

Non-bonded interactions

speeding up the number-crunching

Marcus Elstner and Tomáš Kubař

May 8, 2015

Non-bonded interactions – why care?

- key to understand biomolecular structure and function
 - binding of a ligand
 - efficiency of a reaction
 - color of a chromophore
- two-body potentials \rightarrow computational effort of $\mathcal{O}(N^2)$
 - good target of optimization
- solvent (H_2O) – crucial role, huge amount
 - efficient description needed

How many pair-wise interactions are there?

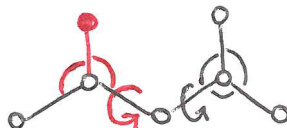
imagine we introduce an additional atom into a system
that already has $N - 1$ atoms

How many pair-wise interactions are there?

imagine we introduce an additional atom into a system
that already has $N - 1$ atoms

bonded interactions

- we add at most (roughly)
2 bonds, 2 angles, 3 dihedrals
- for N atoms, this is at most $7N$
– proportional to N : $O(N)$

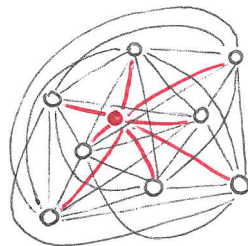


How many pair-wise interactions are there?

imagine we introduce an additional atom into a system
that already has $N - 1$ atoms

non-bonded interactions

- between the new atom and each of the previous atoms:
 $N - 1$ interactions!
- for N atoms, this is $N(N - 1)/2$
– proportional to N^2 : $\mathcal{O}(N^2)$



How many pair-wise interactions are there?

Let us assume that the calculation of every atom–atom interaction takes the same amount of time

Then, the $\mathcal{O}(N^2)$ evaluation of non-bonded interactions will be the most computationally intensive procedure in the entire simulation (the bottle neck)

Coulomb's law

elstat. interaction energy of point charges q and Q on distance r :

$$E^{\text{el}} = \frac{1}{4\pi\epsilon_0} \cdot \frac{q \cdot Q}{r}$$

electrostatic potential (ESP) induced at \vec{r} by Q at \vec{r}_1 :

$$\phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \cdot \frac{Q}{|\vec{r} - \vec{r}_1|}$$

ESP induced by a number of point charges Q_i :

$$\Phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \sum_i \frac{Q_i}{|\vec{r} - \vec{r}_i|}$$

if we know ESP at \vec{r} , and q is placed there, elstat. energy follows as

$$E^{\text{el}}(\vec{r}) = \Phi(\vec{r}) \cdot q$$

→ 'electrostatic potential energy surface'

Coulomb's law

- continuous charge distribution – **charge density** $\rho = \partial Q / \partial V$
- charge in a volume element V_i is then
$$Q_i = \rho(\vec{r}_i) \cdot V_i = \rho(\vec{r}_i) \cdot \Delta V$$
- summing the potential induced by all elements gives

$$\Phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \sum_i \frac{\rho(\vec{r}_i) \cdot \Delta V}{|\vec{r} - \vec{r}_i|}$$

- for infinitesimal volume elements (with $d^3\vec{r} = dV$):

$$\Phi(\vec{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\vec{r}_1)}{|\vec{r} - \vec{r}_1|} d^3\vec{r}_1$$

- elstat. energy of a charge density $\rho(\vec{r})$ follows as

$$E = \frac{1}{2} \int \Phi(\vec{r}) \cdot \rho(\vec{r}) dV = \frac{1}{8\pi\epsilon_0} \iiint \frac{\rho(\vec{r}_1) \cdot \rho(\vec{r})}{|\vec{r} - \vec{r}_1|} d^3\vec{r} d^3\vec{r}_1$$

Poisson's equation

- needs to be solved to get ESP from charge distribution (differential equation for Φ as a function of \vec{r}):

$$\nabla^2 \Phi(\vec{r}) = -\frac{\rho(\vec{r})}{\varepsilon}$$

- if permittivity ε is not constant:

$$\nabla (\varepsilon \nabla \Phi(\vec{r})) = -\rho(\vec{r})$$

example: ESP of Gaussian charge density around \vec{o} with width σ :

$$\rho(r) = Q \cdot \frac{1}{\sigma^3 \sqrt{2\pi}^3} \cdot \exp \left[-\frac{r^2}{2\sigma^2} \right]$$

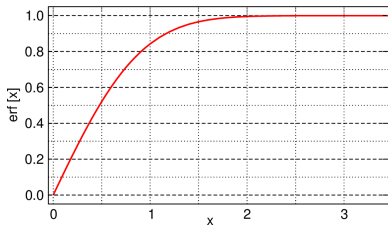
solution of Poisson's equation:

$$\Phi(r) = \frac{1}{4\pi\varepsilon} \cdot \frac{Q}{r} \cdot \operatorname{erf} \left[\frac{r}{\sqrt{2}\sigma} \right]$$

