Molecular Mechanics:

The Ab Initio Foundation

Ju Li

GEM4 Summer School 2006
Cell and Molecular Mechanics in BioMedicine
August 7–18, 2006, MIT, Cambridge, MA, USA

Outline

- Why are electrons quantum?
- Born-Oppenheimer approximation and the energy surface
- Hartree-Fock and density functional theory
- Interatomic potentials

The electron problem: basic facts

- Electrons $\{\mathbf{x}_i\}$, i=1..n: $m_\mathrm{e} = 9.1\times10^{-31}$ kg
- Nuclei $\{\mathbf{X}_I\}$, I=1..N: $M_\mathrm{I} \approx A_\mathrm{I} M_\mathrm{p}$, $M_\mathrm{p} \approx 2000 m_\mathrm{e}$
- They interact electrostatically as

$$V_{iI} = \frac{-Z_I e^2}{|\,\mathbf{x}_i - \mathbf{X}_I\,|}$$

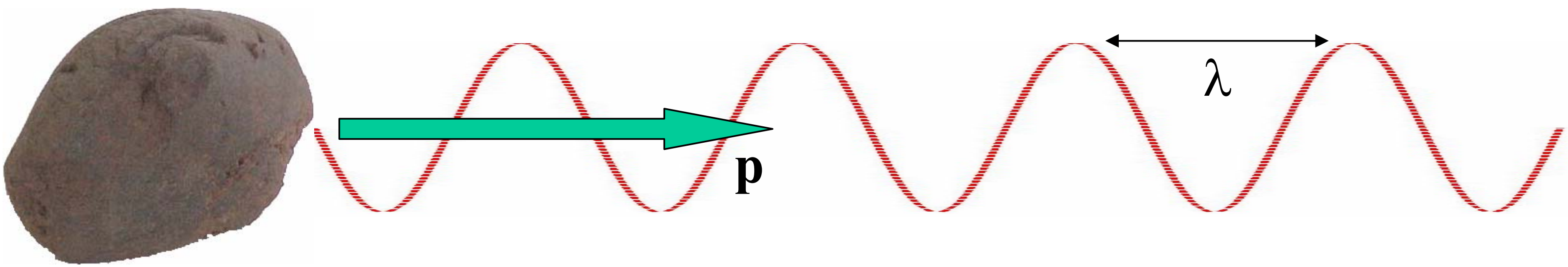
I is a carbon ion: $A_I = 12,\ Z_I = 6$

Why electrons must be considered quantum, while ions are often considered classical, objects?

$$\text{de Broglie wavelength: } \lambda = \frac{h}{|\mathbf{p}|}$$

$$h = 6.6\times10^{-34}\,\mathrm{J\cdot s}: \quad \text{Planck's constant}$$

$$\mathbf{p}: \ \text{momentum}$$



Electrons and ions share the same energy scale:

$$\frac{|\mathbf{P}_I|^2}{2M_I} \sim \ 13.6\ \mathrm{eV} \ \sim \frac{|\mathbf{p}_i|^2}{2m_\mathrm{e}}, \quad \text{energy of C-C bond} \sim 3.5\ \mathrm{eV}$$

Plugging in the numbers, we get de Broglie wavelength of

electron: 3.3 Å

hydrogen: 0.08 Å

carbon: 0.02 Å

Note these are quantum, not thermal, fluctuations

Image removed due to copyright restrictions.
Diagram of an generic amino acid as an example molecule.

- electron's wavelength permeates through several structural units (bonds) and so must be treated quantum mechanically

- carbon's wavelength is well-localized

- hydrogen is a borderline case.

Born-Oppenheimer approximation

Max Born

J. Robert Oppenheimer

The electrons minimize their quantum mechanical energy as if the ions are *immobile*; the resulting total energy (electrons + ions) is $V_{\mathrm{BO}}(\{\mathbf{X}_I\})$, the Born-Oppenheimer energy surface, aka energy landscape, interatomic potential.

