N-body issues: softening, tree codes, falcON

Walter Dehnen

Leicester (UK)

contents

- foundations of collisionless N-body methods
- force softening: motivation and limitation
- force approximation: tree code and FMM
- force approximation: falcON

the collisionless N-body code

want to solve:

$$\mathrm{d}_t f = \partial_t f + \boldsymbol{v} \cdot \partial_{\boldsymbol{x}} f - \partial_{\boldsymbol{x}} \Phi \cdot \partial_{\boldsymbol{v}} f = 0$$

$$\partial_{\boldsymbol{x}}^2 \Phi(x, t) = 4\pi G \int f(\boldsymbol{x}, \boldsymbol{v}) \, \mathrm{d}^3\boldsymbol{v}$$

$f(\boldsymbol{x}, \boldsymbol{v})$ continuous, one-particle DF (lowest in BBGKY hierarchy)

the collisionless N-body code

want to solve:

$$\begin{aligned} \mathrm{d}_t f &= \partial_t f + \boldsymbol{v}\cdot\partial_{\boldsymbol{x}} f - \partial_{\boldsymbol{x}}\Phi\cdot\partial_{\boldsymbol{v}} f = 0 \\ \partial_{\boldsymbol{x}}^2 \Phi(x,t) &= 4\pi G \int f(\boldsymbol{x},\boldsymbol{v})\,\mathrm{d}^3\boldsymbol{v} \end{aligned}$$

$f(\boldsymbol{x},\boldsymbol{v})$ continuous, one-particle DF (lowest in BBGKY hierarchy)

- applicable if $t_{\mathrm{relax}} \gg \mathrm{age} \Rightarrow N \gg 10\,\mathrm{age}/t_{\mathrm{dyn}}$
- no analytic non-trivial non-equilibrium solutions exist
- $f(\boldsymbol{x},\boldsymbol{v})$ is 6D & very inhomogeneous $\Rightarrow$ grid methods difficult
- mixing invalidates continuum limit $\Rightarrow$ grid methods fail
- instead: use the Lagrangian ‘method of characteristics’

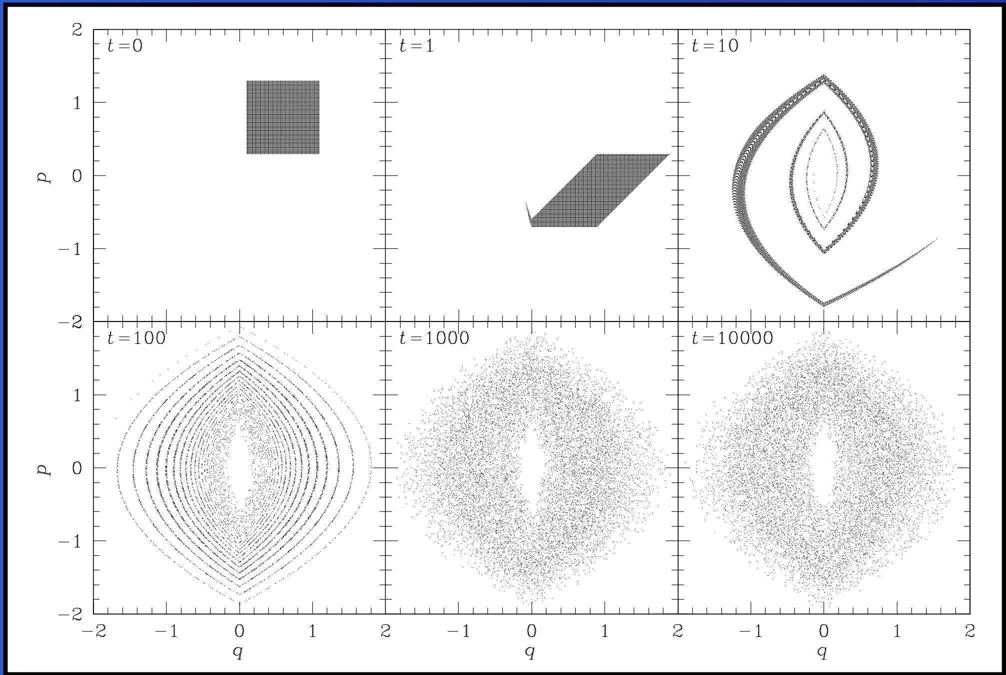
the collisionless N-body code

want to solve:

$$\mathrm{d}_t f = \partial_t f + \boldsymbol{v} \cdot \partial_{\boldsymbol{x}} f - \partial_{\boldsymbol{x}} \Phi \cdot \partial_{\boldsymbol{v}} f = 0$$

$$\partial_{\boldsymbol{x}}^2 \Phi(x, t) = 4\pi G \int f(\boldsymbol{x}, \boldsymbol{v})\, \mathrm{d}^3\boldsymbol{v}$$

$f(\boldsymbol{x}, \boldsymbol{v})$ continuous, one-particle DF (lowest in BBGKY hierarchy)



Phase-mixing of 10^4 points in 1D Hamiltonian $H = p^2/2 + |q|$ of a point mass in 1D gravity. The fine-grained DF is either 1 or 0, but at late times a smooth distribution appears.

How to solve the CBE?

- sample N trajectories $\{\mu_i, \boldsymbol{x}_i, \boldsymbol{v}_i\}$ from $f(\boldsymbol{x}, \boldsymbol{v}, t = 0)$
- solve equations of motion $\ddot{\boldsymbol{x}}_i = -\partial_x \Phi(\boldsymbol{x}_i, t)$
- CBE: $\mu_i = \mathrm{const}$ along trajectories

How to solve the CBE?

- sample N trajectories $\{\mu_i, \boldsymbol{x}_i, \boldsymbol{v}_i\}$ from $f(\boldsymbol{x}, \boldsymbol{v}, t = 0)$
- solve equations of motion $\ddot{\boldsymbol{x}}_i = -\partial_x \Phi(\boldsymbol{x}_i, t)$
- CBE: $\mu_i = \mathrm{const}$ along trajectories

this implies

- $f(\boldsymbol{x}, \boldsymbol{v}, t)$ is unknown (except $f(\boldsymbol{x}_i, \boldsymbol{v}_i, t) = f(\boldsymbol{x}_i, \boldsymbol{v}_i, t = 0)$), but represented by $\{\mu_i, \boldsymbol{x}_i(t), \boldsymbol{v}_i(t)\}$
- automatic *coarse-graining* ($\Rightarrow$ no problems with mixing)
- *moments* of f can be *estimated*
- N $\ll N$ is a *numerical parameter* (unlike collisional N-body)
- $\Rightarrow$ artificial two-body relaxation

How to solve Poisson's equation?

- grid techniques (FFT, multigrid)
  - fast: $\mathcal{O}(n_{\mathrm{grid}} \log n_{\mathrm{grid}})$
  - periodic ($\Rightarrow$ cosmology)
  - problem: inhomogeneity (but: adaptive multigrid)
- basic functions (using Y_{lm})
  - fast: $\mathcal{O}(\mathrm{N} n_{\mathrm{basis}})$
  - problems: central singularity, assumed symmetry

How to solve Poisson's equation?

- Greens-function approach:

$$\Phi(\boldsymbol{x},t) = -G \int \frac{f(\boldsymbol{x}',\boldsymbol{v}',t)}{|\boldsymbol{x}-\boldsymbol{x}'|}\, \mathrm{d}^3\boldsymbol{x}'\, \mathrm{d}^3\boldsymbol{v}'$$

  - general & adaptive
  - problem: $f(\boldsymbol{x},\boldsymbol{v},t)$ unknown $\Rightarrow$ *estimate*

$$\Phi(\boldsymbol{x}_i,t) \approx \hat{\Phi}(\boldsymbol{x}_i,t) = -\sum_{j=1}^{\mathrm{N}} \frac{G\mu_j}{\epsilon}\, \varphi\left(\frac{|\boldsymbol{x}_i-\boldsymbol{x}_j|}{\epsilon}\right)$$

  with **softening kernel** $\varphi(r)$ & **softening length** ϵ

  - $\varphi(r) \sim r^{-1}$ for $r \gg 1$
  - slow: $\mathcal{O}(N^2)$, *but* fast approximations

Why softening?

