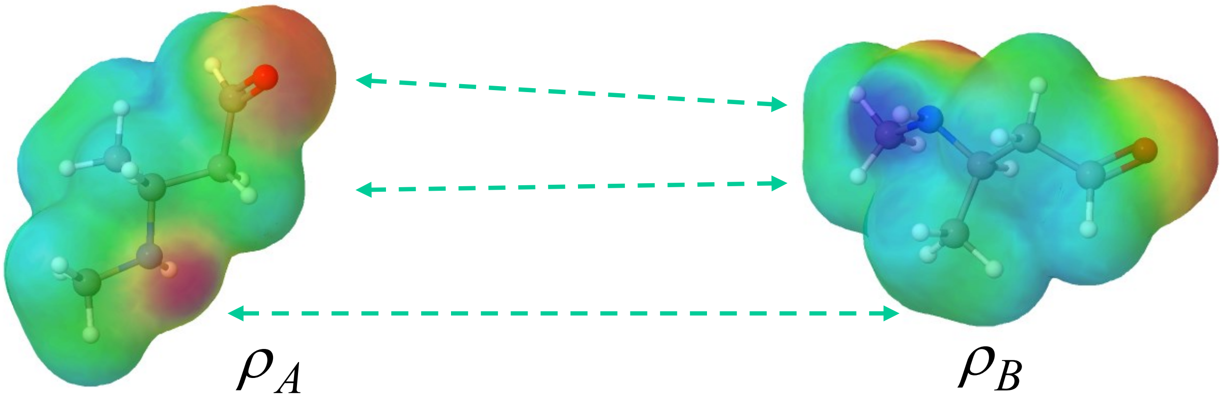
Electrostatic Interactions Between Molecules

Electrostatic interactions between molecules can be strong and long ranged. These must be taken into account as a separate term in the force field.



Molecules have a complex electronic charge density distribution, ρ, surrounding them.

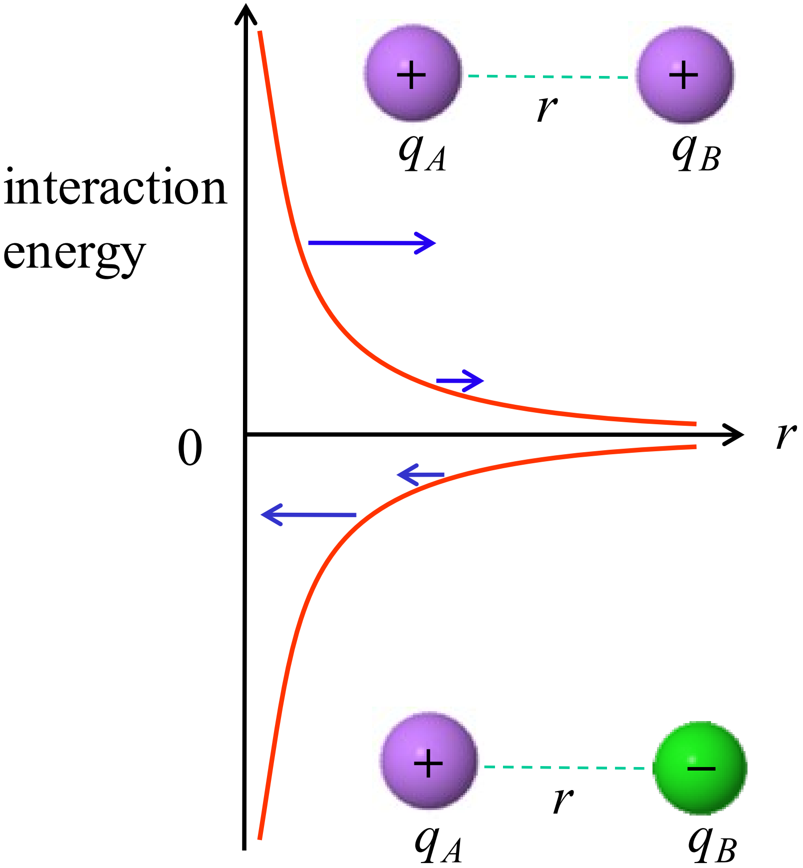
There are also complex polarization effects – i.e., where the interaction changes the molecules charge distribution.

Without doing an expensive quantum mechanical calculation, we don’t have access to the ‘true’ charge distribution or polarization effects.

