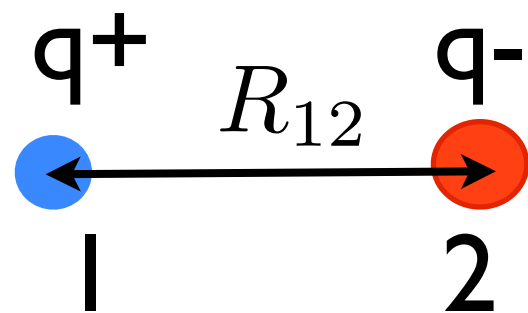


Lecture 3 & 4

Non bonding Interactions

Electrostatic Interactions



$$U(R_{12}) = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{R_{12}}$$

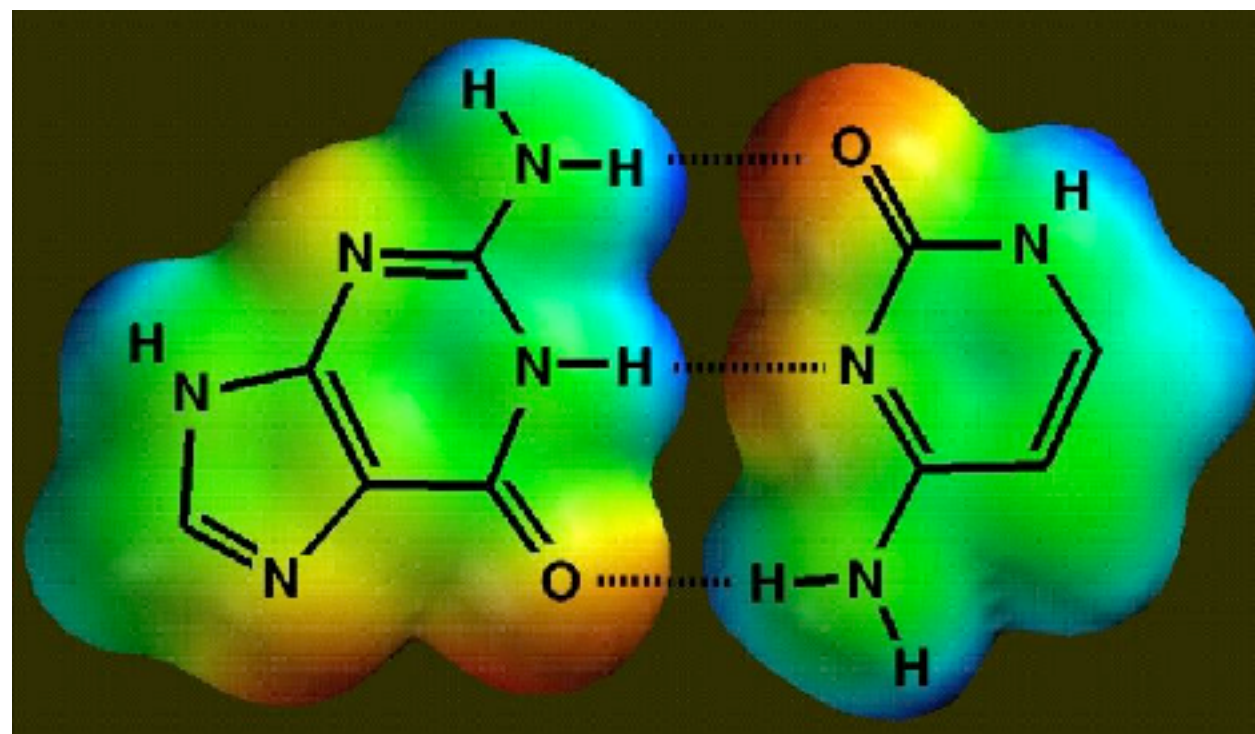
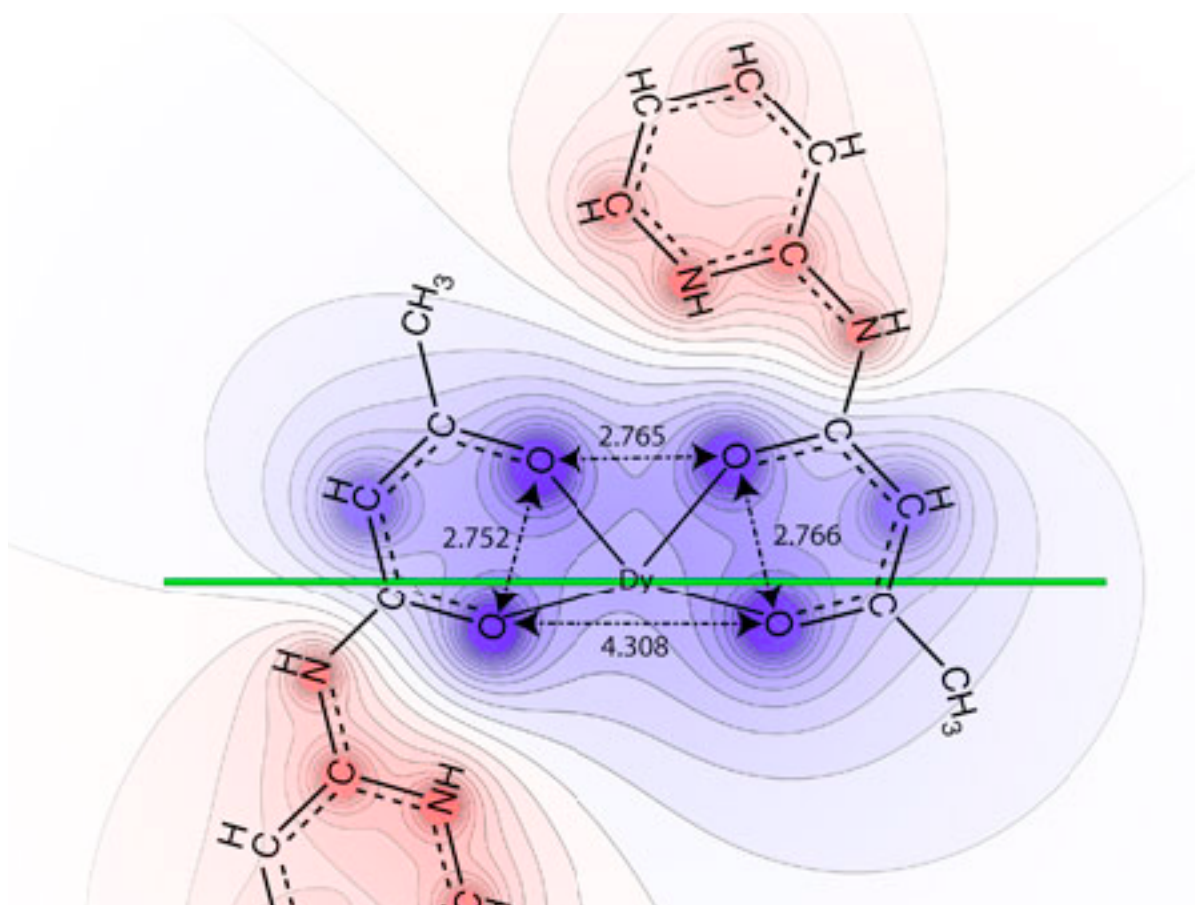
● Point charges: due to electronic redistribution when a molecule is formed

● Charges are only conceptual

● Charges are not observables

● Electronic density is an observable

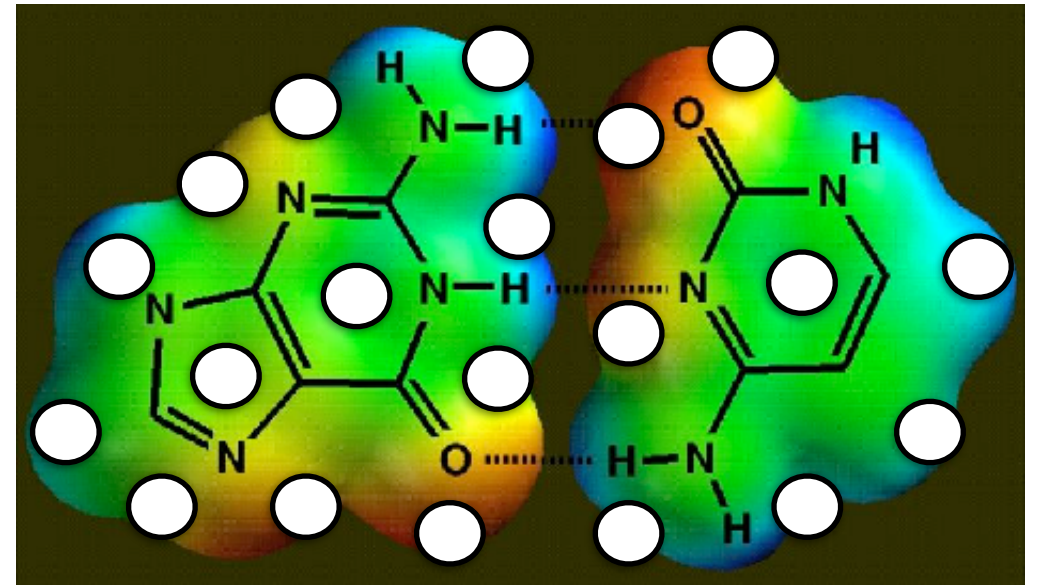
Electrostatic Potential Derived Point Charges (ESP Charges)



Can we get point charges on atoms (i.e. atomic positions) such that we can reproduce the electrostatic potential?