- Without softening ($\epsilon=0$): N-body instead of N-body problem
- close encounters are
  - hard to integrate (ask Sverre Aarseth)
  - artificially so, since $F \propto \mu/d^2 \propto \mathrm{N}^{-1}/(\mathrm{N}^{-1/3})^2 \propto \mathrm{N}^{-1/3}$.

Why softening?

- Without softening (ϵ=0): N-body instead of N-body problem
- close encounters are
  - hard to integrate (ask Sverre Aarseth)
  - artificially so, since $F \propto \mu/d^2 \propto \mathrm{N}^{-1}/(\mathrm{N}^{-1/3})^2 \propto \mathrm{N}^{-1/3}$.
- thus, softening
  - significantly simplifies the time integration
  - avoids artifacts due to *close* encounters (such as formation of binaries)
  - but *does not avoid artificial two-body relaxation*, because that is driven by encounters on *all* scales $\Rightarrow$ softening can only reduce it by a factor $\sim$2

Softening as Optimal Force Estimation

The *estimated* force

$$\hat{\boldsymbol{F}}(\boldsymbol{x}) = -\sum_{j=1}^{\mathrm{N}} \frac{G\mu_j}{\epsilon^2}\,\varphi'\left(\frac{|\boldsymbol{x}-\boldsymbol{x}_j|}{\epsilon}\right)\frac{\boldsymbol{x}-\boldsymbol{x}_j}{|\boldsymbol{x}-\boldsymbol{x}_j|}$$

is a random variable (like the $\boldsymbol{x}_i$) with mean-square error

$$\mathrm{MSE}(\hat{\boldsymbol{F}}) \equiv \left\langle (\hat{\boldsymbol{F}}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}))^2 \right\rangle = \mathrm{bias}(\hat{\boldsymbol{F}}(\boldsymbol{x}))^2 + \mathrm{var}(\hat{\boldsymbol{F}}(\boldsymbol{x}))$$

with

$$\begin{aligned} \mathrm{bias}(\hat{\boldsymbol{F}}(\boldsymbol{x})) &= \langle \hat{\boldsymbol{F}}(\boldsymbol{x}) \rangle - \boldsymbol{F}(\boldsymbol{x}) \\ \mathrm{var}(\hat{\boldsymbol{F}}(\boldsymbol{x})) &= \left\langle (\hat{\boldsymbol{F}}(\boldsymbol{x}) - \langle \boldsymbol{F}(\boldsymbol{x}) \rangle)^2 \right\rangle = \langle \hat{\boldsymbol{F}}^2(\boldsymbol{x}) \rangle - \langle \hat{\boldsymbol{F}}(\boldsymbol{x}) \rangle^2 \end{aligned}$$

Softening as Optimal Force Estimation

The *estimated* force

$$\hat{\boldsymbol{F}}(\boldsymbol{x}) = -\sum_{j=1}^{\mathrm{N}} \frac{G\mu_j}{\epsilon^2}\,\varphi'\left(\frac{|\boldsymbol{x}-\boldsymbol{x}_j|}{\epsilon}\right)\frac{\boldsymbol{x}-\boldsymbol{x}_j}{|\boldsymbol{x}-\boldsymbol{x}_j|}$$

is a random variable (like the $\boldsymbol{x}_i$) with mean-square error

$$\mathrm{MSE}(\hat{\boldsymbol{F}}) \equiv \left\langle (\hat{\boldsymbol{F}}(\boldsymbol{x}) - \boldsymbol{F}(\boldsymbol{x}))^2 \right\rangle = \mathrm{bias}(\hat{\boldsymbol{F}}(\boldsymbol{x}))^2 + \mathrm{var}(\hat{\boldsymbol{F}}(\boldsymbol{x}))$$