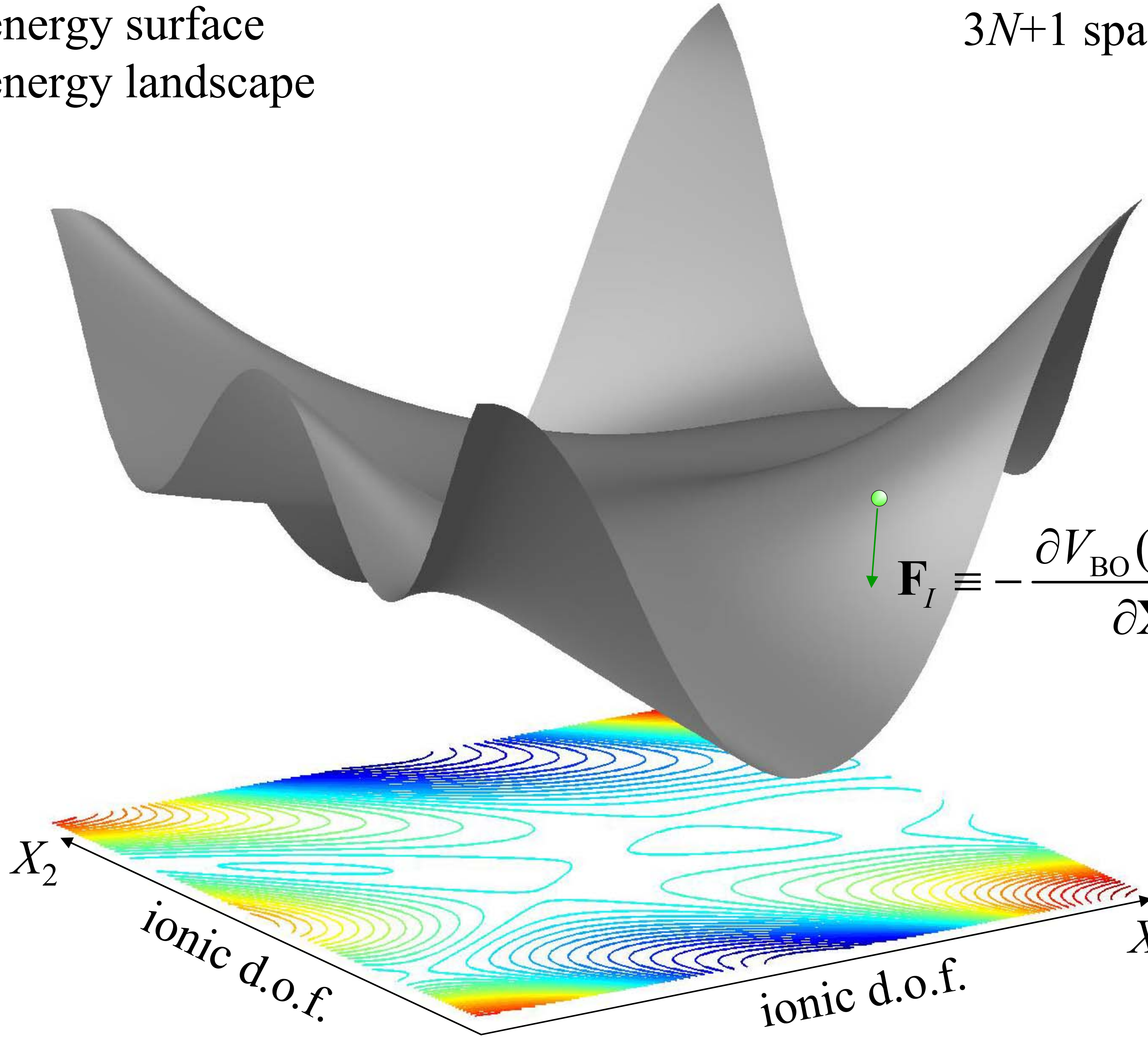
B-O approximation is also called the adiabatic approximation. The idea is that electrons "move" so much faster than the ions, that they are at their ground state $\Psi_{\mathrm{G}}(\{\mathbf{x}_i\})$ for a given ionic configuration $\{\mathbf{X}_I\}$.

Addendum: The ions move classically on the BO energy surface according to Newton's 2nd law:

$$M_I \ddot{\mathbf{X}}_I = -\frac{\partial V_{\mathrm{BO}}(\{\mathbf{X}_I\})}{\partial \mathbf{X}_I}$$

