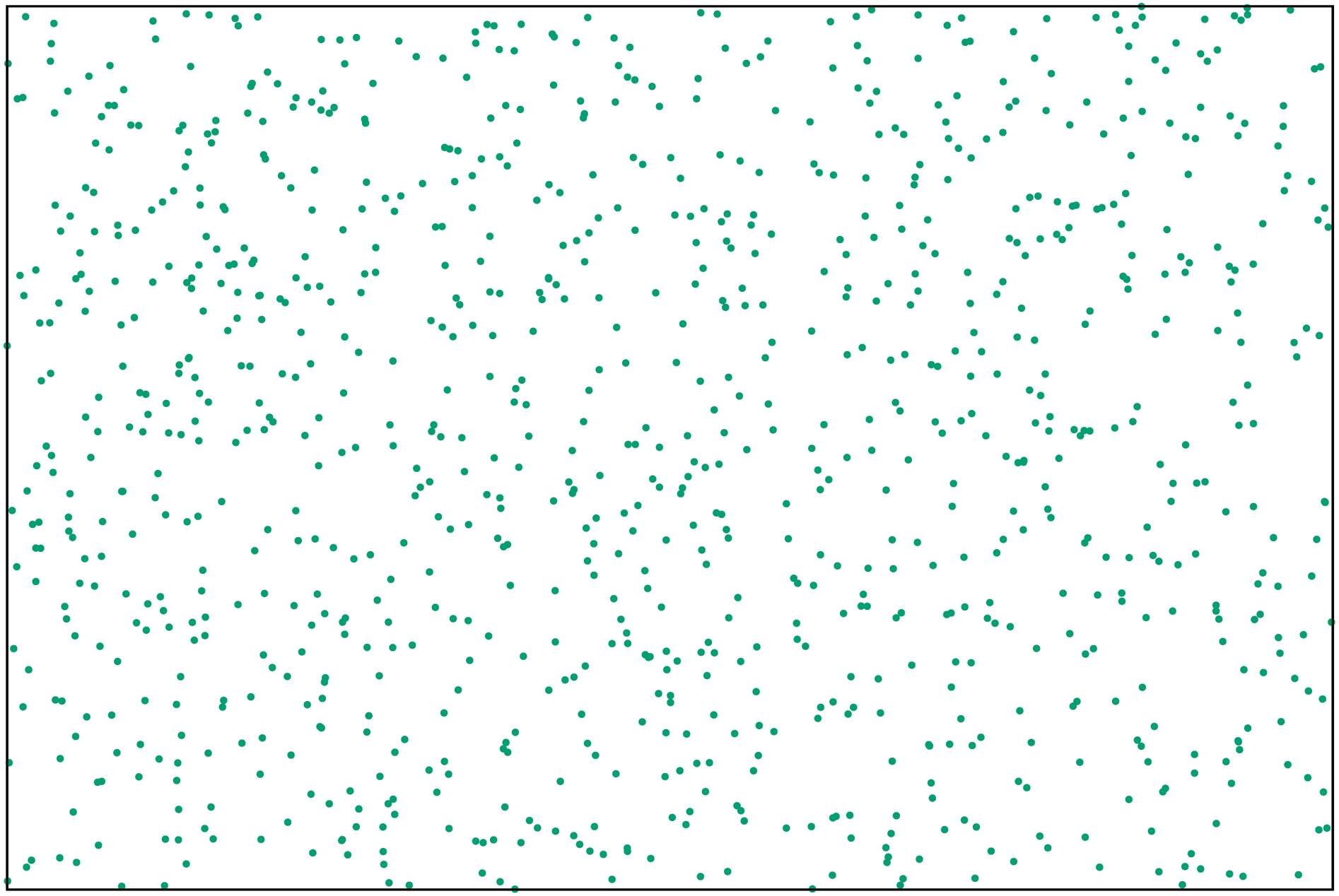
data overhead compared to direct N-Body

computational cost much lower N log(N)