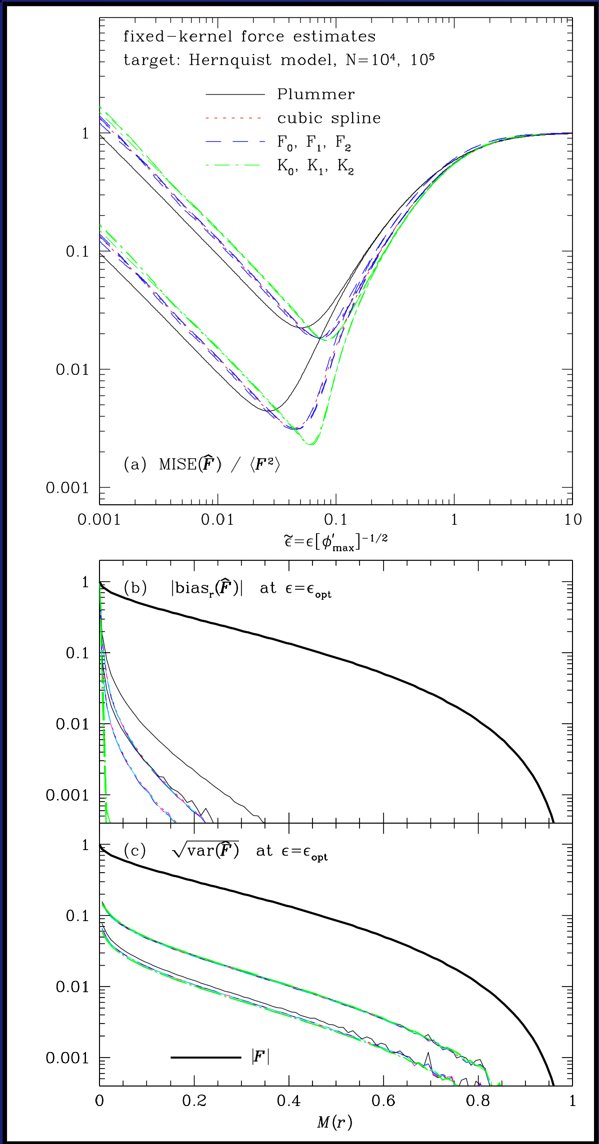
with

$$\begin{aligned} \mathrm{bias}(\hat{\boldsymbol{F}}(\boldsymbol{x})) &= -a_0\epsilon^2 G\partial_x\rho(\boldsymbol{x}) - a_2\epsilon^4 G\partial_x(\partial_x^2\rho(\boldsymbol{x})) + \mathcal{O}(\epsilon^6) \\ \mathrm{N}\,\mathrm{var}(\hat{\boldsymbol{F}}(\boldsymbol{x})) &= b_F\epsilon^{-1}G^2M\rho(\boldsymbol{x}) + \mathcal{O}(\epsilon^0) \end{aligned}$$

and

$$\begin{aligned} a_k &= \frac{4\pi}{(k+1)!}\int_0^\infty \mathrm{d}r\, r^{k+2}(r^{-1} - \varphi(r)) \\ b_F &= 4\pi\int_0^\infty \mathrm{d}r\, r^2\varphi'^2(r) \end{aligned}$$

Softening as Optimal Force Estimation



$$\mathrm{MSE}(\hat{F}) \approx a_0^2 \epsilon^4 G^2 (\partial_x \rho)^2 + b_F \mathrm{N}^{-1} \epsilon^{-1} G^2 M \rho$$

is minimal at

$$\epsilon_{\mathrm{opt}} \propto \mathrm{N}^{-1/5}$$

with

$$\mathrm{MSE}_{\mathrm{opt}}(\hat{F}) \propto \mathrm{N}^{-4/5}$$

Is ϵ_{opt} the best choice for the dynamics?