Poisson's equation

solution for Gaussian charge density:

$$\Phi(r) = \frac{1}{4\pi\epsilon} \cdot \frac{Q}{r} \cdot \operatorname{erf}\left[\frac{r}{\sqrt{2}\sigma}\right]$$



Poisson's equation

solution for Gaussian charge density:

$$\Phi(r) = \frac{1}{4\pi\epsilon} \cdot \frac{Q}{r} \cdot \operatorname{erf} \left[\frac{r}{\sqrt{2}\sigma} \right]$$

if we move far from the center of charge density (r is large)

- erf converges to 1, ESP equals that of a point charge placed in \vec{o}
- accordance with experience – a point charge and a well-localized charge density interact with distant charges in the same way
- actually, we need not go far to see that
 - erf = 0.999 already for $x = 2.4\sigma$.

Biomolecule in solution

typical MD simulations – molecular system in aqueous solution

task – make the system as small as possible (reduce cost)

straightforward solution – single molecule of solute (protein, DNA)

with a smallest possible number of H₂O molecules

typical – several thousand H₂O molecules in a cube $n \times n \times n$ nm

issue – everything is close to the **surface**,

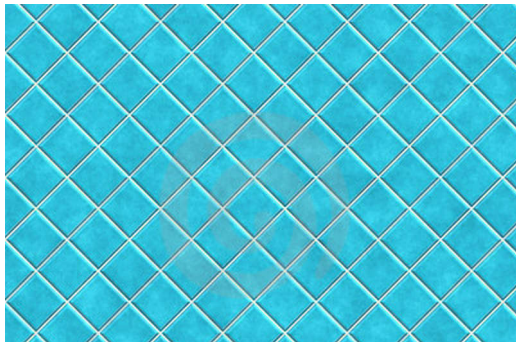
while we are interested in a molecule in **bulk solvent**

so – we may be simulating something else than what we want

Periodic boundary conditions

- elegant way to avoid these problems
- molecular system placed in a regular-shaped **box**
- the box is virtually replicated in all spatial directions

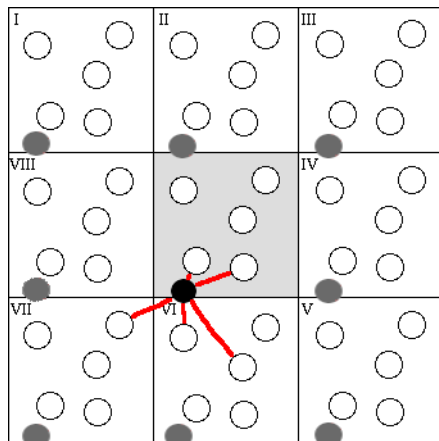
Periodic boundary conditions



Periodic boundary conditions

- elegant way to avoid these problems
- molecular system placed in a regular-shaped **box**
- the box is virtually replicated in all spatial directions
- positions (and velocities) of all particles are **identical** in all replicas, so that we can keep only one copy in the memory
- this way, the system is **infinite** – no surface!
- the atoms near the wall of the **simulation cell** interact with the atoms in the neighboring replica

PBC – example

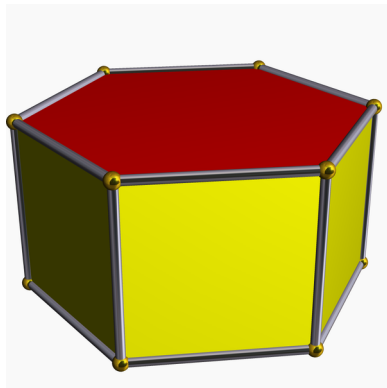
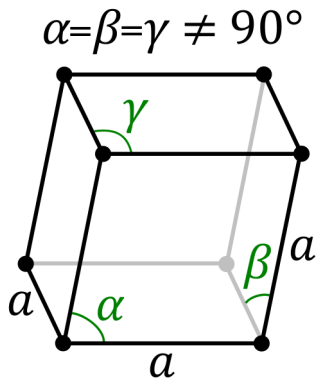


PBC – features

- small problem – artificial periodicity in the system (entropy ☹)
 - still much better than boundary with vacuum
- only coordinates of the unit cell are recorded
- atom that leaves the box enters it on the other side.
- careful accounting of the interactions of atoms necessary!
 - simplest – **minimum image convention**:
 - an atom interacts with the nearest copy of every other
 - interaction with two different images of another atom, or even with another image of itself is avoided

PBC – box shape

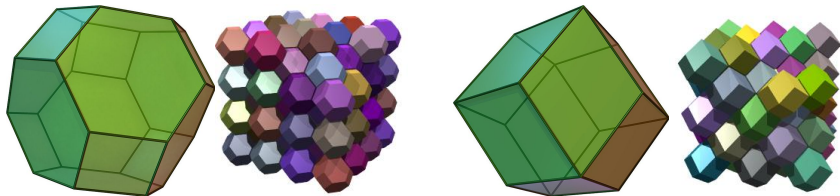
may be simple – cubic or orthorhombic, parallelepiped (specially, rhombohedron), or hexagonal prism