ESP charges: Merz-Singh Kollman Scheme

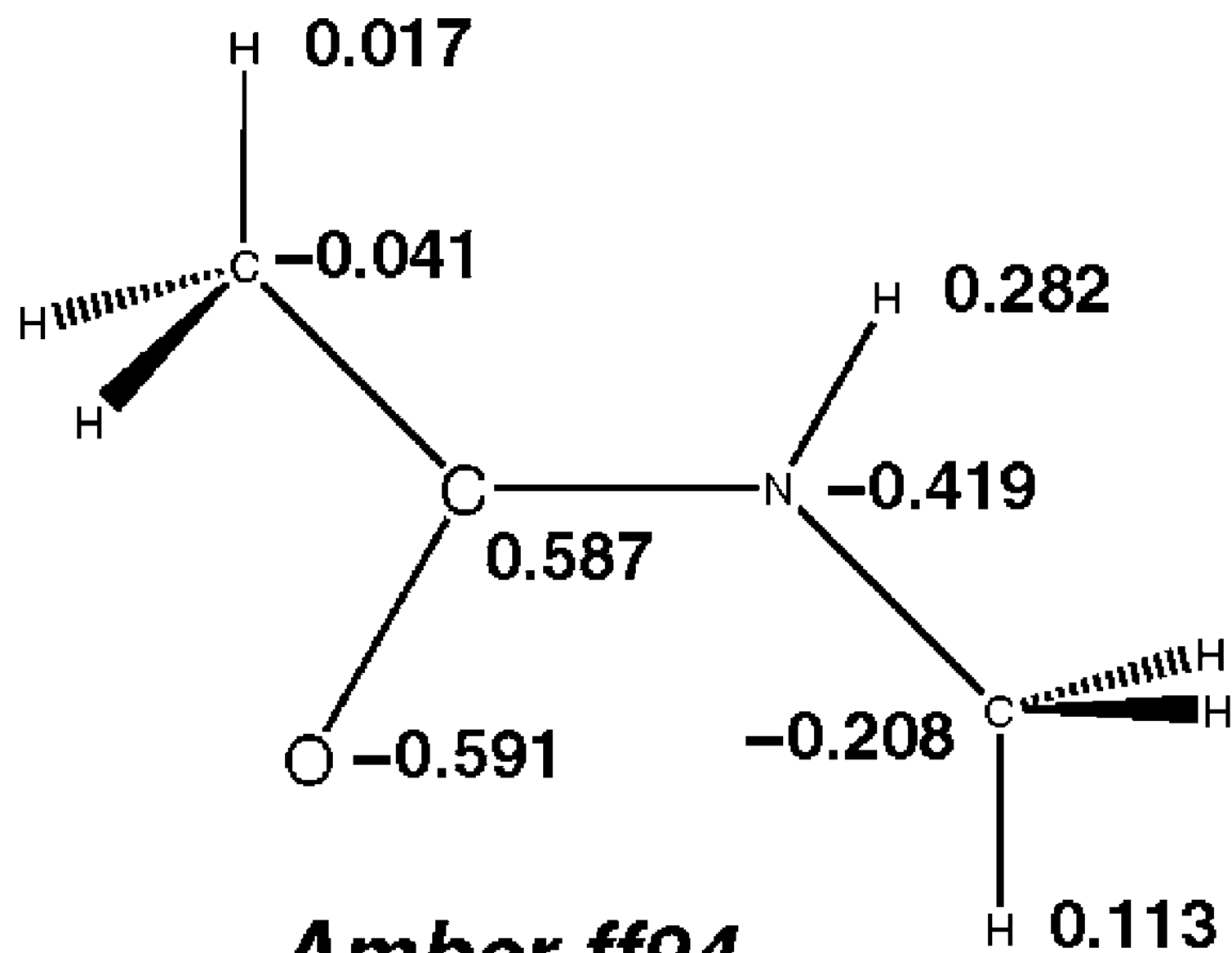
- Comp. Chem., 5 (1984) 129-45
J. Phys. Chem., 1992, 96, 10276
- Point charges are such that the electrostatic potential obtained from a quantum mechanical calculation is mimicked for the selected points
- Fitting procedure is used here



from quantum
mechanics

$$R = \sum_i^{N_{\text{points}}} (\phi_i^0 - \phi_i^{\text{calc}})^2$$

$$\sum_l q_l / |R_l - r|$$



Using Gaussian for Computing ESP Charges

Ref: http://archive.ambermd.org/200812/att-0116/resp_fit.htm

```
%Chk=esp.chk
```

```
# HF/6-31G* SCF=Tight Pop=MK IOP(6/33=2,6/41=10,6/42=17)
```

```
<blank-line>
```

```
ESP Charge Calculation
```

```
<blank-line>
```

```
O 1
```

```
H
```

```
O 1 dOH
```

```
H 2 dOH 1 tHOH
```

```
<blank-line>
```

```
dOH=0.95
```

```
tHOH=104.5
```

```
<blank-line>
```

If you want to restraint the charges: Restrained-ESP or RESP

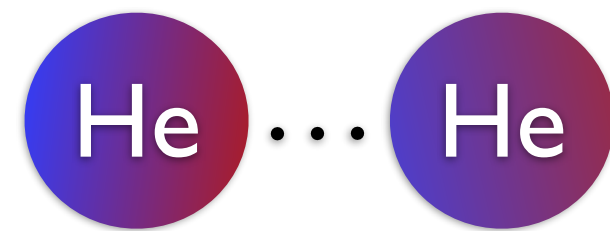
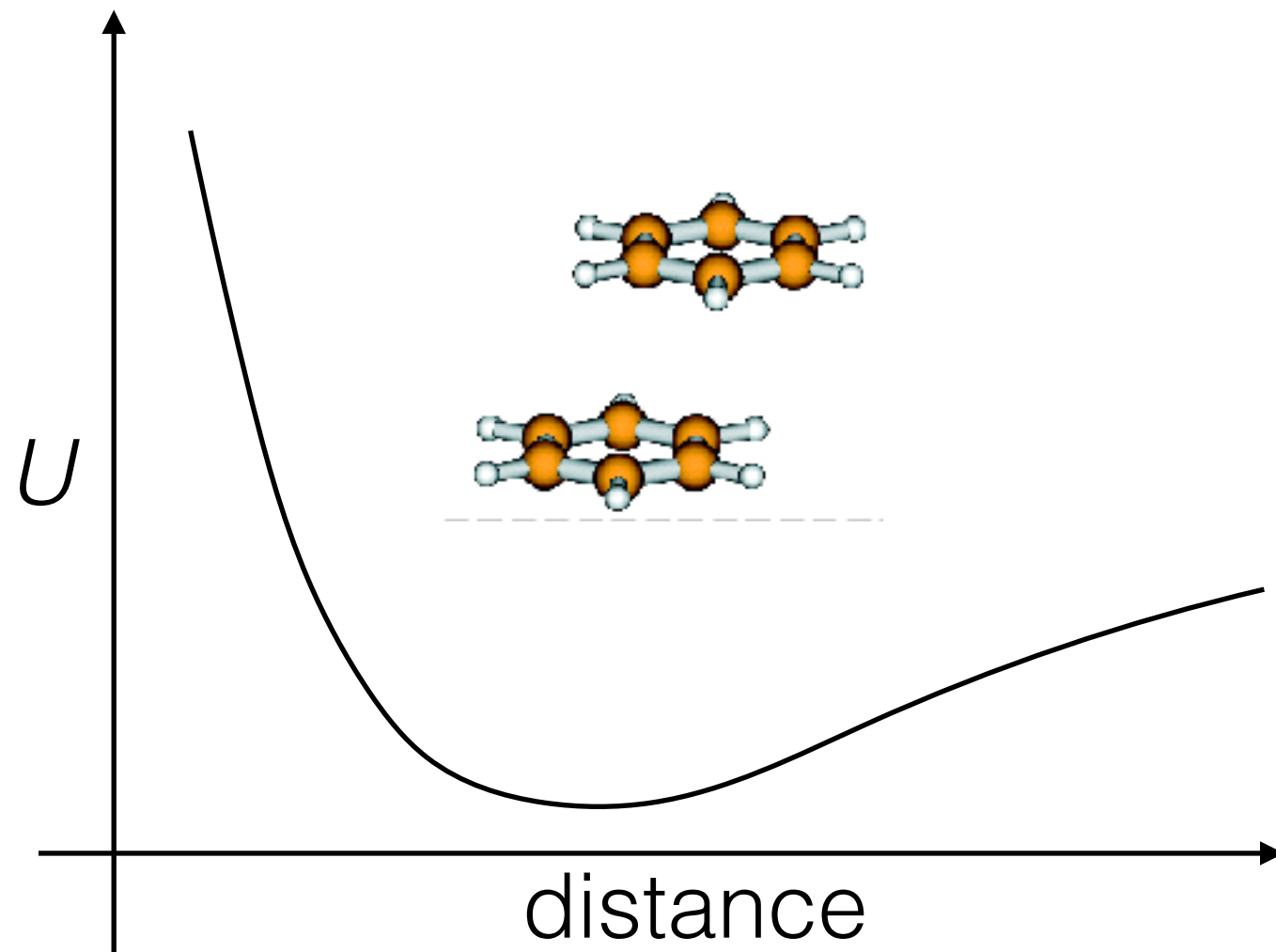
- Run **changept** to create inputs file for resp program from Gaussian output
- Modify resp.in

```
Water
&cntrl  ihfree=0, qwt=0.0005, iqopt=2
&end
1.0
Mol1
0      3
8      0
1      0
1      2
```