Thus, we need to ‘inexpensively’ evaluate the electrostatic interactions!

Point Charge – Point Charge Interactions

For the interaction of simple ions, Coulomb's Law would be a very good approximation.



$$E(r) = \frac{1}{4\pi\varepsilon_o}\frac{q_A q_B}{r}$$

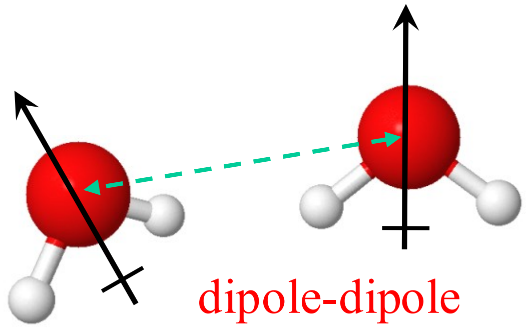
$$F(r) = -\frac{1}{4\pi\varepsilon_o}\frac{q_A q_B}{r^2}$$

Note that Coulomb's law would allow opposite charges to collapse on each other ($r = 0$).

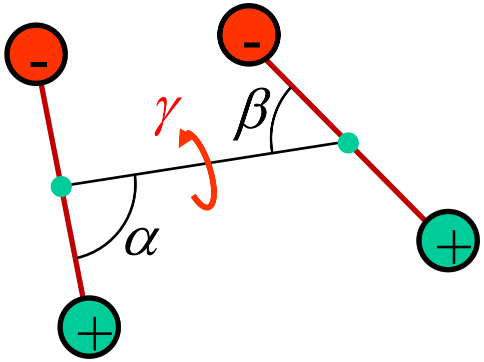
What prevents this from happening?

Dipole Interactions

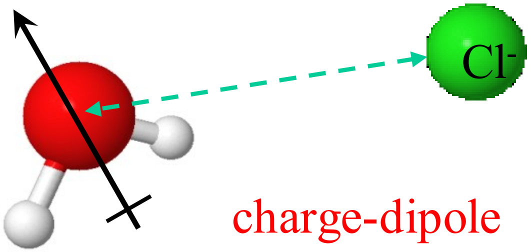
For molecules with a permanent dipole moment, electrostatic interactions can be evaluated as dipole-dipole and dipole-charge interaction terms:



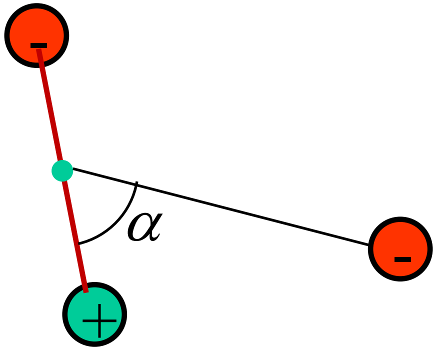
dipole-dipole



$$E_{dipole-dipole} = \frac{1}{4\pi\varepsilon_o}\frac{|\vec{\mu}_A||\vec{\mu}_B|}{r^3}\left(2\cos\alpha\cos\beta + \sin\alpha\sin\beta\cos\gamma\right)$$



charge-dipole



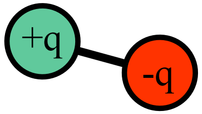
$$E_{\text{dipole-charge}} = \frac{1}{4\pi\varepsilon_o}\frac{q|\vec{\mu}|\cos\alpha}{r^2}$$

don't memorize

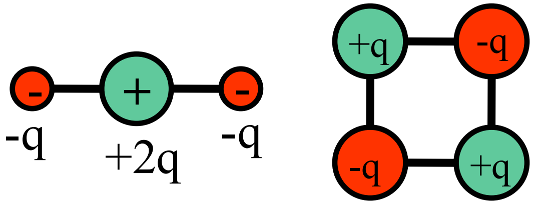
Higher Multipole Moments

There are higher order electric multipole moments:

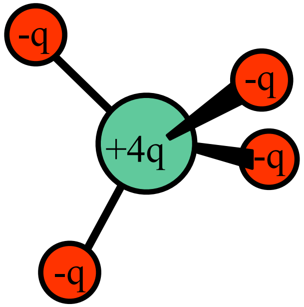
monopole



dipole



quadrupole



octupole

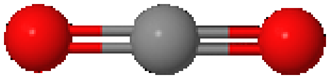
An electric dipole consists of two opposite charges, whose net charge is zero.