PBC – box shape

... but also more complicated

- truncated octahedral or rhombic dodecahedral
- quite complex equations for interactions & eqns of motion

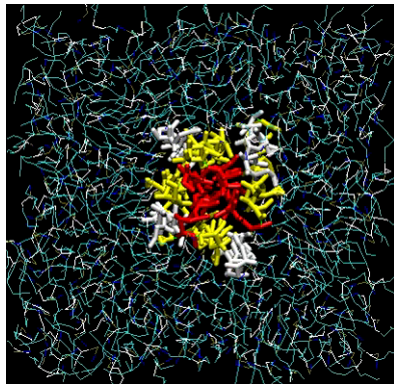
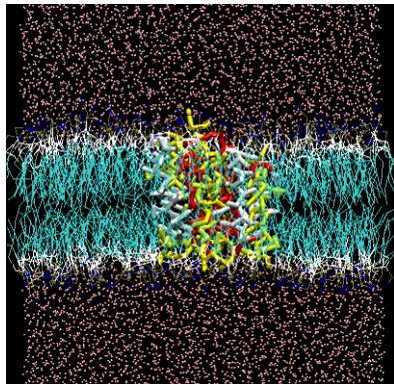


advantage for simulation of spherical objects (globular proteins)

- no corners far from the molecule filled with unnecessary H_2O

PBC – box shape

- 2D objects – phase interfaces, membrane systems
– usually treated in a **slab** geometry



Cut-off – simple idea

non-bonded terms – bottleneck of the calculation
with PBC – infinite number of interaction pairs in principle,
but the interaction gets weaker with distance
simplest and crudest approach to limit the number of calculations
neglect interaction of atoms further apart than r_c – **cut-off**
very good for rapidly decaying LJ interaction ($1/r^6$) ($r_c = 10 \text{ \AA}$)
not so good for slowly decaying electrostatics ($1/r$)
– sudden jump (discontinuity) of potential energy,
disaster for forces at the cut-off distance

Cut-off – better alternatives

shift the whole function by $V(r_c)$ – eliminate the jump at r_c :

$$V^{\text{sh}}(r) = \begin{cases} V(r) - V(r_c), & \text{for } r \leq r_c, \\ 0, & \text{otherwise.} \end{cases}$$

still, the gradients (forces) are at r_c discontinuous!

shift-force potential gets rid of that ($V' \equiv dV/dr$):

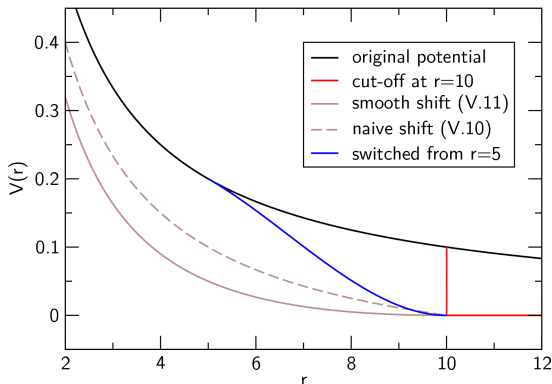
$$V^{\text{sf}}(r) = \begin{cases} V(r) - V(r_c) - V'(r_c) \cdot (r - r_c), & \text{for } r \leq r_c, \\ 0, & \text{otherwise.} \end{cases}$$

drawback – the Coulomb energy is not quite Coulomb anymore

switch potential – starting from a certain distance r_1 , the Coulomb interaction is multiplied by a function going gradually from 1 to 0
– interaction altered in the cut-off region.

Cut-off – better alternatives

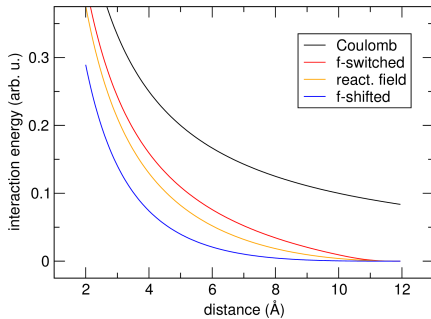
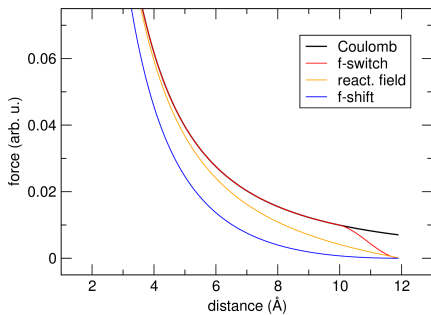
elstat. interaction energy of two unit positive charges



shift / switch – applied to either energy or force

cut-off schemes – based on functional groups, if applicable

Cut-off – better alternatives



Neighbor lists

cut-off – we still have to calculate the distance for **every** two atoms
(to compare it with the cut-off distance)