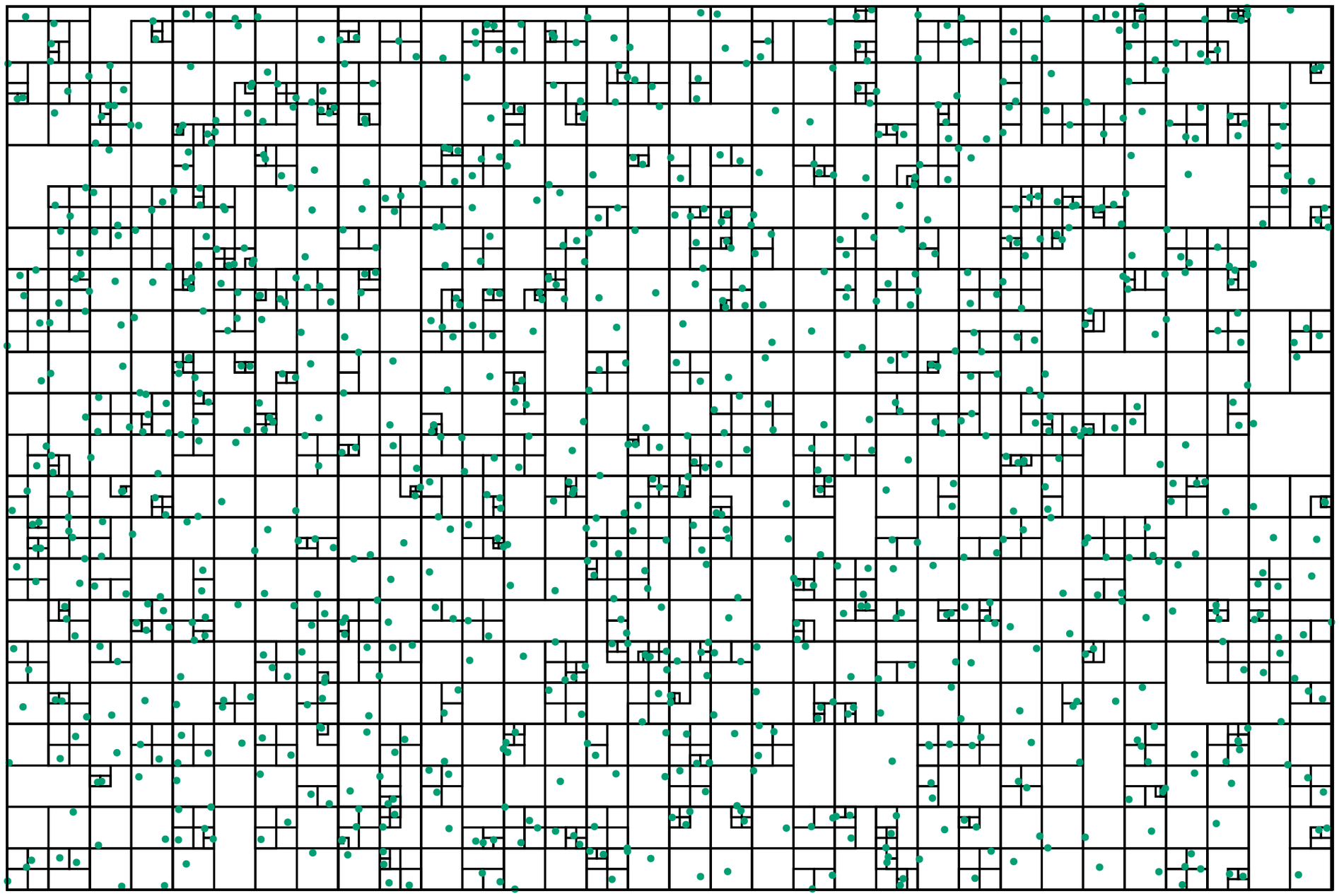




Approximate N-Body - Barnes-Hut Tree