use same charge as atom no. 2

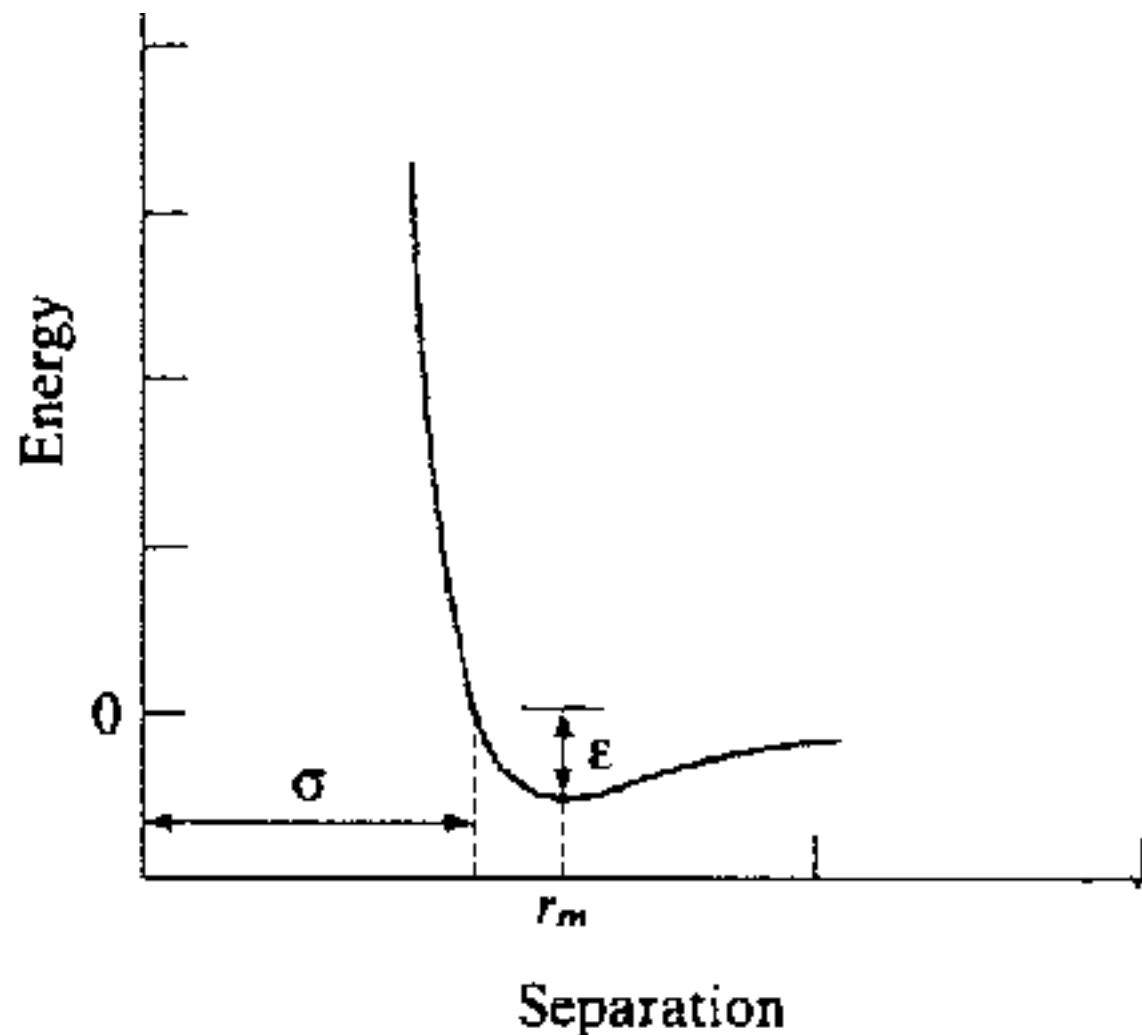
- `resp -O -i resp.in -o resp.out -q qin -e resp.pot`

Dispersive Interactions



- such interactions are due to instantaneous dipole-dipole interactions (dynamic electronic correlation)

- Can be generally mimicked by Lennard-Jones potential



$$U(R) = 4\epsilon \left[\left(\frac{\sigma}{R} \right)^{12} - \left(\frac{\sigma}{R} \right)^6 \right]$$

attractive

repulsive

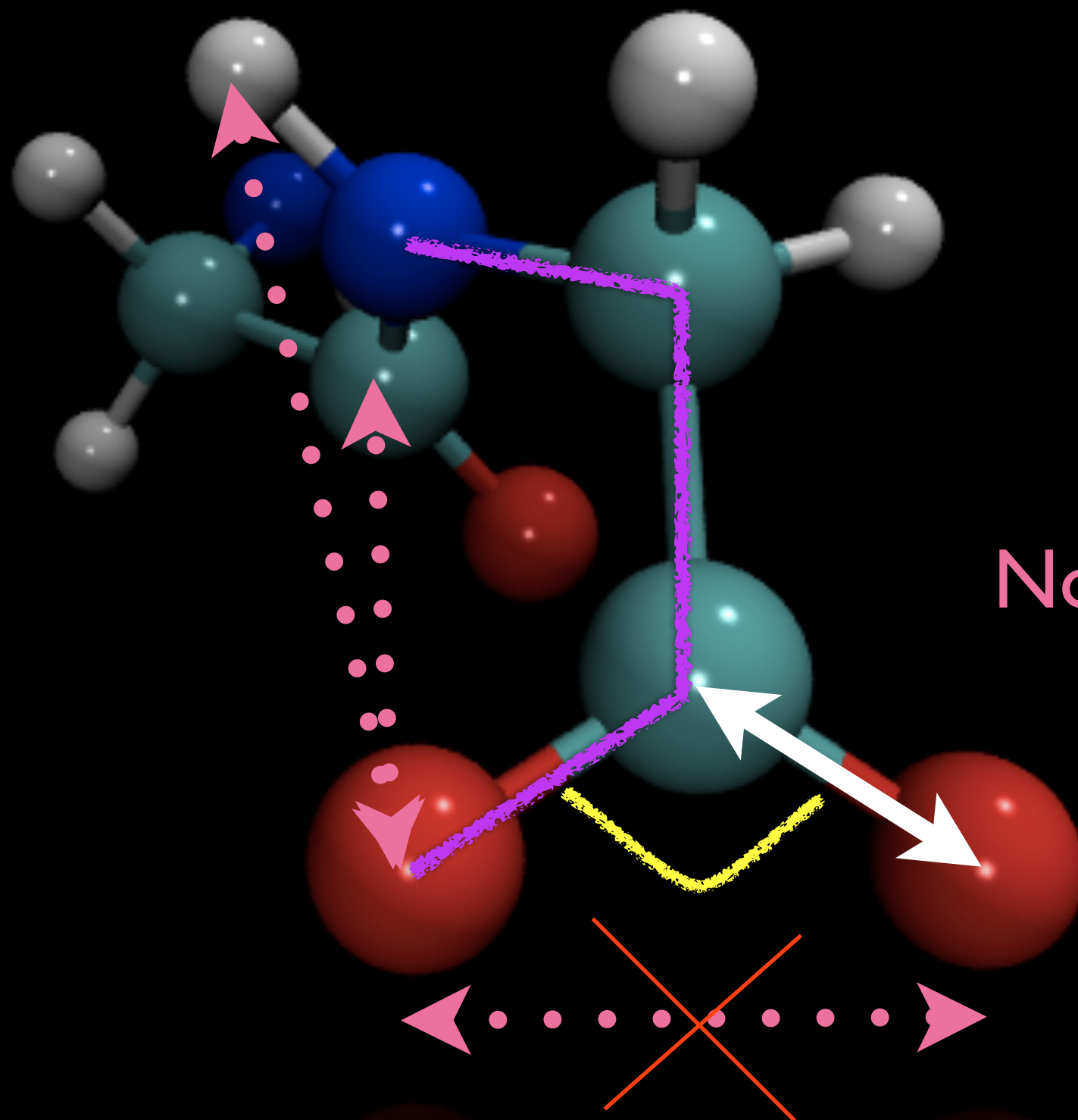
$$\epsilon_{AB} = \sqrt{\epsilon_A \epsilon_B}$$

$$\sigma_{AB} = \frac{\sigma_A + \sigma_B}{2}$$

Molecular Mechanics

Force-Field: Summary

- Bonding terms till fourth bonded neighbour
- “bond” [2-body], “angle” [3-body], “torsion” [4-body] are the bonding terms
- A pair of non-bonding atoms interact by both electrostatic and dispersion interactions



Bond
Angle
Torsion
Non-bonded

Overall potential form:

$$U_{\text{MM}}(\mathbf{R}^N) = \sum_{\text{bonds}} \frac{1}{2} k (x - x_0)^2 + \sum_{\text{angles}} \frac{1}{2} k (\theta - \theta_0)^2 +$$
$$\sum_{\text{torsions}} \frac{V_n}{2} (1 + \cos n\omega) +$$
$$\sum_{IJ} \left\{ 4\epsilon_{IJ} \left[\left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^{12} - \left(\frac{\sigma_{ij}}{R_{IJ}} \right)^6 \right] + \frac{q_I q_J}{R_{IJ}} \right\}$$

non-bonded pairs

- Several parameters are present in the molecular mechanics force-field terms
- These parameters are obtained from experimental observables or by fitting the data to experimental data
- A database with molecular mechanics force-field parameters is called a “force-field library”
- E.g. AMBER (<http://ambermd.org>), GROMACS (<http://www.gromacs.org/>) has well tested “force-field libraries” for proteins, nucleic acids, simple solvents and organic molecules

First empirical force-field

THE JOURNAL OF CHEMICAL PHYSICS VOLUME 14, NUMBER 12 DECEMBER, 1946

The Theory of the Racemization of Optically Active Derivatives of Diphenyl

F. H. WESTHEIMER AND JOSEPH E. MAYER

Department of Chemistry, University of Chicago, Chicago, Illinois

(Received August 2, 1946)

T. L. Hill has recently shown that it is possible to compute the magnitude of steric strain in certain organic molecules. The present paper summarizes the results of a similar but independent investigation directed towards a computation of the rate of racemization of optically active (i.e., sterically hindered) derivatives of diphenyl. The energy of the planar form of a sterically hindered diphenyl can be approximated by the equation:

$$E = \sum_i \frac{1}{2} a_i q_i^2 + A_1 \exp(-d_1/\rho_1) + A_2 \exp(-d_2/\rho_2),$$

where the q_i are the normal coordinates of the unstrained molecule in question and the exponen-

http://ocikbws.uzh.ch/education/molecular_science/mm/

Major breakthrough...

THE JOURNAL OF CHEMICAL PHYSICS

VOLUME 49, NUMBER 11

1 DECEMBER 1968

Consistent Force Field for Calculations of Conformations, Vibrational Spectra, and Enthalpies of Cycloalkane and *n*-Alkane Molecules

S. LIFSON AND A. WARSHEL

Department of Chemical Physics, Weizmann Institute of Science, Rehovot, Israel

(Received 13 May 1968)

An inductive method for a systematic selection of energy functions of interatomic interactions in large families of molecules is suggested and is applied to the family of cycloalkane and *n*-alkane molecules. Equilibrium conformations, vibrational frequencies, and excess enthalpies, including strain energies and vibrational enthalpies, are all derived from the same set of energy functions. The energy-function parameters are optimized by a least-squares algorithm to give the best possible agreement with a large amount and variety of observed data. Analytical derivatives of the various calculated quantities with respect to the energy parameters help to facilitate the computational procedures. The resulting agreement with experiment is used as a measure of success of the energy functions with optimized parameters, referred to as "consistent force field" (CFF). Different CFF's are compared and selected according to their relative success.