BO energy surface
aka energy landscape

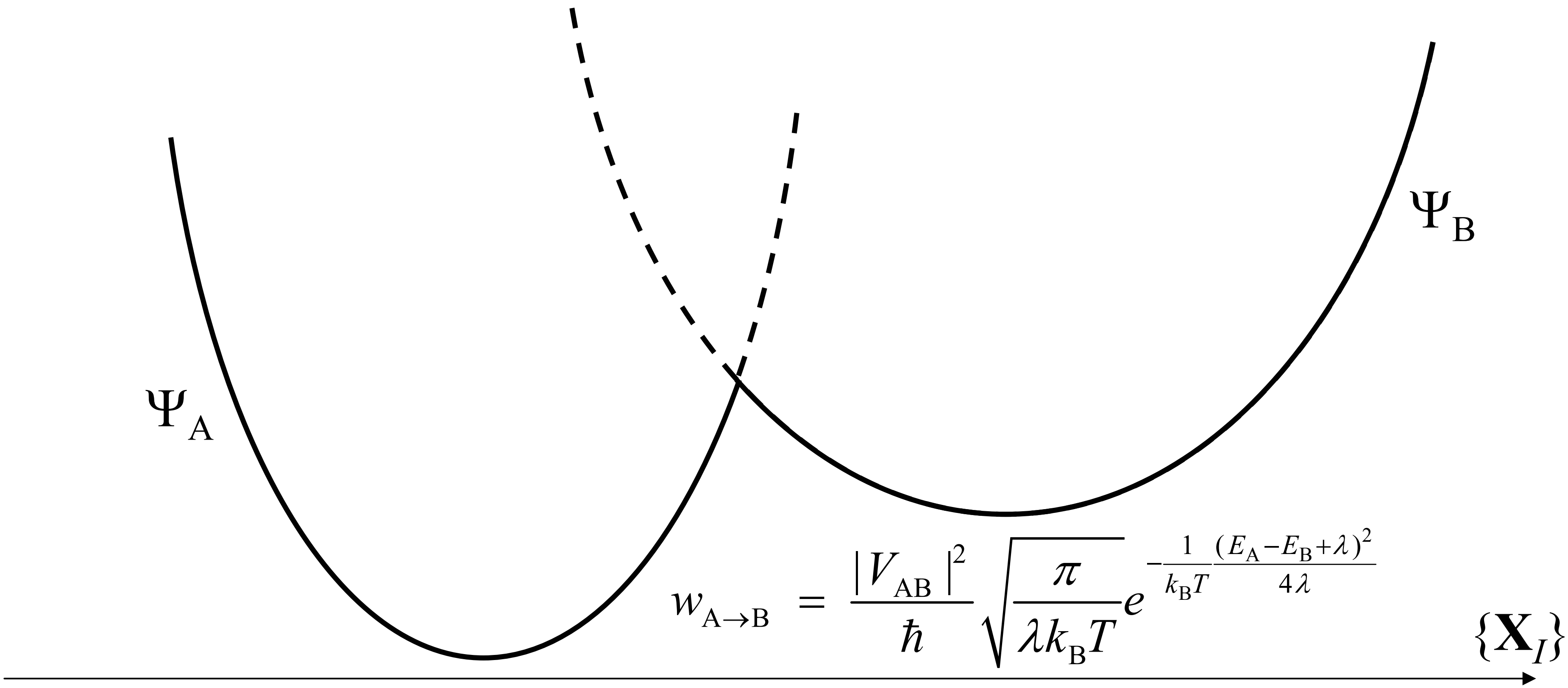
$3N+1$ space



$$\mathbf{F}_I \equiv -\frac{\partial V_{\mathrm{BO}}(\{\mathbf{X}_I\})}{\partial \mathbf{X}_I}$$

BO approximation breaks down when

1. The molecule is optically excited (electronic excited states)
2. During diabatic electron transfer process (Marcus theory)



$$w_{A \to B} = \frac{|V_{AB}|^2}{\hbar}\sqrt{\frac{\pi}{\lambda k_B T}}\, e^{-\frac{1}{k_B T}\frac{(E_A - E_B + \lambda)^2}{4\lambda}}$$

Addendum breaks down when

1. For light-mass ions like hydrogen, or at low temperature $k_B T << \hbar\omega_D$ (Then, even the ions need to be treated quantum mechanically.)

The quantum mechanical life of electrons

$$\widehat{H}\Psi_{\mathrm{G}}(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N)=V_{\mathrm{BO}}\Psi_{\mathrm{G}}(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N)$$

In the non-relativistic limit:

$$\widehat{H}=\sum_{i=1}^{n}\frac{-\hbar^2\nabla_i^2}{2m_{\mathrm{e}}}+\sum_{i\neq j}\frac{e^2}{2|\mathbf{x}_i-\mathbf{x}_j|}-\sum_{i,I}\frac{Z_I e^2}{|\mathbf{x}_i-\mathbf{X}_I|}$$

$$+\sum_{I\neq J}\frac{Z_I Z_J e^2}{2|\mathbf{X}_I-\mathbf{X}_J|}$$

If one attempts to get at V_{BO} by solving the above equation under *rational* and *non-material-specific* approximations, without using any experimental input, the result should then depend on and only depend on numerical values of $\hbar$, m_{e}, e, $\{Z_{\mathrm{I}}\}$.

This is called *ab initio* calculation.

The problem of information explosion

To store a single-variable function $f(x)$, $0<x<1$:

use 10 spline points, each spline data

(in double precision) is 8 bytes: 80 bytes

We have 20 electrons in a box, $0<x<1$, $0<y<1$, $0<z<1$

$$\Psi_G(\mathbf{x}_{1,}\ \mathbf{x}_{2\ ,\ldots,}\ \mathbf{x}_{20}) = \Psi_G(x_{1,} y_{1,} z_{1,} x_{2,} y_{2,} z_{2,\ \ldots,} x_{20,} y_{20,} z_{20})$$

A 60-dimensional function: needs 10^{60} spline data points

total storage required: 8×10^{60} bytes

A CD can store 6×10^{8} bytes: needs 10^{52} CD's

Say each CD is one gram, 10^{52} gram

Mass of the sun: 2×10^{33} gram

There are some symmetry relations:
the electrons are indistinguishable Fermions:

$$\Psi(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_i,\ldots,\mathbf{x}_j,\ldots,\mathbf{x}_n) = -\Psi(\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_j,\ldots,\mathbf{x}_i,\ldots,\mathbf{x}_n)$$

but that does not solve the explosion fundamentally.