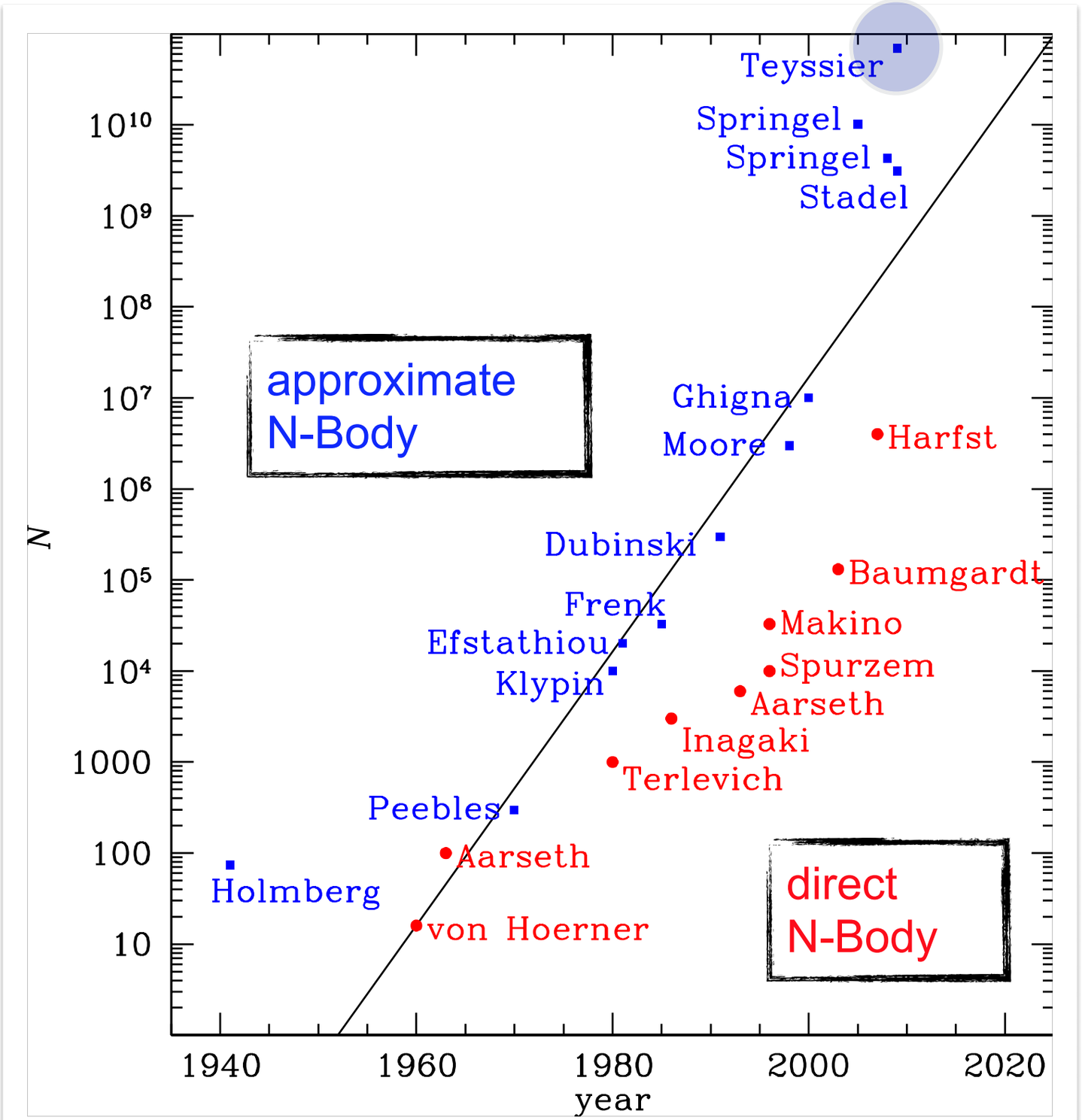


Approximate N-Body - Barnes-Hut Tree



How many particles can we compute?