- * molecules as balls & springs
- * consistent/transferable parameters
- * Cartesian representation of molecules

Refinement of Protein Conformations using a Macromolecular Energy Minimization Procedure

MICHAEL LEVITT AND SHNEIOR LIFSON

*Weizmann Institute of Science
Rehovot, Israel*

(Received 3 December 1968, and in revised form 29 July, 1969)

This paper presents a rapid refinement procedure capable of deriving the stable conformation of a macromolecule from experimental model co-ordinates. All the degrees of freedom of the molecule are allowed to vary and all parts of the structure are refined simultaneously in a general force-field.

The procedure has been applied to myoglobin and lysozyme. The deviations of peptide bonds from planar conformation and of various bond angles from their respective average values are found to contribute significantly to the refined protein conformation. Hydrogen atoms are not included in the present refinement.

A set of non-bonded potential functions, applicable to the equilibrium of a folded protein in an aqueous medium, is described and tested on myoglobin.

Computer simulation of protein folding

Michael Levitt* & Arieh Warshel*

Department of Chemical Physics, Weizmann Institute of Science, Rehovoth, Israel

A new and very simple representation of protein conformations has been used together with energy minimisation and thermalisation to simulate protein folding. Under certain conditions, the method succeeds in 'renaturing' bovine pancreatic trypsin inhibitor from an open-chain conformation into a folded conformation close to that of the native molecule.

protein, in this case myoglobin, w
of cylinders supposed to represent a
not implemented on a computer an
generally to other proteins not built

Here we tackle the problem diffe
the representation of a protein by
details. This is done both to make th
efficient and also to avoid having to

Nature, 1975

Dynamics of folded proteins

J. Andrew McCammon, Bruce R. Gelin & Martin Karplus

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

Abstract

The dynamics of a folded globular protein (bovine pancreatic trypsin inhibitor) have been studied by solving the equations of motion for the atoms with an empirical potential energy function. The results provide the magnitude, correlations and decay of fluctuations about the average structure. These suggest that the protein interior is fluid-like in that the local atom motions have a diffusional character.

Nature, 1977

Remarks:

- Hydrogen bonds (mainly dipole-dipole effects) is described by Coulombic interactions
- Hydrophobic interactions are accounted by dispersive interactions

Cutoffs

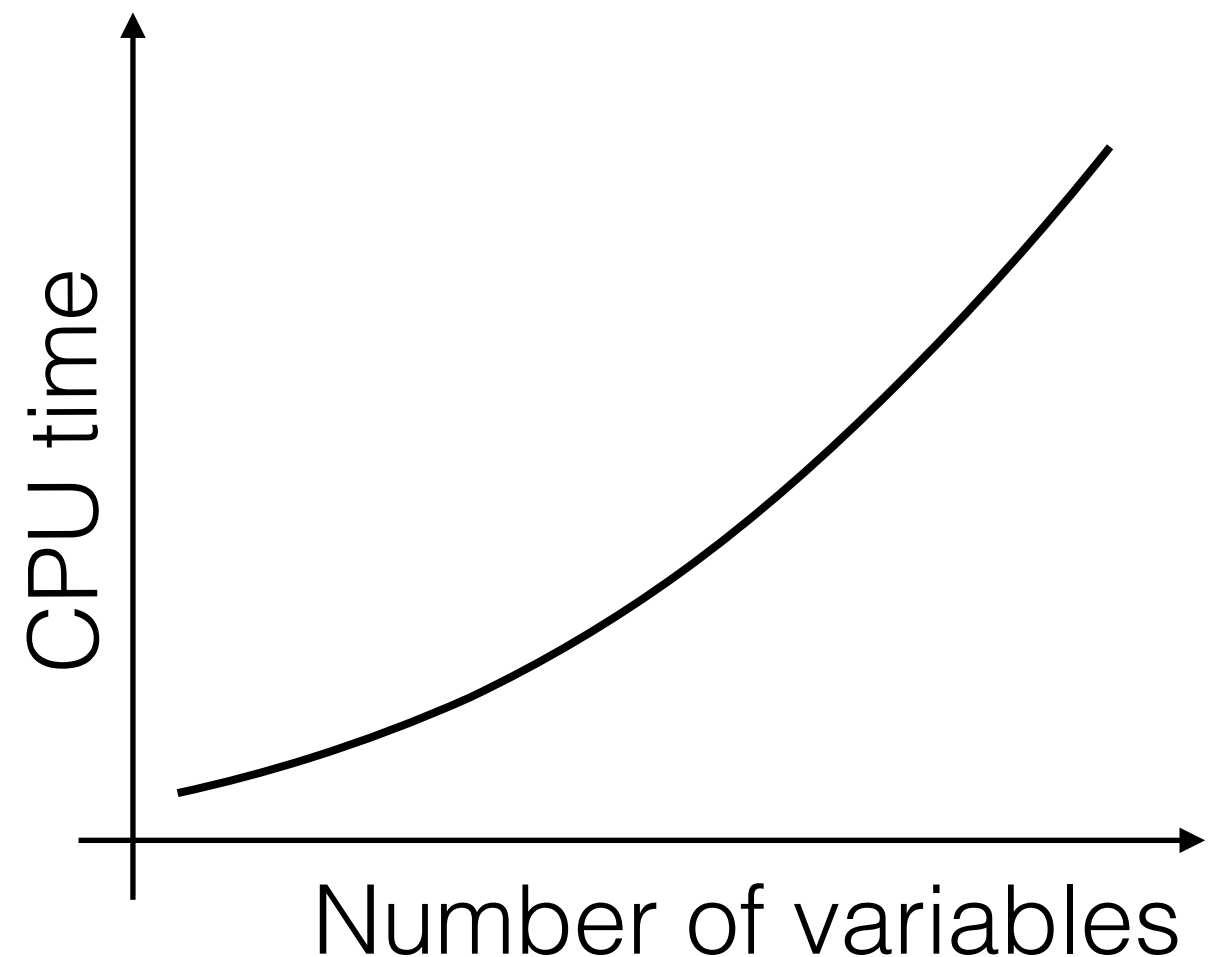
$$U_{\text{MM}}(\mathbf{R}^N) = \sum \frac{1}{2} k (x - x_0)^2 + \sum \frac{1}{2} k (\theta - \theta_0)^2 + \sum \frac{V_n}{2} (1 + \cos n\omega) +$$

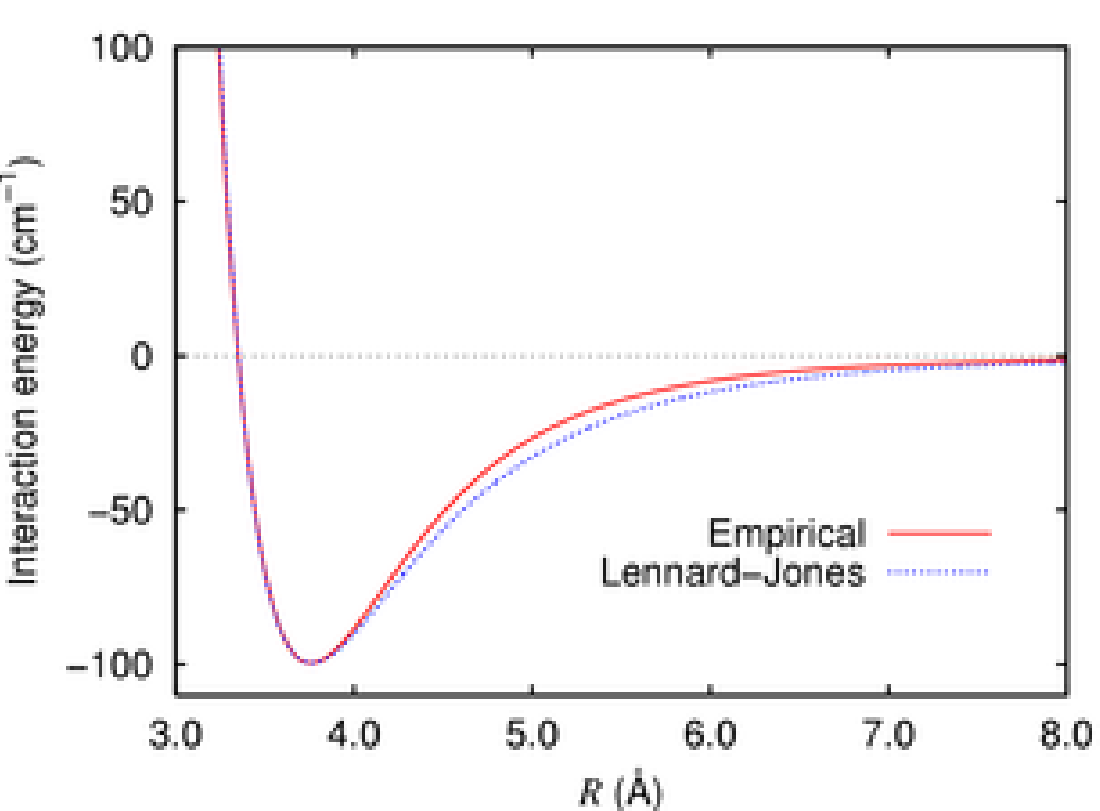
$$\sum_{IJ} \left\{ 4\epsilon_{IJ} \left[\left(\frac{\sigma_{IJ}}{R_{IJ}} \right)^{12} - \left(\frac{\sigma_{ij}}{R_{IJ}} \right)^6 \right] + \frac{q_I q_J}{R_{IJ}} \right\}$$