Top: mean integrated squared error (MISE) of the force for $\mathrm{N} = 10^4$ and $\mathrm{N} = 10^5$ and a Hernquist sphere plotted vs. ϵ for various softening kernels. *Bottom*: run of bias and variance at minimal $\mathrm{MISE}(\hat{F})$.

Force Approximation

- force *estimation error* is unavoidable
- we can tolerate an approximation as long as *approximation error* $\ll$ *estimation error*

Force Approximation

- force *estimation error* is unavoidable
- we can tolerate an approximation as long as
  approximation error $\ll$ *estimation error*
- approximating direct summation: tree code
  - use hierarchical tree (usually: oct-tree) $\Rightarrow$ fully adaptive
  - faster than direct sumation: $\mathcal{O}(\mathrm{N} \log \mathrm{N})$
  - $\boldsymbol{F}_{i \to j} + \boldsymbol{F}_{j \to i} \neq 0 \Rightarrow$ total momentun not conserved

Force Approximation

- force *estimation error* is unavoidable
- we can tolerate an approximation as long as *approximation error* $\ll$ *estimation error*
- approximating direct summation: tree code
  - use hierarchical tree (usually: oct-tree) $\Rightarrow$ fully adaptive
  - faster than direct sumation: $\mathcal{O}(\mathrm{N}\log\mathrm{N})$
  - $\boldsymbol{F}_{i\to j} + \boldsymbol{F}_{j\to i} \neq 0 \Rightarrow$ total momentun not conserved
- approximating direct summation: fast multiple method
  - use hierarchy of cartesian grids $\Rightarrow$ not fully adaptive
  - uses spherical multipoles & complex Y_{lm}
  - numerics complicated & cumbersome
  - high orders $\Rightarrow$ high accuracy
  - formally $\mathcal{O}(\mathrm{N})$ *but* slower than tree code

tree code: details

- preparation phase
  - build a hierarchical tree ⇒ cost: $\mathcal{O}(N \log N)$
  - pre-compute multipole moments etc
- force computation: ‘tree walk’
  - for each body: compute forces due to root cell
  - to compute force from a cell:
    - if body is *well-separated* from cell:
      compute force from multipole moments
    - otherwise: forces from daughter cells (recursive)
  - cost: $\mathcal{O}(\log N)$ per body ⇒ $\mathcal{O}(N \log N)$
- this algorithm clearly is *sub-optimal*:
  forces of neighbours similar yet independently computed

fast multipole method: details

- preparation phase
  - build a hierarchy of cartesian grids
  - pre-compute multipole moments etc (*upward pass*)
- force computation
  - on each grid level: 'intermediate-field' interactions: compute & accumulate multipoles of gravity field
  - *downward pass*
    - pass field-multipoles down the hierarchy
    - compute forces on finest grid
- interaction criterion purely geometric $\Rightarrow$ errors unbalanced
- theoretically $\mathcal{O}(\mathrm{N})$, but not demonstrated in practice
- not competetive with tree code in low-accuracy regime

falcON: details

falcON = force algorithm with complexity $\mathcal{O}(N)$

hybrid of tree & FMM, takes the better of each.

- preparation phase (same as for tree code)
  - build a hierarchical tree ⇒ cost: $\mathcal{O}(N \log N)$
  - pre-compute multipole moments etc

falcON: details

falcON = force algorithm with complexity $\mathcal{O}(N)$

hybrid of tree & FMM, takes the better of each.

- preparation phase (same as for tree code)
- *interaction phase*:
  - compute root-root interaction
  - to compute node-node interaction:
    - if node-node interaction well-separated: accumulate field tensors
    - otherwise: split & continue with child interactions (recursive)
  - cost: (better than) $\mathcal{O}(N)$, dominates

falcON: details

falcON = force algorithm with complexity $\mathcal{O}(N)$
hybrid of tree & FMM, takes the better of each.