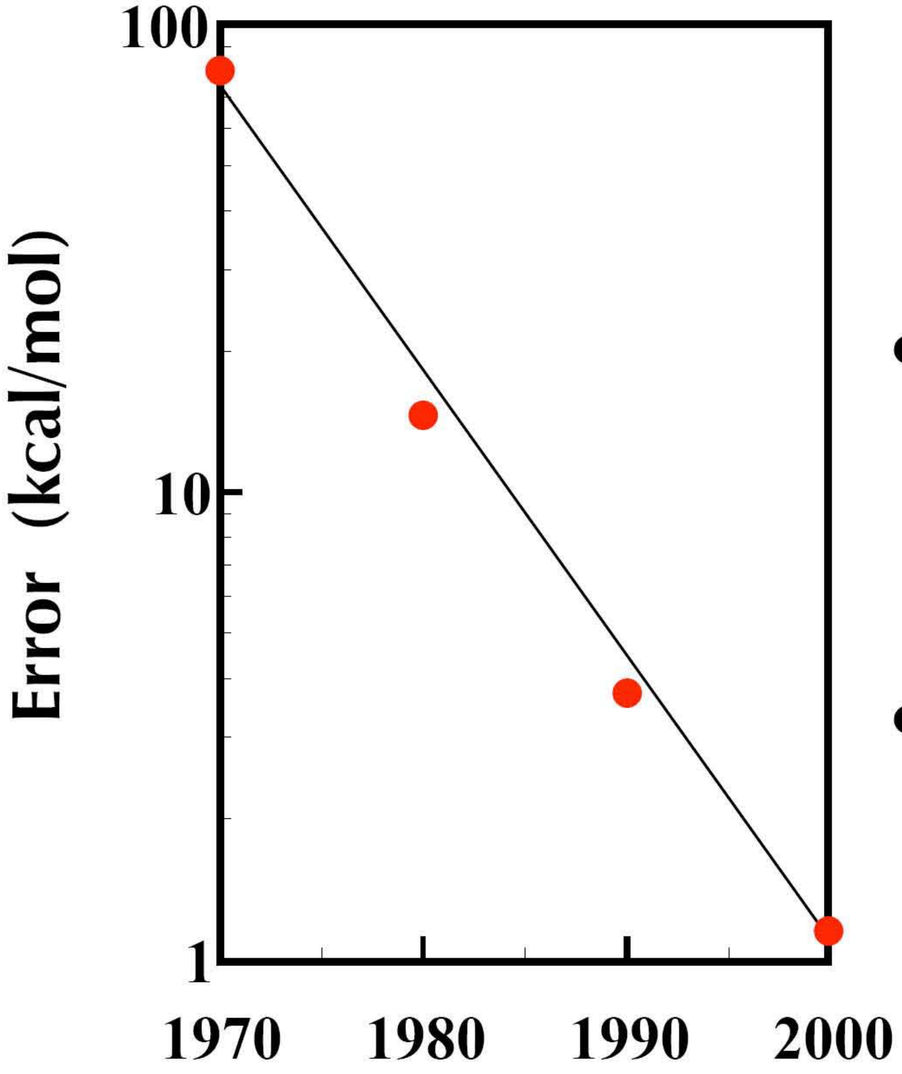
1998 Nobel Prize in Chemistry

Image removed due to copyright restrictions.