pair-wise interactions

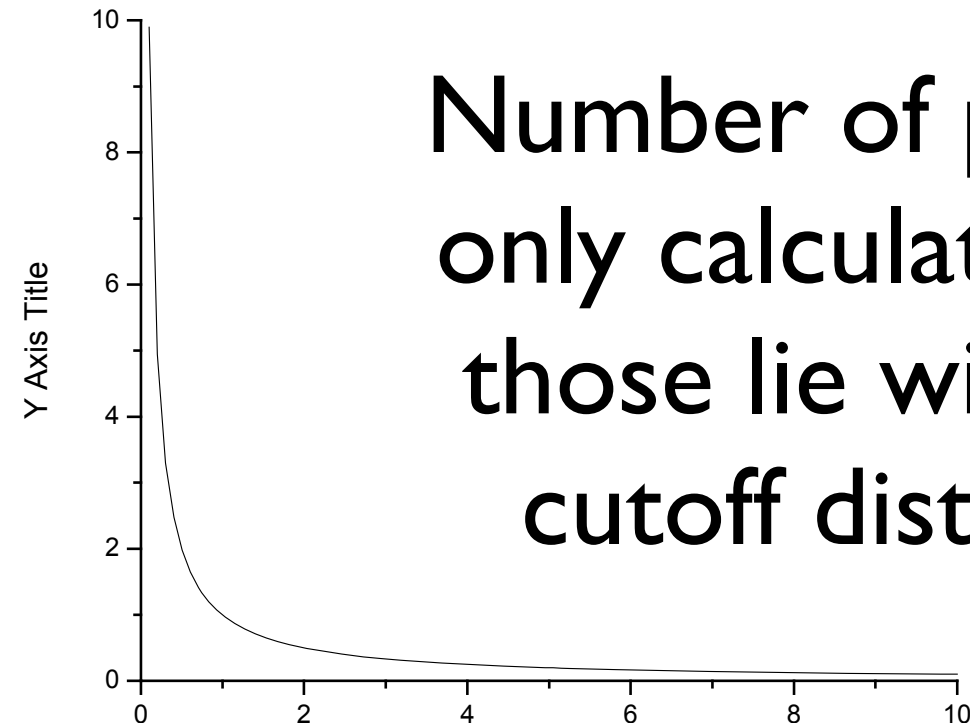
Computational time of
nonbonded terms scale
as $\mathcal{O}(N^2)$

Computational time of
bonded terms scale as
 $\mathcal{O}(N)$



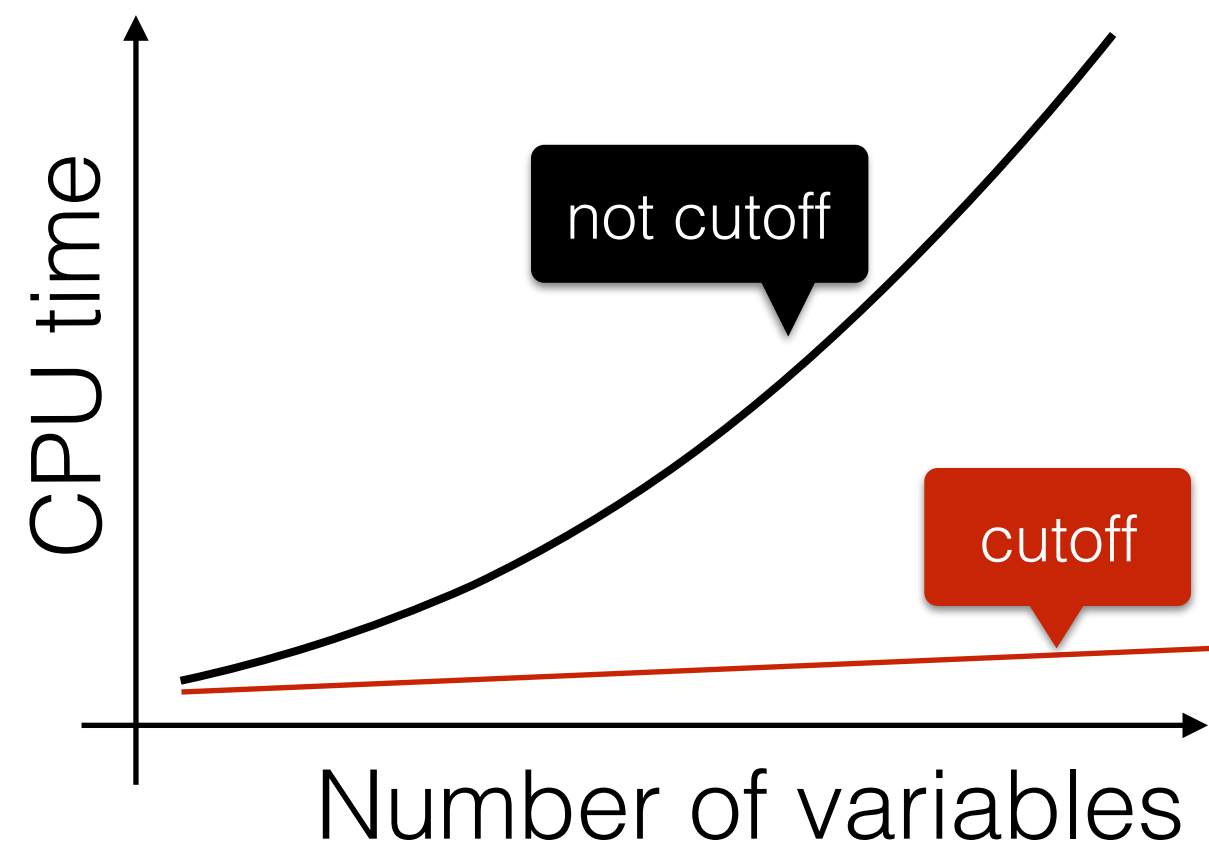
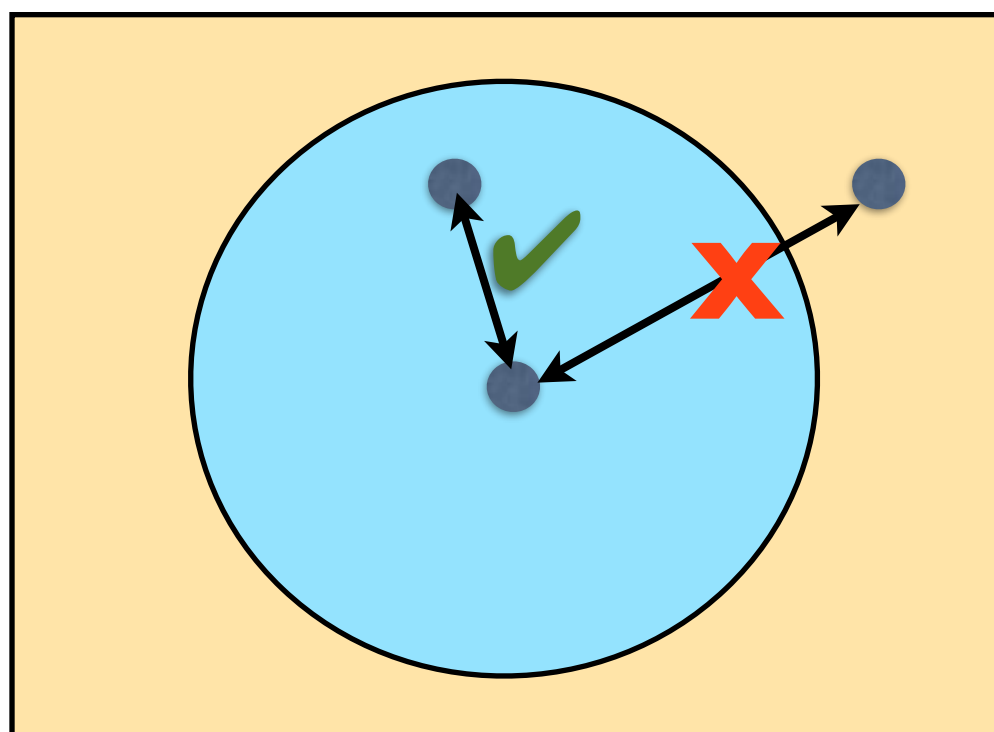


Lennard-Jones



Number of pairs is
only calculated for
those lie within a
cutoff distance

Coulombic



convince yourself

\$ gnuplot

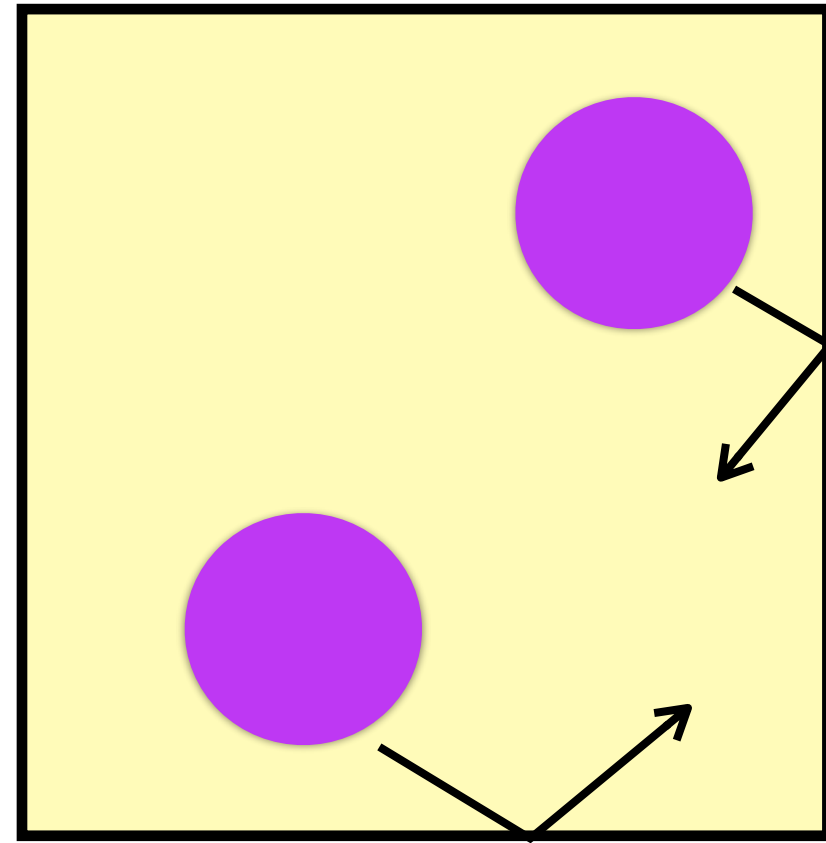
> plot [0:15] 0.3*0.2/x

> plot [0:15] 0.001 * ((1/x)**12 - (1/x)**6)