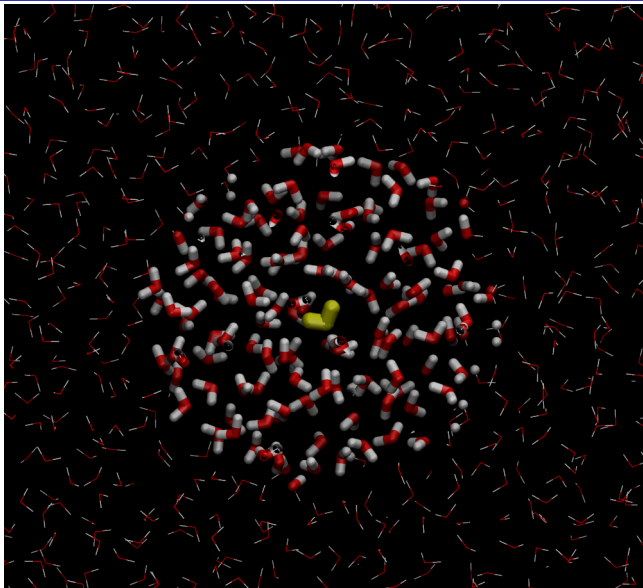
→ we do not win much yet – there are still $\mathcal{O}(N^2)$ distances

observation: pick an atom A.

the atoms that are within cut-off distance r_c around A,
remain within r_c for several consecutive steps of dynamics,
while no other atoms approach A that close

idea: maybe it is only necessary to calculate the interactions
between A and these close atoms – **neighbors**

Neighbor lists



Neighbor lists

what will we do? calculate the distances for every pair of atoms
less frequently, i.e. every 10 or 20 steps of dynamics, and
record the atoms within cut-off distance in a **neighbor list**

| atom | how many? | list of neighboring atoms | | | | | | | | | | | | | | | |
|------|-----------|---------------------------|------|------|------|------|------|------|------|------|------|------|-----|--|--|--|--|
| 1 | 378 | 2191 | 408 | 1114 | 1802 | 262 | 872 | 649 | 805 | 1896 | 2683 | 114 | 189 | | | | |
| 2 | 403 | 1788 | 1624 | 1048 | 1745 | 2546 | 506 | 203 | 288 | 2618 | 1445 | 880 | 133 | | | | |
| 3 | 385 | 779 | 2869 | 800 | 2246 | 1252 | 570 | 454 | 1615 | 1656 | 1912 | 2395 | 152 | | | | |
| 4 | 399 | 367 | 2143 | 1392 | 1448 | 1460 | 1411 | 2921 | 2725 | 429 | 845 | 2601 | 181 | | | | |
| 5 | 406 | 1385 | 425 | 1178 | 2112 | 1689 | 1897 | 1650 | 1747 | 1028 | 1366 | 605 | 176 | | | | |
| 6 | 388 | 1748 | 130 | 2244 | 631 | 1677 | 1748 | 2566 | 303 | 552 | 562 | 1142 | 256 | | | | |
| 7 | 379 | 20 | 15 | 1322 | 196 | 1590 | 655 | 552 | 1401 | 2177 | 411 | 2904 | 236 | | | | |
| 8 | 395 | 888 | 1074 | 786 | 2132 | 1703 | 218 | 1846 | 337 | 1683 | 1917 | 2005 | 94 | | | | |
| 9 | 396 | 2433 | 934 | 1055 | 1518 | 2750 | 2534 | 1697 | 2006 | 769 | 2407 | 1478 | 123 | | | | |
| 10 | 381 | 2461 | 1910 | 459 | 2628 | 2523 | 1709 | 2069 | 1151 | 1710 | 2107 | 1909 | 13 | | | | |
| 11 | 400 | 1029 | 756 | 670 | 1592 | 612 | 676 | 1473 | 2859 | 392 | 986 | 155 | 265 | | | | |

then – calculate the interaction for each atom
only with for the atoms in the neighbor list – formally $\mathcal{O}(N)$
note – the build of the neighbor list itself is $\mathcal{O}(N^2)$,
which can be reduced with further tricks ('cell lists')

Accounting of all of the replicas

cut-off – often bad approximation, e.g. with highly charged molecular systems (DNA, some proteins)

artificial forces with switching function

→ e.g. artificial accumulation of ions around cut-off

only way – abandon the minimum image convention and cut-off
– sum up the long-range Coulomb interaction

between **all** the replicas of the simulation cell

introduce \vec{n} running over all the replicas

- for $|\vec{n}| = 0$, we have $\vec{n} = (0, 0, 0)$ – the central unit cell.
- for $|\vec{n}| = L$: $\vec{n} = (0, 0, \pm L)$, $\vec{n} = (0, \pm L, 0)$, $\vec{n} = (\pm L, 0, 0)$
– the six neighboring unit cells.
- continue with $|\vec{n}| = \sqrt{2}L$: 12 cells touching with edge...