An electric quadrupole consists of point charges, whose net charge and dipole moment is zero.

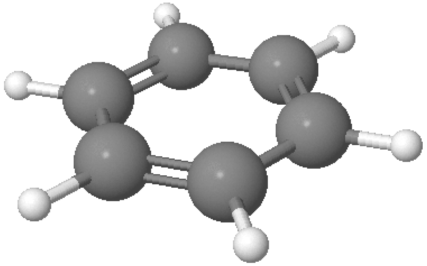
These are examples of point charge distributions whose lowest multipole moment is listed.

If a molecule does not have a net charge or dipole moment, the larger multipole moment interactions will typically dominate its interactions with other molecules.

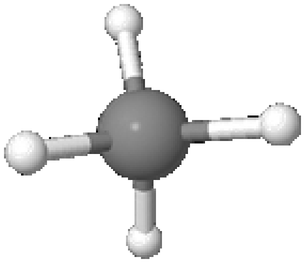
However, some molecules have no net charge, or dipole moment, but higher electrostatic moments.



CO_2
Quadrupole



benzene
Quadrupole (Why?)

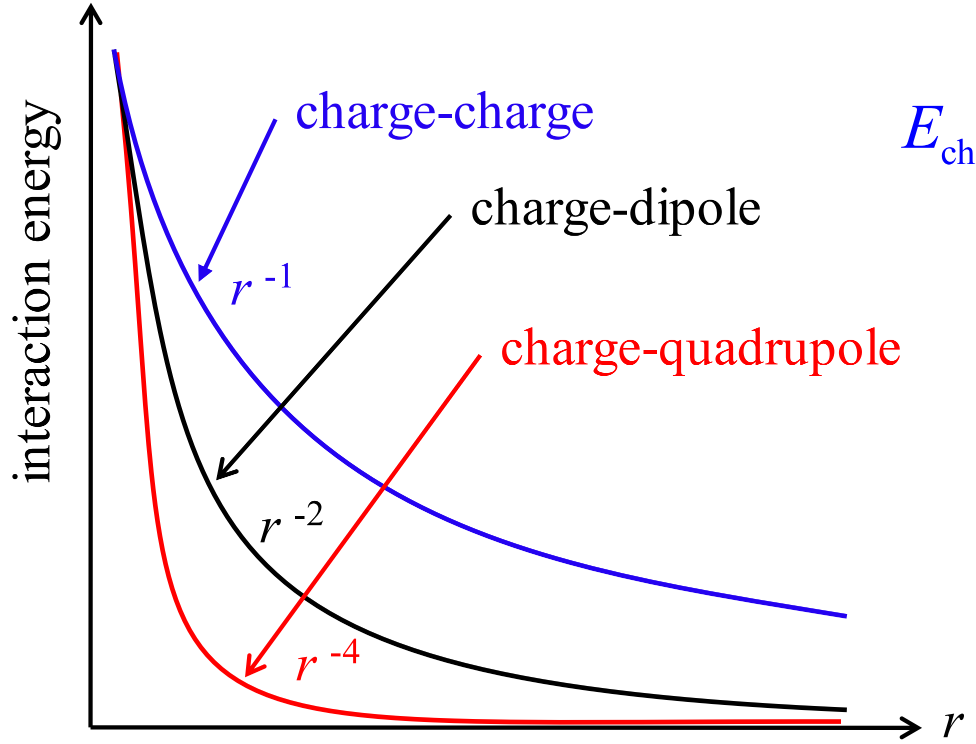


CH_4
Octupole

These higher multipole moment interactions between molecules are weak ~ same order of the dispersion interactions.

For certain systems, i.e. liquid CO_2, the higher multipole moments can be important.

Equations for the interaction energy of multipole-multipole interactions are available:



$$E_{\text{charge-charge}} = \frac{1}{4\pi\varepsilon_o}\frac{q_A q_B}{r} \propto \frac{1}{r}$$