Bulk vs Finite System

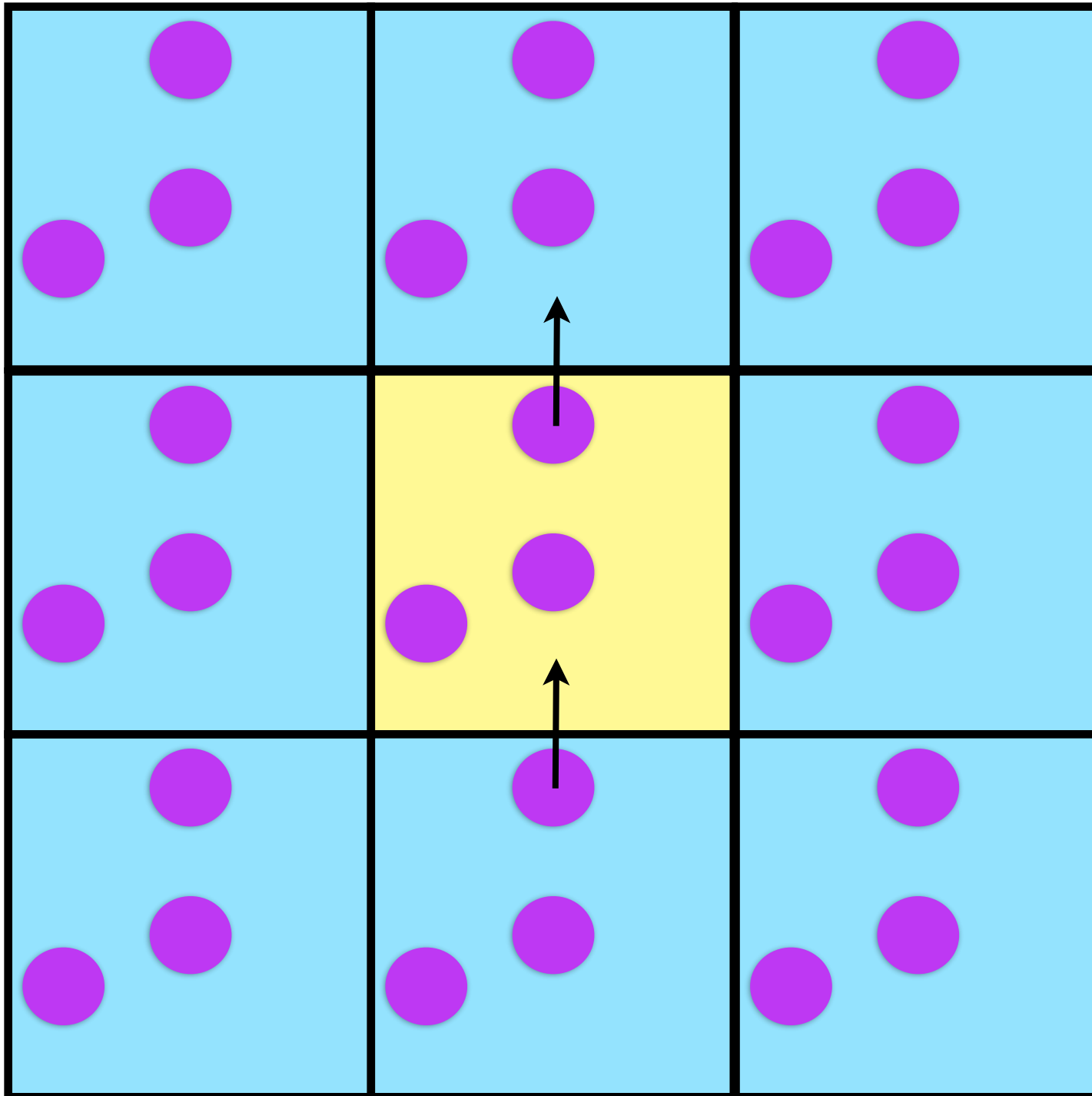
$\sim 10^{23}$ atoms cannot
be treated
computationally

$\sim 10 - 10^6$ atoms can
be usually treated
computationally



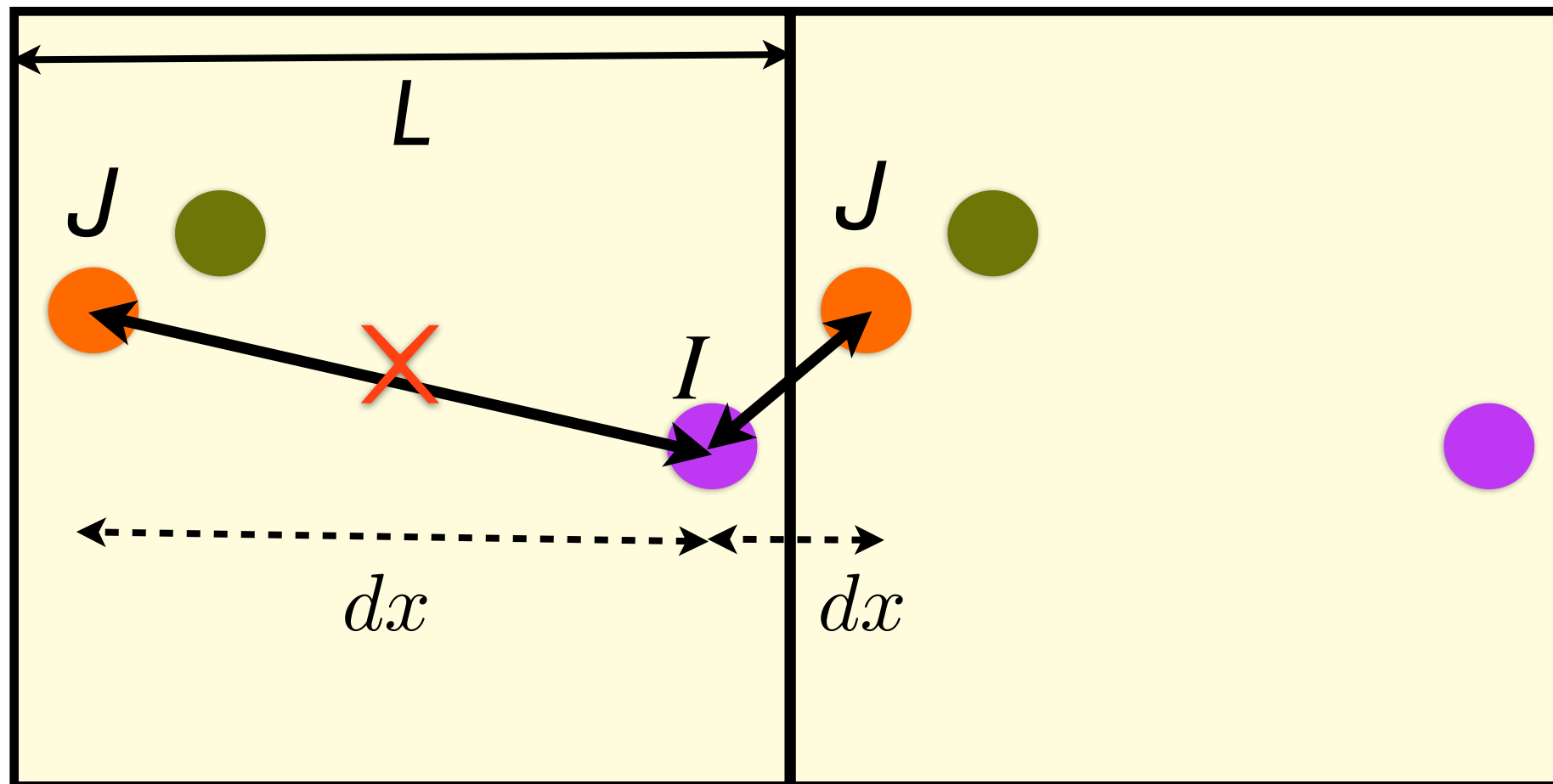
such boundary
effects should be
avoided!

Periodic Boundary Conditions



if an atom goes
out of the simulation
box, the same atom
should come into the
box from the opposite
side

Minimum image convention:



longest
distance
should not be
larger than $L/2$

○

correct distance:

$$dx = x_I - x_J$$

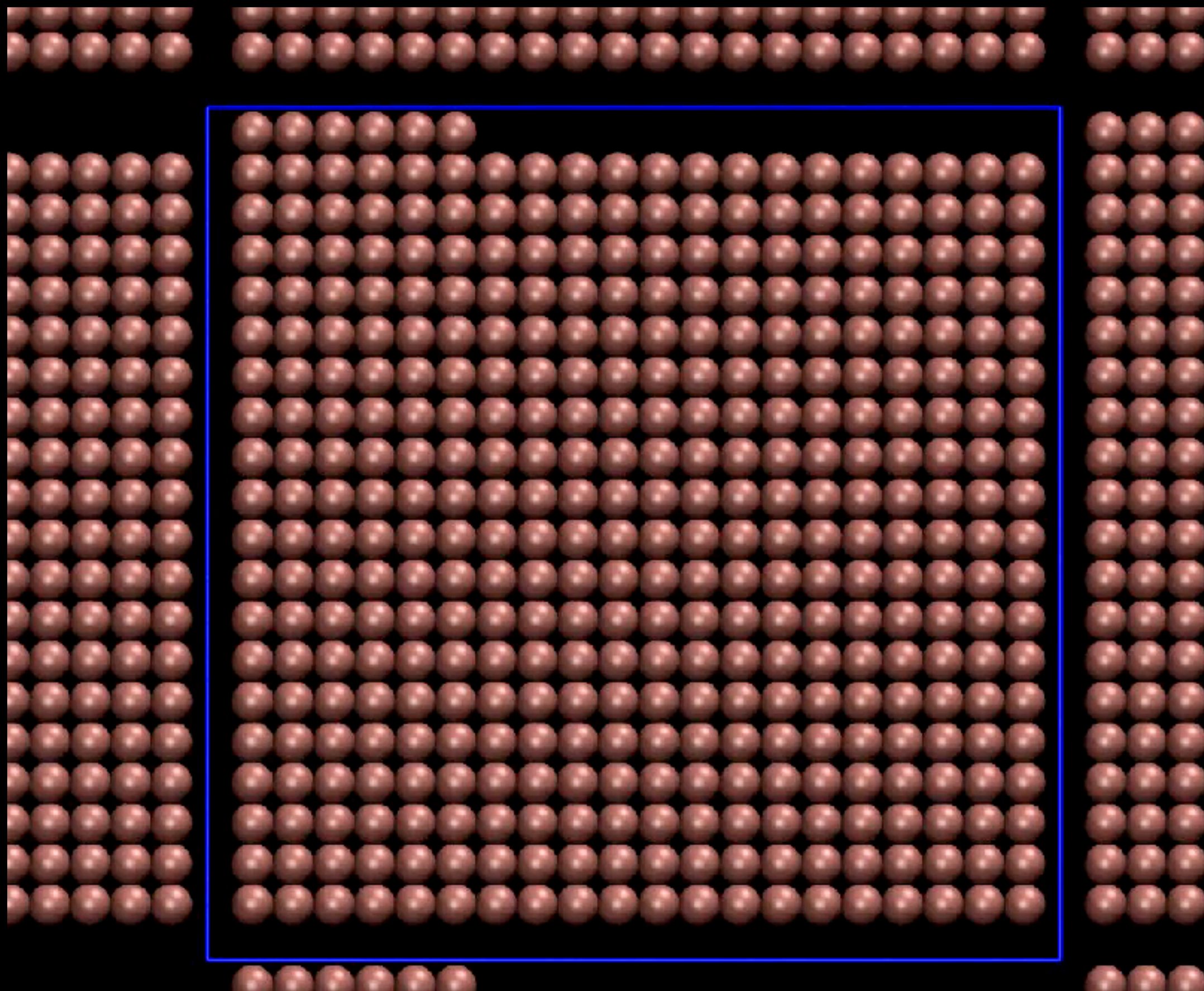
If $|dx| > L/2$, then

$$dx = dx - L \operatorname{sgn}(dx)$$

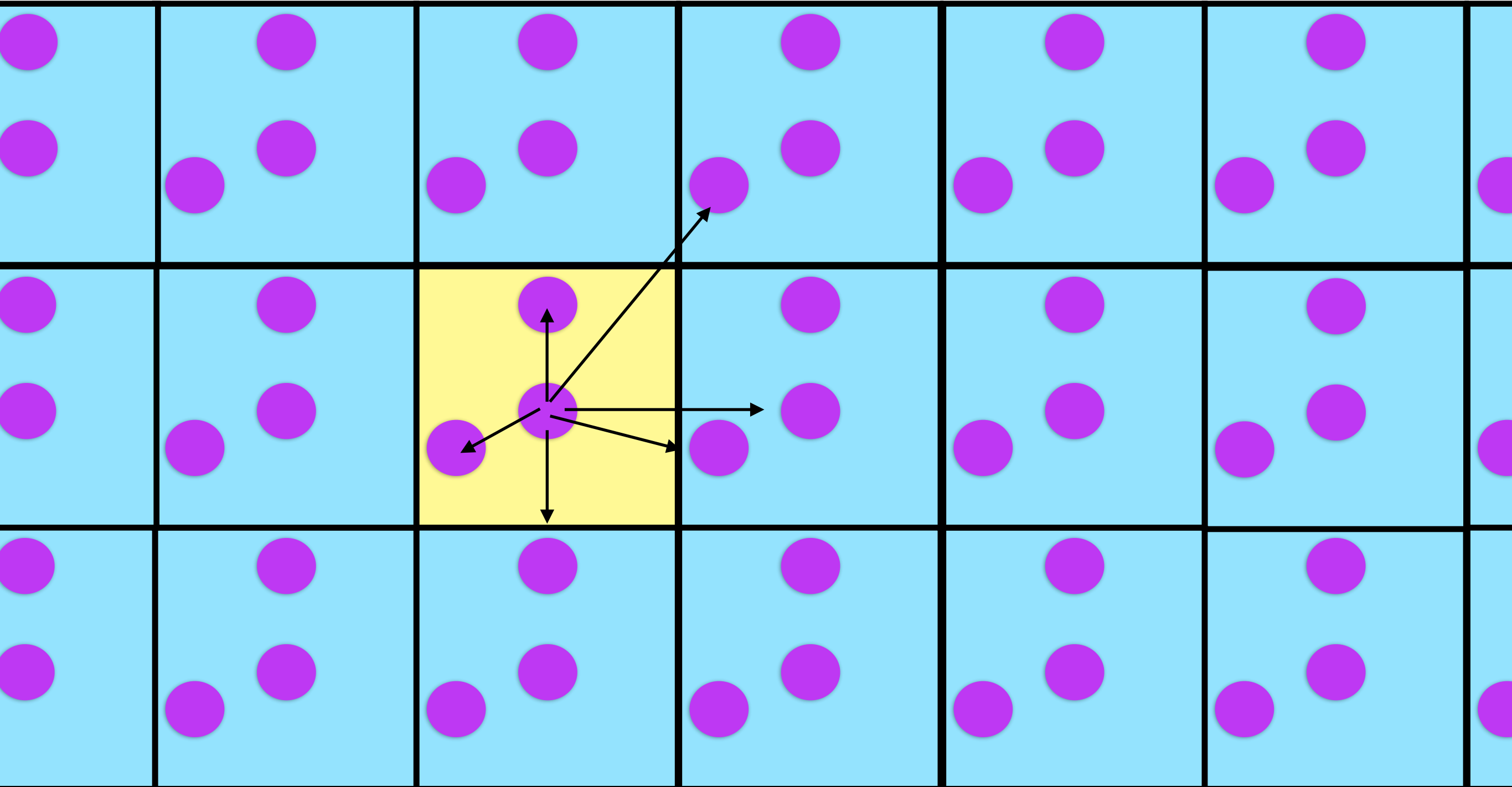
**wrapping the
coordinates:**

If $(x \geq L)$ then $x = x - L$

If $(x < 0)$ then $x = x + L$



- A given particle is interacting with all other particles in the infinite periodic lattice of the same cell



$$U = \frac{1}{2} \sum_{i,j,\mathbf{n}} u(|\mathbf{R}_{ij} + \mathbf{n}L|)$$

$$i \neq j, \text{ if } \mathbf{n} \neq 0$$

L is the size of the periodic box

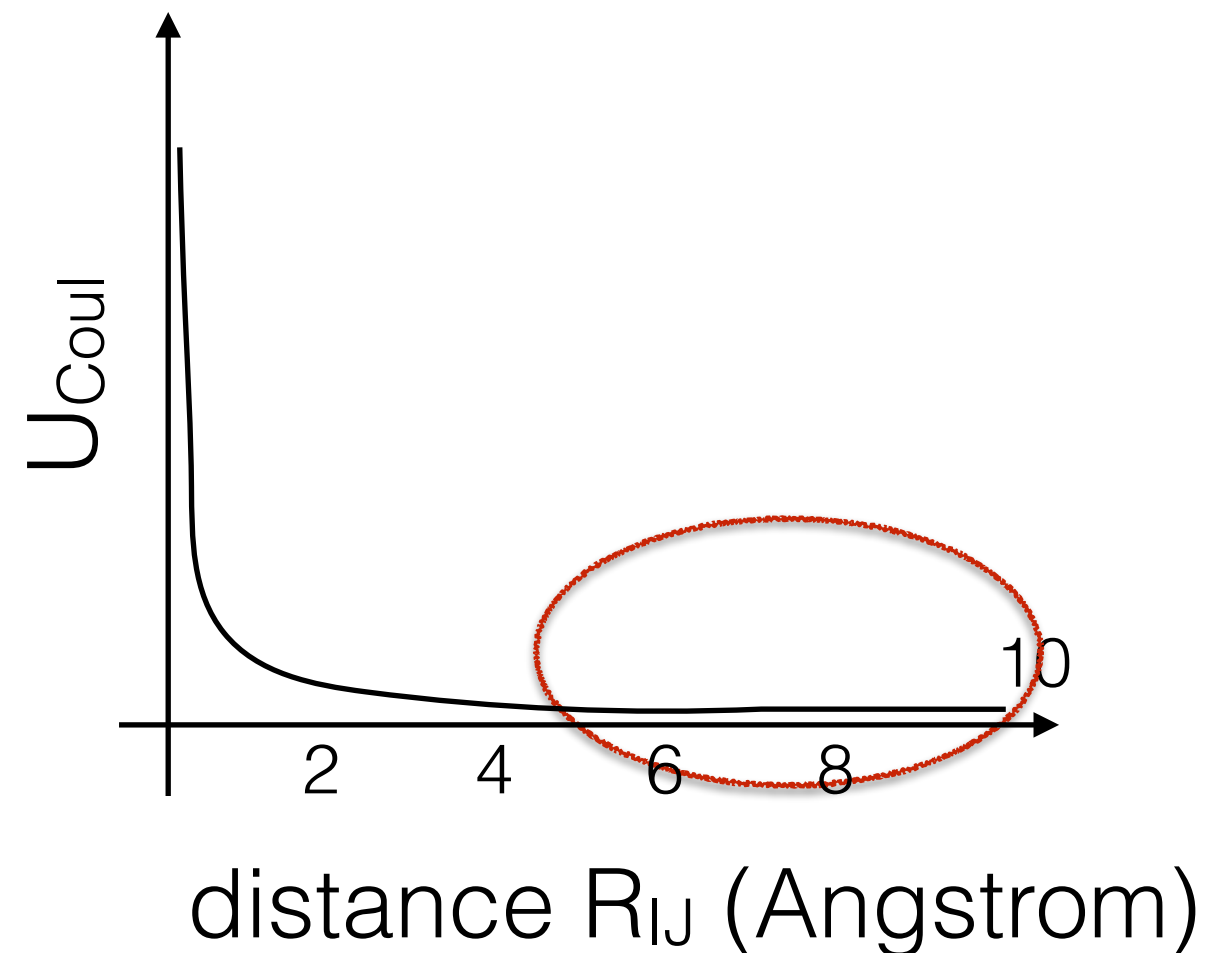
\mathbf{n} is an arbitrary vector of three integers (which decides the periodic image)

In principle, infinite periodic images to be considered!

$$U_{\text{Coul}} = \frac{1}{4\pi\epsilon_0} \sum_{I \neq J} \frac{q_I q_J}{R_{IJ}}$$

- Electrostatic interactions are long ranged
- Introducing short-range cutoff leads to spurious errors, especially treating bulk systems

How to compute long range contributions efficiently?