70 x 10^9



Dehnen & Read 2011

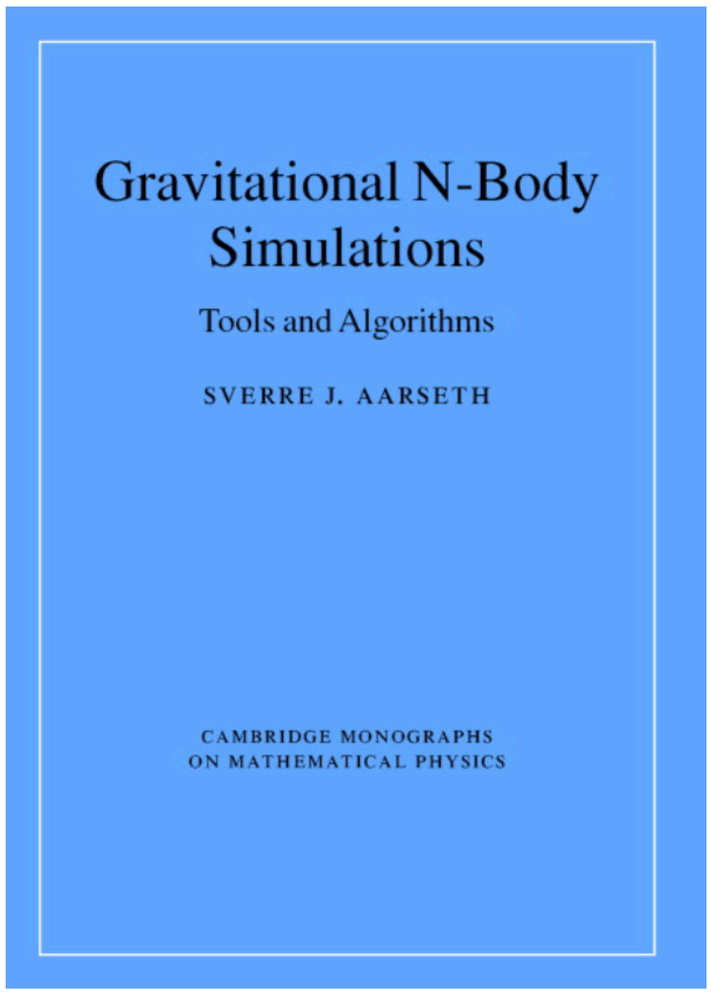
Some N-Body codes you may know

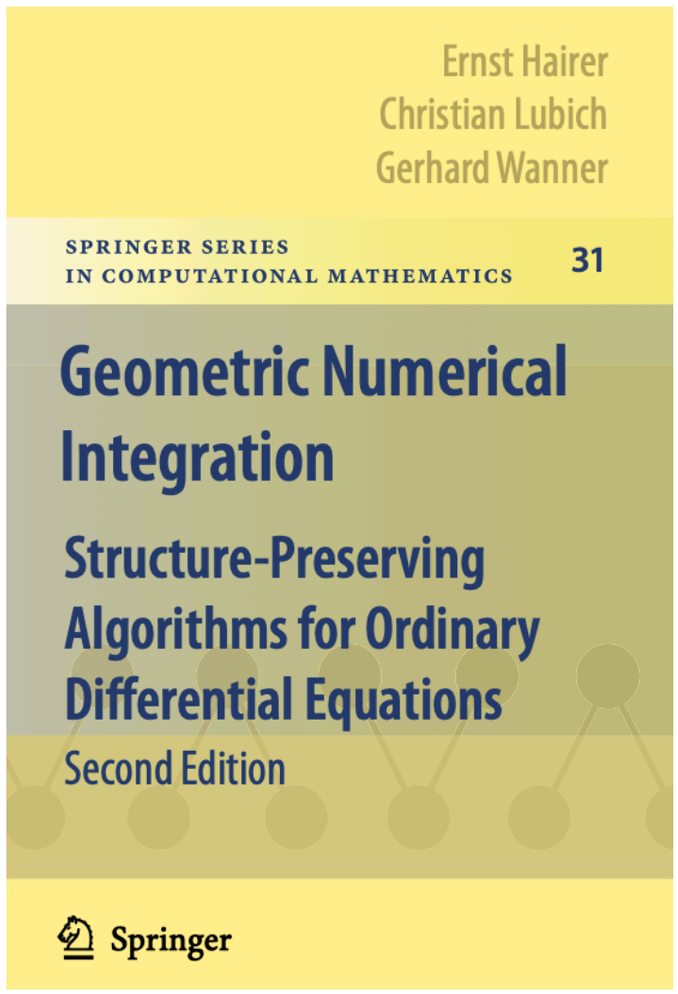
Code	Author	Application	Features
MERCURY	Chambers et al. (2000)	celestial mechanics	symplectic, newer versions from different groups, direct N-Body
NBODY	Aarseth et al. (1985)	star clusters, celestial mechanics	various versions, GPU version, direct N-Body
SyMBA	Duncan et al. (1998)	celestial mechanics	symplectic, direct N-Body
GADGET-2.0	Springel et al. (2005)	galactic dynamics	SPH tree code
AREPO	Springel et al. (2010)	galactic dynamics	moving mesh code
FLASH	Flash consortium	galactic dynamics, accretion discs, star formation	different solvers for gravity: multigrid, tree, …
REBOUND	Rein et al. (2012)	celestial mechanics	😍