- preparation phase (same as for tree code)
- *interaction phase*:
  - compute root-root interaction
  - cost: (better than) $\mathcal{O}(N)$, dominates
- *evaluation phase*:
  - pass field tensors down the tree
  - compute forces at body positions
  - cost: $\mathcal{O}(N)$
- ~ 10 times faster than tree or FMM (at low accuracy)
- simple error balancing (could be done with tree too)

falcON: numerics

wanted: $\Phi(\boldsymbol{x}_i) = -\sum_{j\neq i} \mu_j\, g(\boldsymbol{x}_i - \boldsymbol{y}_j)$. Taylor expand g about $\boldsymbol{R} = \boldsymbol{x}_0 - \boldsymbol{y}_0$:

$$g(\boldsymbol{x}-\boldsymbol{y}) = \sum_{n=0}^{p} \frac{1}{n!}\,(\boldsymbol{x}-\boldsymbol{y}-\boldsymbol{R})^{(n)} \odot \partial_{\boldsymbol{x}}^{(n)} g(\boldsymbol{R}) \;+\; \mathcal{R}_p(g).$$

Insert & sum over source cell B (Warren & Salmon 1995, CPC, 87, 266):

$$\begin{aligned}
\Phi_{\mathsf{B}\to\mathsf{A}}(\boldsymbol{x}) &= -\sum_{m=0}^{p} \frac{1}{m!}\,(\boldsymbol{x}-\boldsymbol{x}_0)^{(m)} \odot \mathbf{C}^{m,p} \;+\; \mathcal{R}_p(\Phi_{\mathsf{B}\to\mathsf{A}}) \\
\mathbf{C}^{m,p} &= \sum_{n=0}^{p-m} \frac{(-1)^n}{n!}\, \partial_{\boldsymbol{x}}^{(n+m)} g(\boldsymbol{R}) \odot \mathbf{M}_{\mathsf{B}}^{n}, \\
\mathbf{M}_{\mathsf{B}}^{n} &= \sum_{\boldsymbol{y}_i\in\mathsf{B}} \mu_i\,(\boldsymbol{y}_i-\boldsymbol{y}_0)^{(n)}.
\end{aligned}$$

falcON: numerics

wanted: $\Phi(\boldsymbol{x}_i) = -\sum_{j\neq i}\mu_j\, g(\boldsymbol{x}_i - \boldsymbol{y}_j)$. Taylor expand g about $\boldsymbol{R} = \boldsymbol{x}_0 - \boldsymbol{y}_0$:

$$g(\boldsymbol{x}-\boldsymbol{y}) = \sum_{n=0}^{p} \frac{1}{n!}\,(\boldsymbol{x}-\boldsymbol{y}-\boldsymbol{R})^{(n)} \odot \partial_{\boldsymbol{x}}^{(n)} g(\boldsymbol{R}) \;+\; \mathcal{R}_p(g).$$

Insert & sum over source cell B (Warren & Salmon 1995, CPC, 87, 266):

$$\Phi_{\mathrm{B}\to\mathrm{A}}(\boldsymbol{x}) = -\sum_{m=0}^{p} \frac{1}{m!}\,(\boldsymbol{x}-\boldsymbol{x}_0)^{(m)} \odot \mathbf{C}^{m,p} \;+\; \mathcal{R}_p(\Phi_{\mathrm{B}\to\mathrm{A}}) \qquad \text{evaluation}$$

$$\mathbf{C}^{m,p} = \sum_{n=0}^{p-m} \frac{(-1)^n}{n!}\,\partial_{\boldsymbol{x}}^{(n+m)} g(\boldsymbol{R}) \odot \mathbf{M}_{\mathrm{B}}^{n}, \qquad \text{interaction}$$

$$\mathbf{M}_{\mathrm{B}}^{n} = \sum_{\boldsymbol{y}_i\in\mathrm{B}} \mu_i\,(\boldsymbol{y}_i-\boldsymbol{y}_0)^{(n)}. \qquad \text{preparation}$$