John A. Pople — quantum chemistry

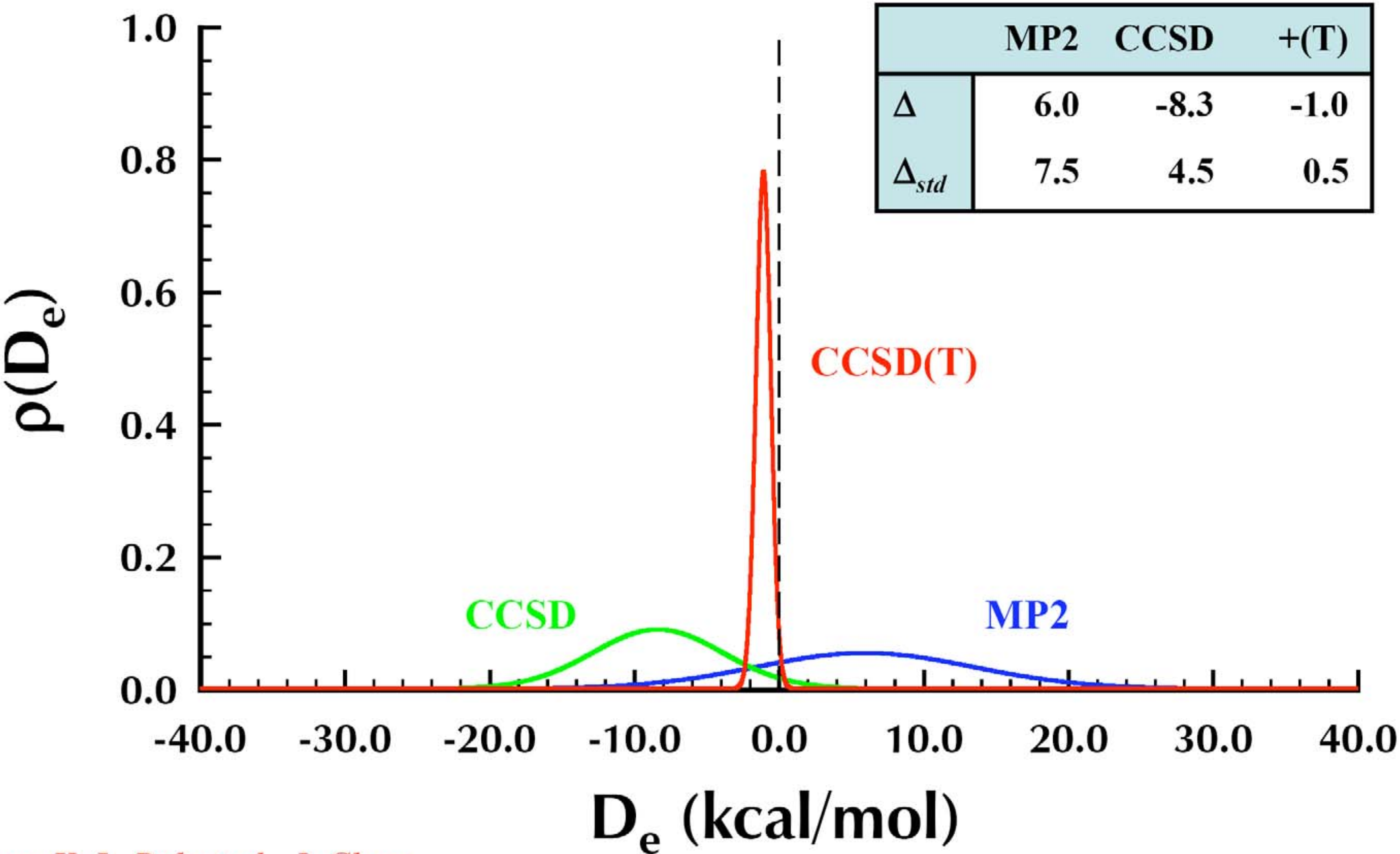
Walter Kohn — density-functional theory (DFT)

Progress in the last 50 years has been tremendous.
Significant number of researchers "most cited chemists",
"most cited physicists".



from data by Thom H. Dunning, Jr.

- Bond energies critical for describing many chemical phenomena; difficult to determine experimentally
- Accuracy of calculated bond energies increased dramatically from 1970-2000
- Due to advances in
  - Theoretical methodology
  - Computational techniques
  - Computing technology



	MP2	CCSD	+(T)
Δ	6.0	-8.3	-1.0
Δ_{std}	7.5	4.5	0.5

Data from K. L. Bak et al., J. Chem. Phys. ***112****, 9229-9242 (2000)*

MP2: Møller-Plesset perturbation theory 2nd order
CCSD: Coupled cluster with single- and double-excitations
CCSD(T): plus triple excitations calculated by perturbation theory

Hartree-Fock Theory

Assume $\Psi(\mathbf{x}_{1,}\ \mathbf{x}_{2\ ,\ldots,}\ \mathbf{x}_n)$ can be well-approximated by a *single* Slater determinant:

$$S = \frac{1}{\sqrt{n!}} \begin{vmatrix} \tilde{\psi}_1(\mathbf{x}_1) & \tilde{\psi}_2(\mathbf{x}_1) & \cdots & \tilde{\psi}_n(\mathbf{x}_1) \\ \tilde{\psi}_1(\mathbf{x}_2) & \tilde{\psi}_2(\mathbf{x}_2) & \cdots & \tilde{\psi}_n(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \tilde{\psi}_1(\mathbf{x}_n) & \tilde{\psi}_2(\mathbf{x}_n) & \cdots & \tilde{\psi}_n(\mathbf{x}_n) \end{vmatrix}$$

$3n$-dimensional function $\rightarrow$ n 3-dimensional functions:

$$\tilde{\psi}_1(\mathbf{x}), \tilde{\psi}_2(\mathbf{x}), \ldots, \tilde{\psi}_n(\mathbf{x})$$

If n=20, $20\times10^3 \times 8$ = 160 kilobytes

8×10^{60} bytes $\rightarrow$ 160 kilobytes compression

Each trial wave function also contains spin information:

$$\tilde{\psi}(\mathbf{x}) \;=\; \psi(\mathbf{x})\,|\!\uparrow\rangle \;\text{ or }\; \psi(\mathbf{x})\,|\!\downarrow\rangle: \;\; \text{just 1-bit extra}$$