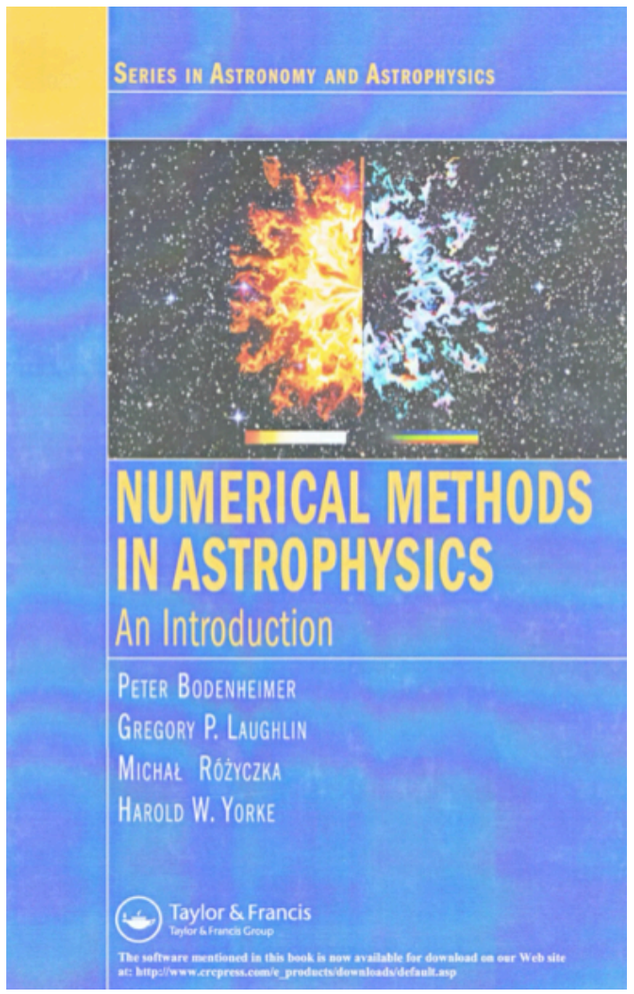
and many more…

More literature for the interested reader:

- Aarseth, Gravitational N-Body Simulations: Tools and Algorithms
- Hairer, Lubich, Wanner, Geometric Numerical Integration
- Bodenheimer, Laughlin, Rozycka, Yorke, Numerical Methods in Astrophysics
- Dehner & Read, N-body simulations of gravitational dynamics, astro-ph







Goal for this experiment: Implement your own code

- Programming language C, C++, python, MATLAB, FORTRAN, julia.
- Read object data (positions, velocities, masses).
- Integration over given time range with a selected integrator and a selected time step size.
- Output of object data to file.
- Postprocessing: plot using gnuplot, matplotlib, … and analyse the data (energy conservation, momentum conservation, …)