Ewald Sums

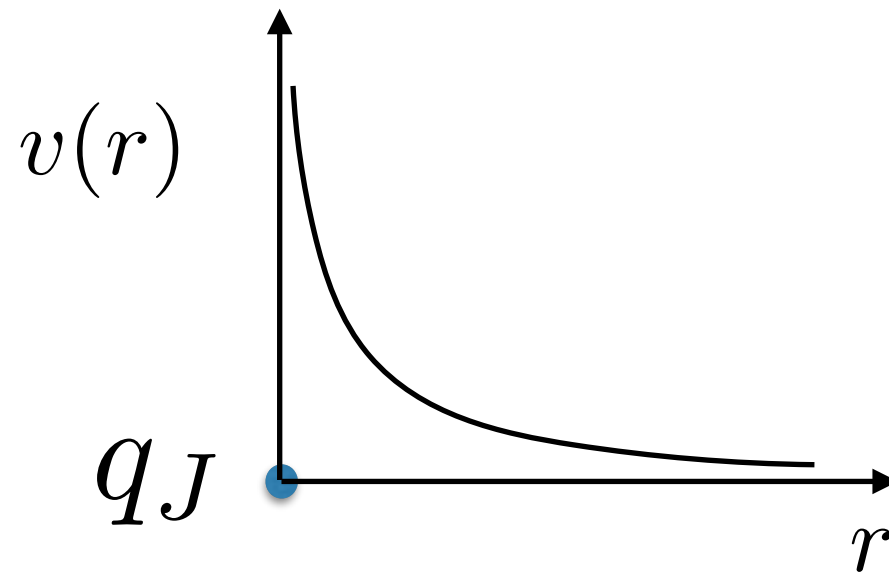
$$U_{\text{Coul}} = \frac{1}{2} \sum_{IJ}^N \frac{q_I q_J}{R_{IJ}}$$

$$= \frac{1}{2} \sum_I^N q_I v(\mathbf{R}_I)$$

where $v(\mathbf{R}_I) = \sum_{J, \mathbf{n}} \frac{q_J}{|\mathbf{R}_{IJ} + \mathbf{n}L|}$

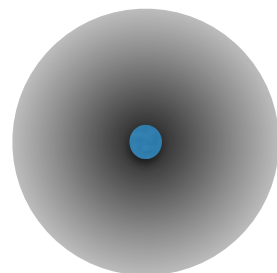
electrostatic potential

This sum is poorly convergent and thus not usable in computations as such

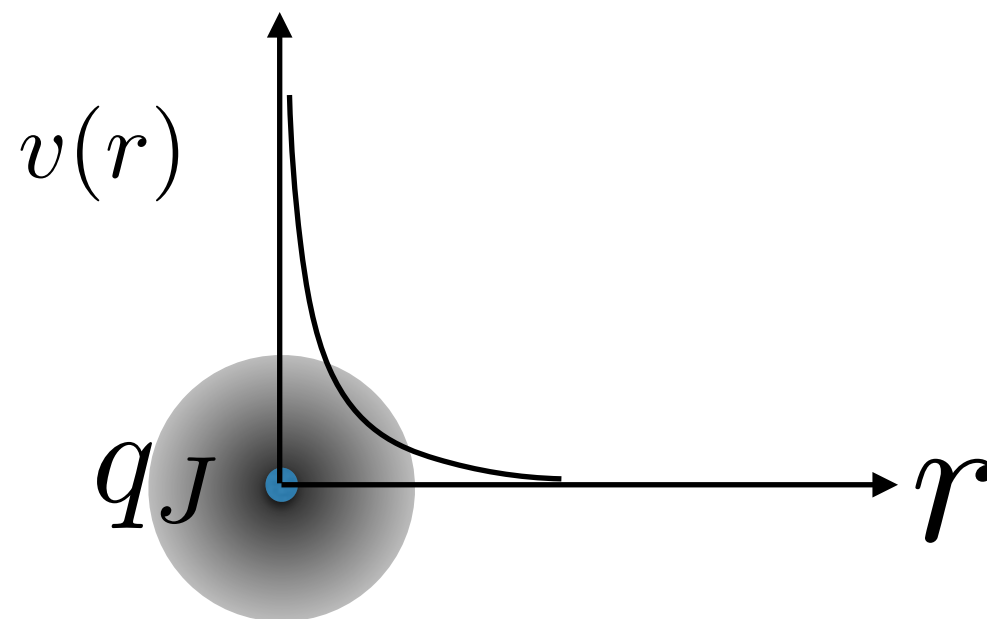


Here charges are treated as point charges - thus the sum to get the potential at a point r can be considered as sum of some delta functions

Let us assume that every point charge J is surrounded by a diffuse charge distribution with opposite sign (total charge of the cloud exactly cancels q_J)



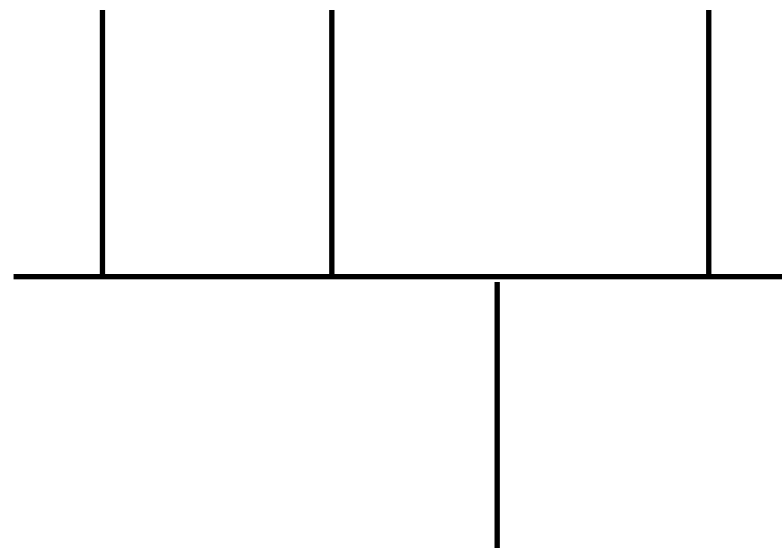
- In the presence of the background diffused charge, the electrostatic potential rapidly goes to zero (as the charge of q_j is screened by the cloud)



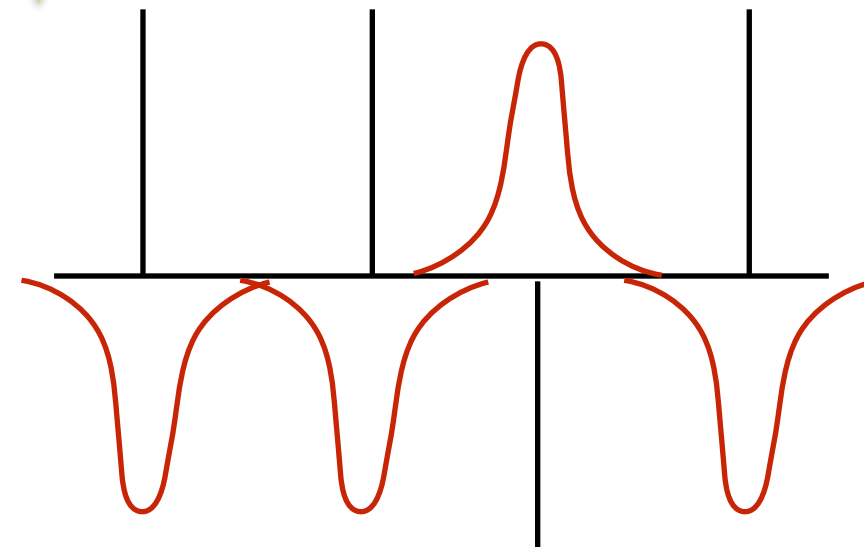
How fast it will decay, will depends on the functional form of the screening charge cloud

Finally, we correct the potential by subtracting the potential ONLY due to the charge distribution - Usually one uses Gaussian distribution

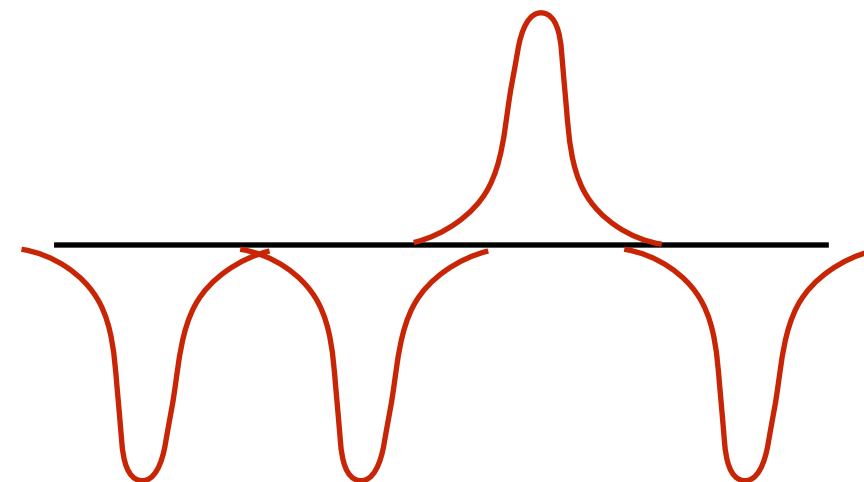
fast converging (direct sum - short range)



=



— (minus)



computed by Fourier sums (converging sums!)