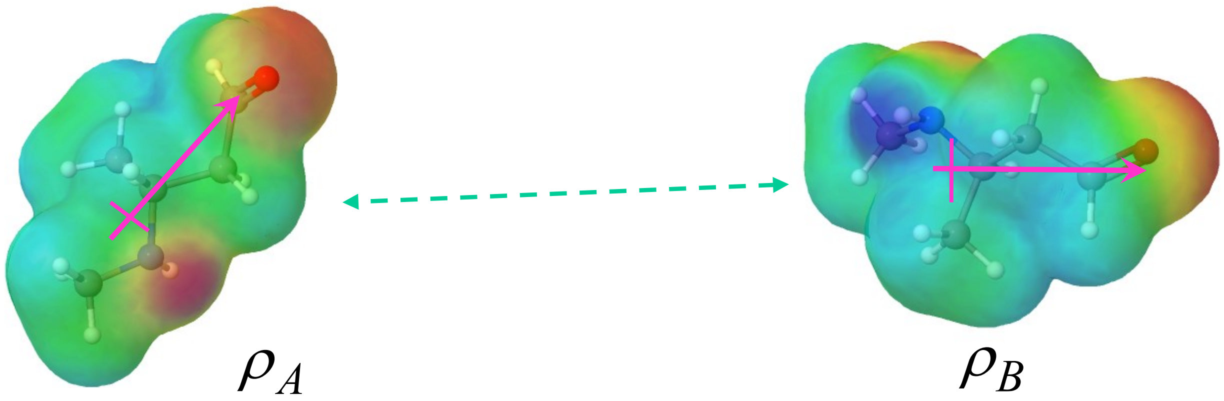
$$E_{\text{charge \_ dipole}} \propto \frac{1}{r^2}$$

$$E_{\text{dipole-dipole}} \propto \frac{1}{r^3}$$

$$E_{\text{quadrupole-quadrupole}} \propto \frac{1}{r^6}$$

Typically, the higher the multipole moment, the shorter range the electrostatic interaction energy. **(Why?)**

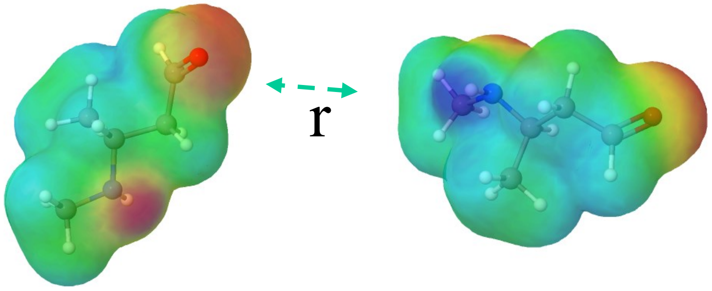
Electrostatic Interactions Between Molecules



If we know the net charge and dipole moment of two molecules, can we evaluate the total electrostatic interaction between them accurately?

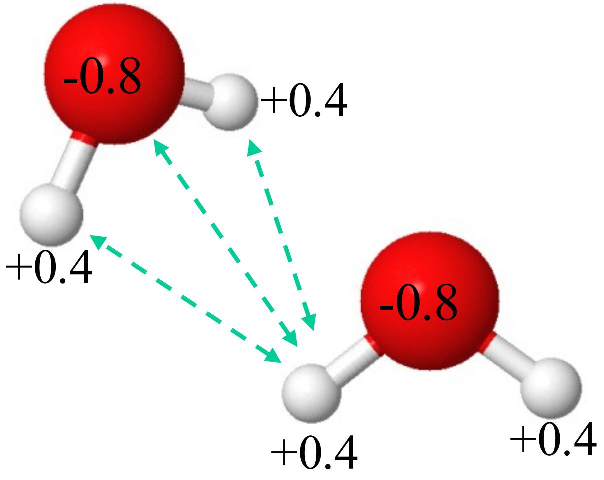
This avoids a complicated charge density calculation to evaluate the electrostatic interaction energy associated with the complex shape of the molecules.

This simple multipole representation breaks down at close distances, particularly for large molecules. Local distribution of charge is important in determining behavior.



Partial Atomic Charge Approximation

In most common force fields, electrostatic interactions by are handled by assigning fixed **partial atomic charges** at the nucleus position to each atom in the molecule.



Partial charges are not 'formal' integer charges, but rather are non-integer numbers.