Contribution to total energy due to $\tilde{\psi}_i$ and $\tilde{\psi}_j$

$$= \;\; E_{\text{Hartree}}[\tilde{\psi}_i,\tilde{\psi}_j] - E_{\text{Exchange}}[\tilde{\psi}_i,\tilde{\psi}_j]$$

$$E_{\text{Hartree}}[\tilde{\psi}_i,\tilde{\psi}_j] \;\equiv\; e^2 \iint d\mathbf{x}d\mathbf{x}' \frac{\overbrace{(\tilde{\psi}_i^*(\mathbf{x})\tilde{\psi}_i(\mathbf{x}))}^{\rho_i(\mathbf{x})}\overbrace{(\tilde{\psi}_j^*(\mathbf{x}')\tilde{\psi}_j(\mathbf{x}'))}^{\rho_j(\mathbf{x}')}}{|\,\mathbf{x}-\mathbf{x}'\,|}$$

$$E_{\text{Exchange}}[\tilde{\psi}_i,\tilde{\psi}_j] \;\equiv\; e^2 \iint d\mathbf{x}d\mathbf{x}' \frac{(\tilde{\psi}_i^*(\mathbf{x})\tilde{\psi}_j(\mathbf{x}))(\tilde{\psi}_i^*(\mathbf{x}')\tilde{\psi}_j(\mathbf{x}'))}{|\,\mathbf{x}-\mathbf{x}'\,|}$$

$E_{\text{Hartree}}[\tilde{\psi}_i, \tilde{\psi}_j] > 0$ does not care if $\tilde{\psi}_i, \tilde{\psi}_j$ have the same spin

$$E_{\text{Exchange}}[\tilde{\psi}_i, \tilde{\psi}_j] = \begin{cases} 0 \text{ if } \tilde{\psi}_i, \tilde{\psi}_j \text{ have different spin} \\ >0 \text{ if } \tilde{\psi}_i, \tilde{\psi}_j \text{ have same spin} \end{cases}$$

An occupied wavefunction does not see itself in Hartree-Fock theory:

$$E_{\text{Hartree}}[\tilde{\psi}_i, \tilde{\psi}_i] = E_{\text{Exchange}}[\tilde{\psi}_i, \tilde{\psi}_i]$$

$$\rightarrow E_{\text{Hartree}} \ \tilde{\psi}_i \ \tilde{\psi}_i] - E_{\text{Exchange}}[\tilde{\psi}_i, \tilde{\psi}_i] = 0$$

$\rightarrow$ No self-interaction

[,

Exchange interaction stabilizes occupation of same-spin wavefunction with large spatial overlap.

Due to the structure of the Slater determinant, two electrons of the same spin "automatically" avoid each other, creating so-called "exchange-hole" in their pair-correlation function.

Hartree-Fock is beautiful. But in reality:

$$\Psi_{\mathrm{G}} = a_1 S_1 + a_2 S_2 + a_3 S_3 + \dots$$

each S_μ is a Slater determinant

By optimizing the coefficients a_1, a_2, a_3, … and also the Slater determinants, one can further reduce the energy beyond the best single-determinant (Hartree-Fock) energy.
This energy reduction is called correlation energy.

The brute-force way of doing above is called configuration interaction (CI).

Both exchange and correlation energies stablize a many-electron system beyond naïve Coulomb interactions. Exchange energy tends to be larger in magnitude (10×) than correlation energy.

Full CI is formally exact, but it has very bad scaling, something like $O(n^{10})$

The present "gold standard", CCSD(T), maintains most of the accuracy and has better scaling. But it is still expensive, something like $O(n^7)$, so still limited to small molecules, say ~20 atoms.

In this sense, density functional theory (DFT) is a "poor man's way" of taking account of both exchange and correlation. It tends to be cheaper than even Hartree-Fock.

DFT is the technology that underlies most of the condensed-matter research, as well as a very significant part of biomolecular modeling.

Hohenberg and Kohn, Phys. Rev. 136 (1964) B864:

$$V_{\mathrm{BO}}(\{\mathbf{X}_I\}) = V_{\mathrm{BO}}[\rho(\mathbf{x})]$$

$\rho(\mathbf{x})$: single-electron density at ground state

$\leftrightarrow$ $v(\mathbf{x})$: ion-electron (external) potential

$\leftrightarrow$ $\Psi_{\mathrm{G}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)$

So from informatics point of view, instead of treating n 3-dimensional functions, one formally only needs to treat one 3-dimensional function. 8×10^{60} bytes $\leftrightarrow$ 8 kbyte

Kohn and Sham, Phys. Rev. 140 (1965) A1133:

It is still exact, and physically expedient, to decompose

$$V_{\mathrm{BO}}[\rho(\mathbf{x})] = V_{\mathrm{Independent}}[\rho(\mathbf{x})] + V_{\text{e-e Hartree}}[\rho(\mathbf{x})] + V_{\text{Exchange-correlation}}[\rho(\mathbf{x})]$$

where $V_{\mathrm{Independent}}[\rho(\mathbf{x})]$ is the energy of a fictitious, *independent-electrons* system having the same $\rho(\mathbf{x})$, where

$$V_{\mathrm{independent}}[\rho(\mathbf{x})] = \boxed{\sum_{i=1}^{n} \frac{-\hbar^2}{2m_{\mathrm{e}}} \int d\mathbf{x}\, \psi_i^*(\mathbf{x}) \nabla^2 \psi_i(\mathbf{x})}$$

$$-\sum_{i,I} \int d\mathbf{x} \frac{Z_I e^2 \rho_i(\mathbf{x})}{|\mathbf{x} - \mathbf{X}_I|} + \sum_{I \neq J} \frac{Z_I Z_J e^2}{2|\mathbf{X}_I - \mathbf{X}_J|},$$

$$\rho(\mathbf{x}) \equiv \sum_{i=1}^{n} \psi_i^*(\mathbf{x}) \psi_i(\mathbf{x}) \equiv \sum_{i=1}^{n} \rho_i(\mathbf{x}).$$

kinetic energy: difficult to *express* well (either in fictitious- or real-electrons system) as a *local* functional of $\rho(\mathbf{x})$.

$\psi_1(\mathbf{x}), \psi_2(\mathbf{x}), ..., \psi_n(\mathbf{x})$ are called Kohn-Sham wave functions. 8×10^{60} bytes $\leftrightarrow$ 160 kbyte.

In appearance, they look similar to Hartree-Fock trial wave functions. But their interpretations are shadowy, and it would be inappropriate to call DFT a "single configuration" (single determinant) method.

Local density approximation (LDA):

$$\mathrm{V}_{\text{Exchange-correlation}}[\rho(\mathbf{x})] \approx \int d\mathbf{x} \rho(\mathbf{x}) v_{\mathrm{XC}}(\rho(\mathbf{x}))$$

Perdew and Zunger parameterized $v_{\mathrm{XC}}(\rho)$ using the Quantum Monte Carlo data by Ceperley and Alder (1980) for *homogeneous electron gas*.

By definition, LDA works well when the electron density is nearly uniform, for instance inside a simple metal.

But when the electron density varies violently, for instance in gas-phase molecules, LDA could fail.

Various attempts of Generalized Gradient Approximation (GGA), such as PW91 and PBE96, improve results somewhat in condensed phases such as water, but serious problems remain for molecules.

An important reason is LDA/GGA sometimes underestimates exchange energy → self-interaction.

Comes in hybrid functionals (*Becke, J. Chem. Phys. 98 (1993) 5648*): The DFT exchange energy is mixed with Hartree-Fock exact exchange (nonlocal).

Hybrid functionals such as B3LYP work well for biomolecular systems, but because of the mixing parameter, many people do not consider them true *ab initio* methods.

DFT typically treat hundreds of atoms with good basis.

Currently very active developments:
orbital-dependent density functionals, LDA+U,
self-interaction correction (SIC),
time-dependent DFT (TDDFT), …

Planewave codes: VASP, PWSCF, CPMD, ABINIT, DACAPO, CASTEP…

Local orbital basis: Gaussian, NWChem, GAMESS, DMol, SIESTA, …

Semi-Empirical Electronic-Structure Methods (Hückel Theory, Linear Combination of Atomic Orbitals, Molecular Orbital Theory, Empirical Valence Bond, Tight-Binding, …)

Typically treat thousands of atoms with minimal basis.

Experimental or *ab initio* information used to fit intrinsic *electronic* quantities: orbital overlap, hopping integral, etc.

For a given ion configuration $\{\mathbf{X}_I\}$: an *electronic Hamiltonian* is first assembled, and then diagonalized, in usage.

Parameter sets: AM1, PM3, PM5

codes: MOPAC

Also implemented in: Gaussian, GAMESS, CAChe …

Interatomic Potential / Force Field

Direct parameterization of $V_{\mathrm{BO}}(\{\mathbf{X}_I\})$ without touching the electronic degrees of freedom, in usage.

$I = 1..N$: typically N=10^4-10^8

Images removed due to copyright restrictions.
Examples of bond stretching, rotation, and non-bonding interactions.

Leach, Molecular modelling: principles and applications (Prentice-Hall, New York, 2001).

A typical Force Field for macromolecules looks like:

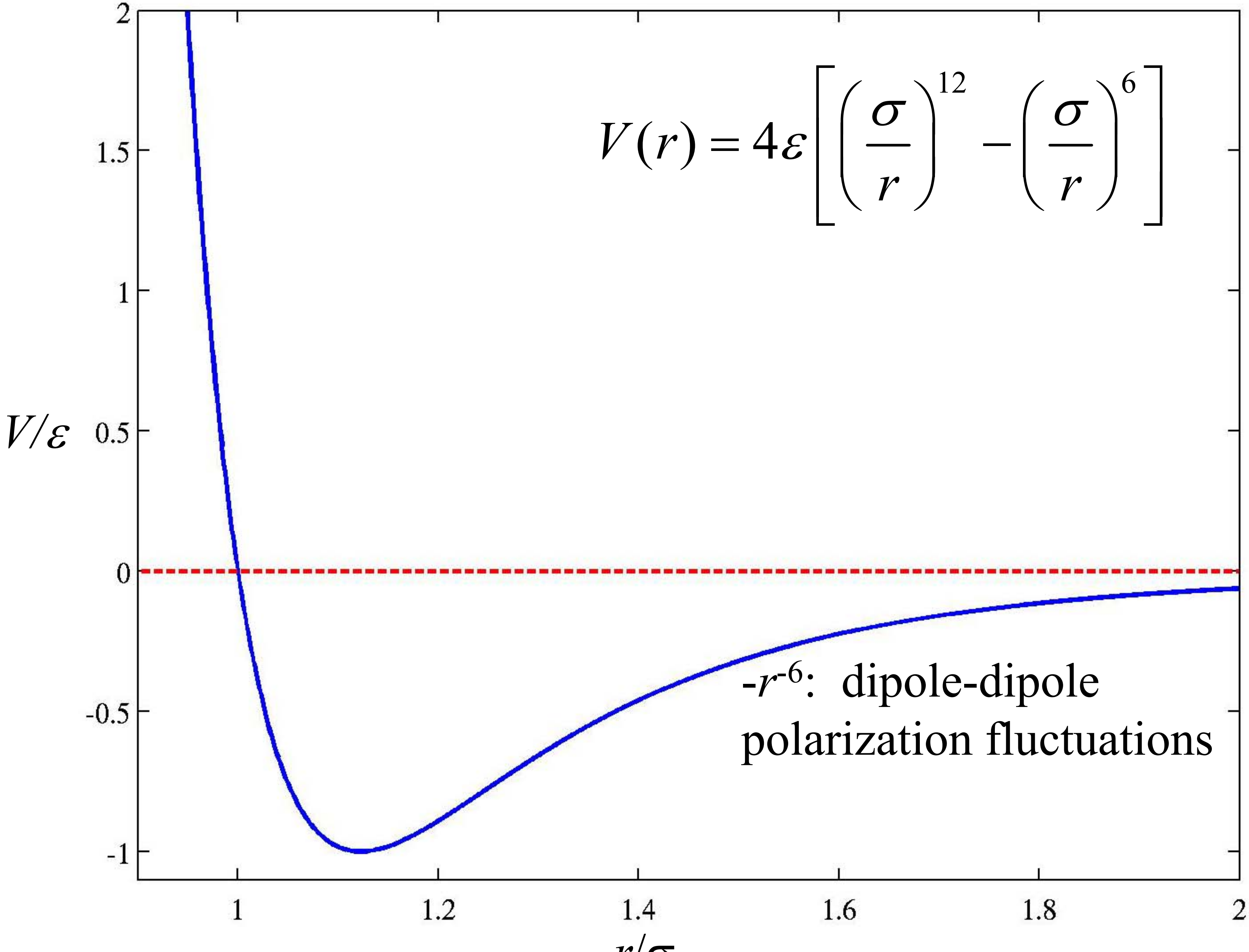
$$V_{\mathrm{BO}}(\{\mathbf{X}_I\}) = \sum_{\text{bonds}} \overbrace{\frac{k_I}{2}(l_I - l_I^0)^2}^{\text{bond stretch: pair}} + \sum_{\text{angles}} \overbrace{\frac{h_I}{2}(\theta_I - \theta_I^0)^2}^{\text{bending: triplet}} \leftarrow \text{covalent}$$

$$+ \sum_{\text{dihedral angles}} \overbrace{\frac{g_I}{2}(1+\cos(m(\omega_I - \omega_I^0)))^2}^{\text{torsion: quartet}} \leftarrow \begin{array}{l}\text{covalent \&}\\ \text{delocalized}\\ \pi\ \text{bonding}\end{array}$$

$$+ \sum_{I>J} \overbrace{4\varepsilon_{IJ}\left[\left(\frac{\sigma_{IJ}}{r_{IJ}}\right)^{12} - \left(\frac{\sigma_{IJ}}{r_{IJ}}\right)^{6}\right]}^{\text{non-bond dispersive interaction: pair}} \leftarrow \begin{array}{l}\text{long-range}\\ \text{correlation}\end{array}$$

$$+ \sum_{I>J} \overbrace{\frac{q_I q_J e^2}{r_{IJ}}}^{\text{non-bond Coulomb interaction: pair}} \leftarrow \text{ionic}$$

Dispersive / van der Waals interaction:



Treatment of Long-Range Electrostatic Interactions:

Ewald sum: Decompose $1/r$ into
Long-range smooth + Short-range sharp contributions
Long-range smooth part is summed in reciprocal (**k-**) space
Short-range sharp part is summed in real space

Modern techniques such as Fast Multipole, Particle Mesh Ewald methods further enhance the efficiency to nearly $O(N)$.

Further Readings:

Leach, Molecular modelling: principles and applications (Prentice-Hall, New York, 2001).

Jensen, Introduction to computational chemistry (Wiley, New York, 1999).

Schlick, Molecular Modeling and Simulation (Springer-Verlag, Berlin, 2002).