Can we sum it up simply?

sum of Coulomb interactions over all replicas:

$$E^{\text{Coul}} = \frac{1}{2} \sum_{i,j} \sum_{\text{replicas } \vec{n}} \frac{q_i \cdot q_j}{|\vec{r}_{ij} + \vec{n}|}$$

i and j run over all atoms in the unit cell (r_{ij} – their distance)

infinite sum with special convergence problems

alternating harmonic series $\sum_n (-1)^n / n$ – **conditionally convergent**:

it converges $\sum_{i=1}^{\infty} a_i < \infty$, but

does not converge absolutely: $\sum_{i=1}^{\infty} |a_i| = \infty$

– convergence is slow and dependent on the order of summation

BTW: conditionally convergent series

$$\text{I: } S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \dots$$

$$\text{II: } \frac{1}{2}S = +\frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \dots$$

$$\text{I + II: } \frac{3}{2}S = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \dots$$

$$= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots = S \quad (\text{sic!})$$

Sum it up not so simply – Ewald

any smart way to calculate ESP induced by **all** images of **all** atoms?

$$\Phi(\vec{r}_i) = \sum_j \sum_{\text{replicas } |\vec{n}|} \frac{q_j}{|\vec{r}_{ij} + \vec{n}|}$$

to get the Coulomb energy of charges q_i in the unit cell

$$E^{\text{Coul}} = \frac{1}{2} \sum_i q_i \cdot \Phi(\vec{r}_i)$$

idea: pass to a sum of two series that will converge rapidly:

$$\sum \frac{1}{r} = \sum \frac{f(r)}{r} + \sum \frac{1-f(r)}{r}$$

may seem awkward, but will work well ☺

Sum it up not so simply – Ewald

summing over **point** charges – difficult (convergence problem)

Ewald method – **Gaussian densities** of the same magnitude:

$$q_j \rightarrow q_j \cdot \left(\frac{\alpha}{\sqrt{\pi}} \right)^3 \exp \left[-\alpha^2 \cdot |\vec{r}_j|^2 \right]$$

get ESP – with Poisson's equation

for 1 Gaussian – ESP has a form of **error function**

error function: defined as definite integral of a Gaussian

$$\text{erf}[x] = \frac{2}{\sqrt{\pi}} \int_0^x \exp[-t^2] dt$$

and the complementary error function as

$$\text{erfc}[x] = 1 - \text{erf}[x]$$

Sum it up not so simply – Ewald

ESP for a Gaussian charge density:

$$\Phi(\vec{r}) = q_j \cdot \frac{\text{erf}[\alpha \cdot r]}{r}$$

and, in the special case of 'self-interaction' $\vec{r} = \vec{o}$:

$$\Phi(\vec{o}) = q_j \cdot \frac{2\alpha}{\sqrt{\pi}}$$

with Ewald: summing ESP induced by all charges, we obtain

$$\Phi(\vec{r}_i) = \sum_j \sum_{\text{replicas } |\vec{n}|} q_j \cdot \frac{\text{erf}[\alpha \cdot |\vec{r}_{ij} + \vec{n}|]}{|\vec{r}_{ij} + \vec{n}|}$$

do not forget: we have to compare this with the **full** ESP induced by point charges; the difference – $\text{erfc}[]$

Sum it up not so simply – Ewald

the full ESP induced by all replicas of all charges:

$$\begin{aligned}
 \Phi(\vec{r}_i) &= \sum_j \sum_{\text{replicas } |\vec{n}|} q_j \cdot \frac{\text{erfc}[\alpha \cdot |\vec{r}_{ij} + \vec{n}|]}{|\vec{r}_{ij} + \vec{n}|} \\
 &+ \sum_j \sum_{\text{replicas } |\vec{n}|} q_j \cdot \frac{\text{erf}[\alpha \cdot |\vec{r}_{ij} + \vec{n}|]}{|\vec{r}_{ij} + \vec{n}|} \\
 &= \Phi^{\text{real}}(\vec{r}_i) + \Phi^{\text{rec}}(\vec{r}_i)
 \end{aligned}$$

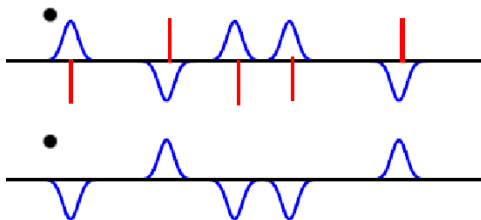
$\Phi^{\text{real}}(\vec{r}_i)$ – **real-space contribution**