$\sum_m$: evaluation of gravity, represented by the *field tensors* $\mathbf{C}^{m,p}$, at position $\boldsymbol{x}$

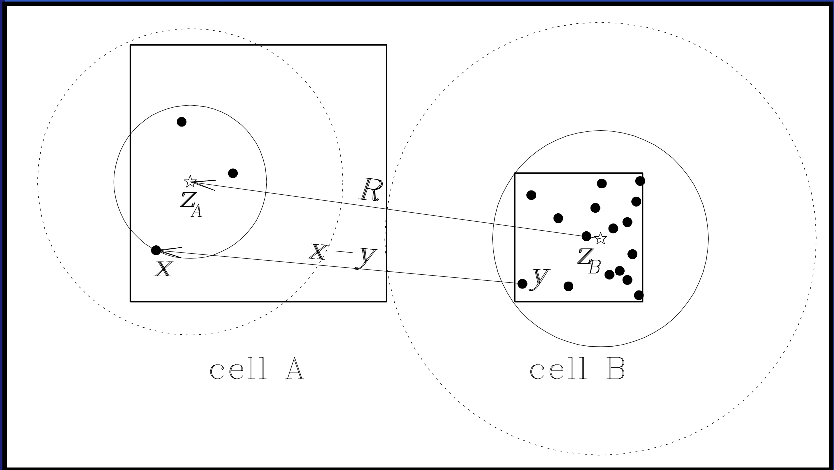
$\sum_n$: interaction between source cell B, represented by the *multipoles* $\mathbf{M}_{\mathrm{B}}^{n}$, and the sink cell A.

Difference to tree code:

expansion in $\boldsymbol{x}$ (tree code: $\boldsymbol{x} \equiv \boldsymbol{x}_0$)

mutuality of *inter*actions

falcON: error balancing



cells are *well-separated* if

$|\boldsymbol{R}| > r_{\mathrm{A,crit}} + r_{\mathrm{B,crit}}$ with $r_{\mathrm{crit}} = r_{\max}/\theta$,

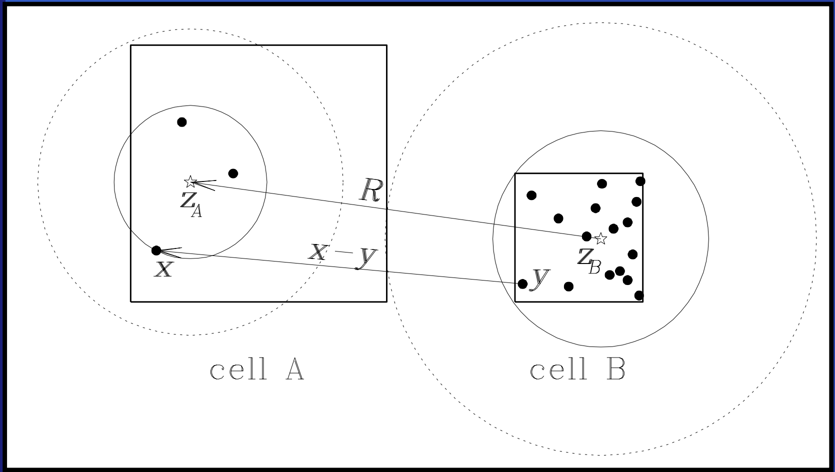
$\Rightarrow |\boldsymbol{x} - \boldsymbol{y} - \boldsymbol{R}| < \theta |\boldsymbol{R}| \; \forall \, \boldsymbol{x} \in \mathrm{A}, \; \boldsymbol{y} \in \mathrm{B}$