- from a certain, quite small distance (depending on α):
point charges and the charge densities cancel each other
- this contribution vanishes and we can use cut-off here

Ewald – two contributions

real-space contribution to the Ewald sum

- original point charges (red) and Gaussian charge densities (blue) of the same magnitude but opposite sign

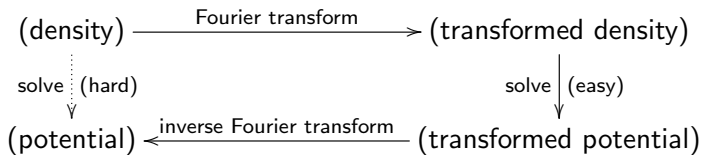


- a distant Gaussian ‘looks much like’ a point charge, and the difference of ESP goes to zero – cut-off is justified

reciprocal-space contribution

Ewald – 2nd contribution

the total charge density is **periodic** → it may be meaningful to Fourier-transform the calculation to the reciprocal space



$$\hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) \cdot \exp[-2\pi i \cdot x \cdot \xi] dx$$

$$\text{Re } \hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) \cdot \cos[x \cdot \xi] dx$$

$$\text{Im } \hat{f}(\xi) = \int_{-\infty}^{\infty} f(x) \cdot \sin[x \cdot \xi] dx \quad (1)$$

Ewald – 2nd contribution

$\Phi^{\text{rec}}(\vec{r}_i)$ – **reciprocal-space contribution**

- with ‘reciprocal’ vector $\vec{k} = \left(k_x \cdot \frac{2\pi}{L_x}, k_y \cdot \frac{2\pi}{L_y}, k_z \cdot \frac{2\pi}{L_z}\right)$, $k_i \in \mathcal{Z}$
- best evaluated in the form

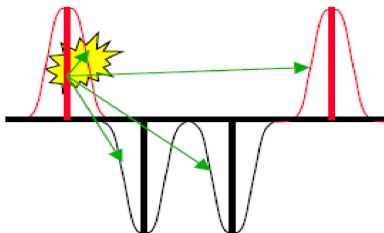
$$\Phi^{\text{rec}}(\vec{r}_i) = \frac{4\pi}{V} \cdot \sum_{\vec{k} \neq \vec{0}} \frac{1}{k^2} \cdot \exp\left[-\frac{|\vec{k}|^2}{4\alpha^2}\right] \cdot \sum_j q_j \cdot \exp[-i \cdot \vec{k} \cdot \vec{r}_{ij}]$$

- terms decrease with increasing $|\vec{k}|$ quickly – cut-off possible
- converges fast with large Gaussian width α
- therefore, the value of α is a **compromise** between the requirements of real- and reciprocal-space calcul.

both contributions – favorable convergence behavior → **we can** evaluate electrostatic interactions with atoms in **all** periodic images

Ewald – the last contribution

broadened charge density interacts with itself
and this energy must be subtracted from the final result



Coulomb self-energy of a broadened Gaussian:

$$E^{\text{self}} = \sum_j q_j \cdot \Phi(\vec{o}) = \sum_j q_j \cdot q_j \cdot \frac{\alpha}{\sqrt{\pi}}$$

Ewald – complete expression for energy

3 contributions:

- ‘real-space’

$$E^{\text{real}} = \frac{1}{2} \sum_j q_j \cdot \Phi^{\text{real}}(\vec{r}_j)$$

- ‘reciprocal-space’

$$E^{\text{rec}} = \frac{1}{2} \sum_j q_j \cdot \Phi^{\text{rec}}(\vec{r}_j)$$

- ‘self-energy’

$$E^{\text{Ewald}} = E^{\text{real}} + E^{\text{rec}} - E^{\text{self}}$$

Ewald – optional additional contribution

Surface / dipole term

- for periodic cells with zero charge and non-zero dipole moment $\vec{\mu}_{\text{tot}}$

$$\Phi^{\text{sur}}(\vec{r}_i) = \frac{4\pi}{3} \frac{\vec{r}_i \cdot \vec{\mu}_{\text{tot}}}{V}$$

- universal application may lead to problems when mobile ions cross the box boundaries (abrupt changes of $\vec{\mu}_{\text{tot}}$)
- this describes the situation with surrounding vacuum
- if this is not present
 - ‘tin-foil’ boundary conditions – surrounding $\epsilon = \infty$

Thinking about Ewald

Ewald summation – correct Coulomb interaction energy
at a quite high computational cost:

- scales with the number of atoms as $\mathcal{O}(N^2)$
- with a better algorithm – $\mathcal{O}(N^{\frac{3}{2}})$
- not efficient enough for large-scale simulations
- goal – improved efficiency of the long-range sum
(reciprocal-space contribution)

Thinking about Ewald

Ewald summation – correct Coulomb interaction energy at a quite high computational cost:

- scales with the number of atoms as $\mathcal{O}(N^2)$
- with a better algorithm – $\mathcal{O}(N^{\frac{3}{2}})$
- not efficient enough for large-scale simulations
- goal – improved efficiency of the long-range sum (reciprocal-space contribution)

particle–mesh Ewald method (1993)

- combines ideas from crystallography (Ewald method) and plasma physics (particle–mesh method)
- key to success – 3D fast Fourier transform technique

Long-range energy with PME

PME works with a regular **grid** constructed in the simulation box

step 1

convert the point charges to Gaussian charge densities and
spread on the grid in the form of splines

practically, we need to have charges discretized on the grid points
if an atom is close to the edge of the box, a part of its charge
must be put to the opposite side of the box (PBC)

Long-range energy with PME

step 2

Fourier transform the charge density on the grid

- discrete 3D fast Fourier transform technique

solve Poisson's eqn in the **reciprocal space**

→ energy and Fourier transform of potential

$$E^{\text{rec}} = \frac{1}{2} \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} \sum_{k_3=0}^{K_3-1} Q(k_1, k_2, k_3) \cdot (\Theta^{\text{rec}} \star Q)(k_1, k_2, k_3)$$

3D-FFT used to calculate the convolution $\Theta^{\text{rec}} \star Q$

- this corresponds to the ESP in reciprocal space

Θ^{rec} depends on box size and character of splines

Long-range energy with PME

step 3

get the potential in real space (inverse Fourier transform),
interpolate its derivative to calculate the **forces**
– expressed in terms of splines – analytical calculation

step 4

get E^{real} and E^{self} – directly from the presented expressions

step 5

attention – the reciprocal energy/forces include contributions from
atom pairs that are connected with bonds
– these have to be subtracted afterwards (**excluded**)

$$E_{\text{excl}}^{\text{rec}} = - \sum_{i,j}^{\text{list}} \frac{1}{4\pi\epsilon_0} \frac{q_i \cdot q_j}{r_{ij}} \text{erf} [\alpha \cdot r_{ij}]$$

Long-range energy with PME

PME parameters: spacing of grid ca. 1 Å,
 α^{-1} ca. 2.5 Å \rightarrow short-range cutoff ≤ 10 Å possible

neighbor lists – every n_l steps, the following procedure:

- measure distances of every pair of atoms in system
- distance smaller than a pre-set value? put the pair on the list

in the following n_l steps, calculate short-ranged interactions
only for the pairs on this neighbor list

\rightarrow linear scaling of short-range interactions ($\mathcal{O}(N)$)

complexity of the long-range PME component:

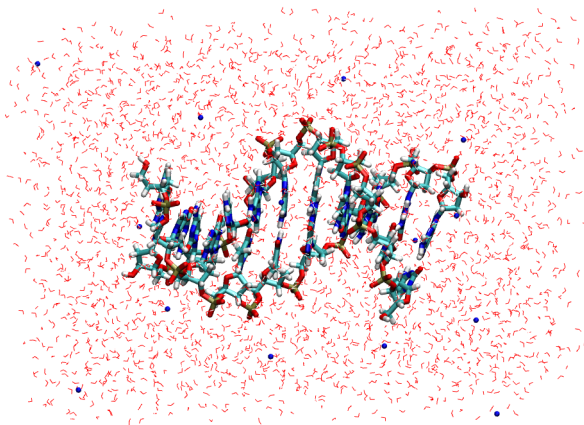
$\mathcal{O}(N \cdot \log N)$ due to the efficiency of FFT

modern implementations – nearly as efficient as cut-off!

Water in biomolecular simulations

most simulations – something in aqueous solutions

H₂O – usually (many) thousands molecules



Water in biomolecular simulations

most simulations – something in aqueous solutions

H₂O – usually (many) thousand molecules

example – simulation of DNA decanucleotide:

- PBC box $3.9 \times 4.1 \times 5.6$ nm (smallest meaningful)
- 630 atoms in DNA, 8346 atoms in water and 18 Na⁺
- concentration of DNA: 18 mmol/L – very high!
- of all **pair interactions**: 86 % are water–water,
most of the others involve water

Water models

most interactions involve H_2O

→ necessary to pay attention to its description

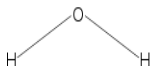
model of water must be simple enough (computational cost)
and accurate enough, at the same time

water models – usually **rigid**

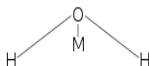
– bond lengths and angles do not vary – **constraints**

molecule with three sites (atoms in this case), or up to six sites

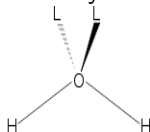
– three atoms and virtual sites corresponding
to a 'center' of electron density or lone electron pairs



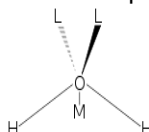
3-site



4-site



5-site



6-site

Water models

TIP3P (very similar is SPC)

- most frequently used
- 3 atoms with 3 rigid bonds, charge on every atom ($-0.834/+0.417$)
- **only** the O possesses non-zero LJ parameters (optimization)

TIP4P

- negative charge placed on virtual site M rather than on the O
- electric field around the molecule described better

TIP5P

- 2 virtual sites L with negative charges near the O – lone pairs
- better description of directionality of H-bonding etc.
(radial distribution function, temperature of highest density)

Water models

Calculated physical properties of the water models

| Model | Dipole moment ^e | Dielectric constant | self-diffusion, 10 ⁻⁵ cm ² /s | Average configurational energy, kJ mol ⁻¹ | Density maximum, °C | Expansion coefficient, 10 ⁻⁴ °C ⁻¹ |
|--------------|----------------------------|---------------------|---|--|---------------------|--|
| SSD | 2.35 [511] | 72 [511] | 2.13 [511] | -40.2 [511] | -13 [511] | - |
| SPC | 2.27 [181] | 65 [185] | 3.85 [182] | -41.0 [185] | -45 [983] | 7.3 [704] ** |
| SPC/E | 2.35 [3] | 71 [3] | 2.49 [182] | -41.5 [3] | -38 [183] | 5.14 [994] |
| SPC/Fw | 2.39 [994] | 79.63 [994] | 2.32 [994] | - | - | 4.98 [994] |
| PPC | 2.52 [3] | 77 [3] | 2.6 [3] | -43.2 [3] | +4 [184] | - |
| TIP3P | 2.35 [180] | 82 [3] | 5.19 [182] | -41.1 [180] | -91 [983] | 9.2 [180] |
| TIP3P/Fw | 2.57 [994] | 193 [994] | 3.53 [994] | - | - | 7.81 [994] |
| iAMOEBA | 2.78 [2031] | 80.7 [2031] | 2.54 [2031] | - | 4 [2031] | 2.5 [2031] |
| QCT ** | 1.85 [1251] | - | 1.5 [1251] | -42.7 [1251] | +10 [1251] | 3.5 [1251] |
| TIP4P | 2.18 [3,180] | 53 ^a [3] | 3.29 [182] | -41.8 [180] | -25 [180] | 4.4 [180] |
| TIP4P-Ew | 2.32 [649] | 62.9 [649] | 2.4 [649] | -46.5 [649] | +1 [649] | 3.1 [649] |
| TIP4P-FQ | 2.64 [197] | 79 [197] | 1.93 [197] | -41.4 [201] | +7 [197] | - |
| TIP4P/2005 | 2.305 [984] | 60 [984] | 2.08 [984] | - | +5 [984] | 2.8 [984] |
| TIP4P/2005f | 2.319 [1765] | 55.3 [1765] | 1.93 [1765] | - | +7 [1765] | - |
| OPC | 2.48 [2168] | 78.4 [2168] | 2.3 [2168] | - | -1 [2168] | 2.7 [2168] |
| SWFLEX-AI | 2.69 [201] | 116 [201] | 3.66 [201] | -41.7 [201] | - | - |
| COS/G3 ** | 2.57 [704] | 88 [704] | 2.6 [704] | -41.1 [704] | -78 [1939] | 7.0 [704] |
| COS/D2 | 2.55 [1617] | 78.9 [1617] | 2.2 [1617] | -41.8 [1617] | - | 4.9 [1617] |
| GCPM | 2.723 [859] | 84.3 [859] | 2.26 [859] | -44.8 [859] | -13 [859] | - |
| SWM4-NDP | 2.461 [933] | 79 [933] | 2.33 [933] | -41.5 [933] | <-53 [1999] | - |
| BK3 | 2.644 [2080] | 79 [2080] | 2.28 [2080] | -43.32 [2080] | +4 [2080] | 3.01 [2080] |
| SWM6 | 2.431 [1999] | 78.1 [1999] | 2.14 [1999] | -41.5 [1999] | -48 [1999] | - |
| TIP5P | 2.29 [180] | 81.5 [180] | 2.62 [182] | -41.3 [180] | +4 [180] | 6.3 [180] |
| TIP5P-Ew | 2.29 [619] | 92 [619] | 2.8 [619] | - | +8 [619] | 4.9 [619] |
| TTM2-F | 2.67 [1027] | 67.2 [1027] | 1.4 [1027] | -45.1 [1027] | - | - |
| POL5/TZ | 2.712 [256] | 98 [256] | 1.81 [256] | -41.5 [256] | +25 [256] | - |
| Six-site * | 1.89 [491] | 33 [491] | - | - | +14 [491] | 2.4 [491] |
| Experimental | 2.95 | 78.4 | 2.30 | -41.5 [180] | +3.984 | 2.53 |

All the data is at 25 °C and 1 atm, except * at 20 °C and ** at 27 °C.