$\Rightarrow$ Taylor series converges

error of individual interaction:

$$|\partial_{\boldsymbol{x}} \mathcal{R}_p(\Phi_{\mathrm{B}\to\mathrm{A}})| \le \frac{(p+1)\theta^p}{(1-\theta)^2} \frac{\mathrm{M_B}}{R^2} \propto \frac{\theta^{p+2}}{(1-\theta)^2} \mathrm{M_B^{1/3}}$$

falcON: error balancing



cells are *well-separated* if

$|\boldsymbol{R}| > r_{\text{A,crit}} + r_{\text{B,crit}}$ with $r_{\text{crit}} = r_{\max}/\theta$,

$\Rightarrow |\boldsymbol{x} - \boldsymbol{y} - \boldsymbol{R}| < \theta|\boldsymbol{R}| \; \forall\, \boldsymbol{x} \in \text{A},\; \boldsymbol{y} \in \text{B}$

$\Rightarrow$ Taylor series converges

error of individual interaction:

$$|\partial_{\boldsymbol{x}} \mathcal{R}_p(\Phi_{\text{B}\to\text{A}})| \le \frac{(p+1)\theta^p}{(1-\theta)^2}\,\frac{\text{M}_\text{B}}{R^2} \propto \frac{\theta^{p+2}}{(1-\theta)^2}\,\text{M}_\text{B}^{1/3}$$

with $\theta = \text{const}$: *relative* error constant, *absolute* error $\propto \text{M}_\text{B}^{1/3}$

$\Rightarrow$ *total error* dominated by few interactions with large cells

better: *balance* absolute errors by $\theta = \theta(\text{M})$ with

$$\frac{\theta^{p+2}}{(1-\theta)^2} = \frac{\theta_{\min}^{p+2}}{(1-\theta_{\min})^2}\left(\frac{\text{M}}{\text{M}_{\text{tot}}}\right)^{1/3}$$

falcON: complexity analysis

- eight-folding N $\Rightarrow N_I \to 8N_I + N_+$ and thus

$$\frac{\mathrm{d}N_I}{\mathrm{dN}} \simeq \frac{N_I}{\mathrm{N}} \frac{\Delta \ln N_I}{\Delta \ln \mathrm{N}} \approx \frac{N_I}{\mathrm{N}} + \frac{N_+}{\mathrm{N} 8 \ln 8},$$

with solution

$$N_I = c_0 \mathrm{N} + \frac{\mathrm{N}}{8 \ln 8} \int \frac{N_+}{\mathrm{N}^2} \,\mathrm{dN}$$

falcON: complexity analysis

- eight-folding N $\Rightarrow N_I \to 8N_I + N_+$ and thus

$$\frac{\mathrm{d}N_I}{\mathrm{dN}} \simeq \frac{N_I}{\mathrm{N}} \frac{\Delta \ln N_I}{\Delta \ln \mathrm{N}} \approx \frac{N_I}{\mathrm{N}} + \frac{N_+}{\mathrm{N}8 \ln 8},$$

with solution

$$N_I = c_0 \mathrm{N} + \frac{\mathrm{N}}{8 \ln 8} \int \frac{N_+}{\mathrm{N}^2} \, \mathrm{dN}$$

- tree code: $N_+ \propto \mathrm{N} \Rightarrow N_I \propto \mathrm{N} \log \mathrm{N}$
- falcON: $N_+(\mathrm{N})$ grow less than $\mathcal{O}(\mathrm{N}) \Rightarrow N_I \propto \mathrm{N}$
- with error-balancing: at error $\simeq$ const: $N_I \propto \mathrm{N}^{0.91}$
- typical for $\mathrm{N} = 10^6$ CPU time: 0.5 + 4.5 sec (Opteron)

to-do & yet-to-understand list

- softening: dynamical effects, resolution & noise
- force approximation:
  - effects of approximation errors, shadow forces (grid)
  - non-conservative force approximation: a problem?
- time integration: what is right & wrong?
- comparative N-body studies (Katrin Heitmann)
- initial conditions