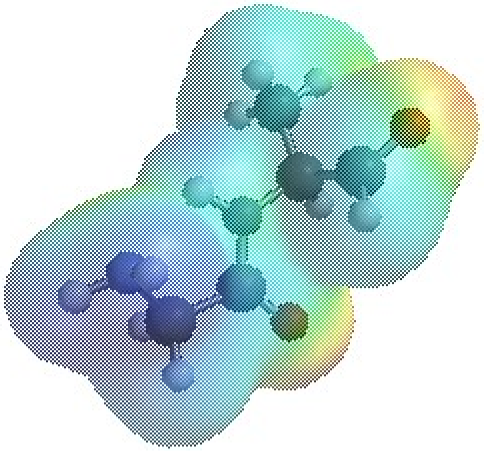
The sum of the electronic charge surrounding an atom and the nuclear charges are condensed into point charges centered on each of the atoms.

$$E_{elstat} = \frac{1}{4\pi\varepsilon_o}\sum \frac{q_i q_j}{r_{ij}}$$

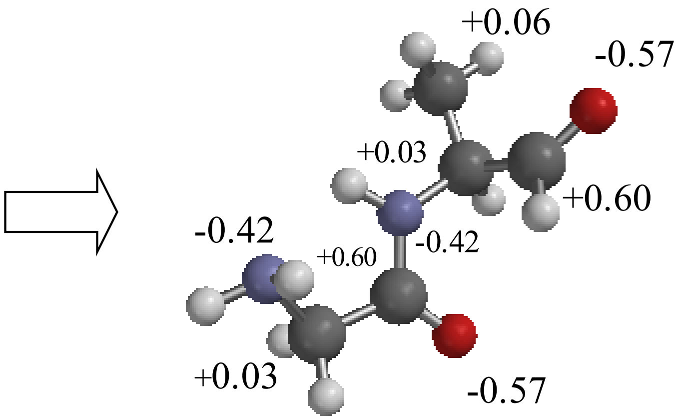
How are the charges assigned?

For small molecules, they can be chosen to reproduce the net charge and permanent dipole moment, for example.

For larger molecules, assigning charges is more problematic.



electronic charge density



Why?

Performing a quantum mechanical calculation to provide an estimate for the wave function is usually not a problem.

- *Many combinations of charges on atoms can reproduce the same multipole moments*
- *Partial atomic charge within a molecule is not a physical observable.*

There is no one “correct” way to assign the charges – there is some ambiguity.

Assigning Partial Atomic Charges

There are many ways to partition the charge within a molecule. These can generally be placed in one of the following categories:

i. Partitioning of orbital space

ii. Partitioning of the total electron density

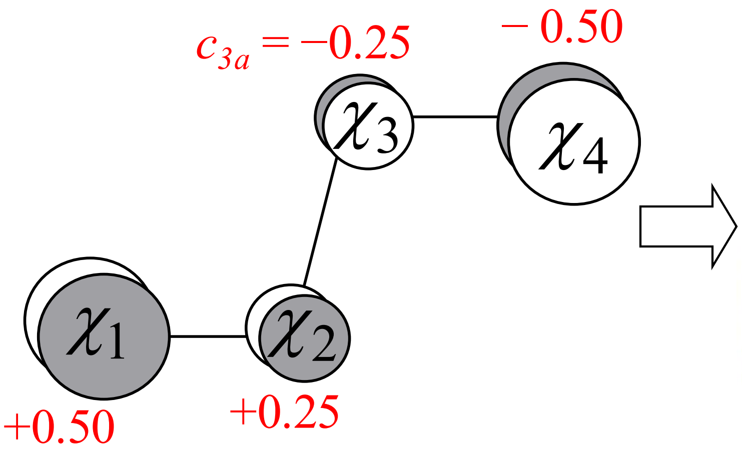
iii. Electrostatic potential (ESP) fitted charges

Category iii is the most physically relevant for molecular simulations, but we should be aware of the other categories because sometimes people use them.

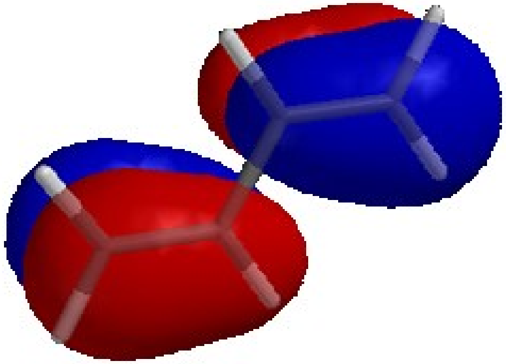
Mulliken Charges – An Orbital Based Partitioning Scheme

Mulliken charges are widely used in quantum chemistry since they are easy to calculate.

Mulliken charges take advantage of the fact that the “basis functions” used to construct the wave function are centered on individual nuclei.



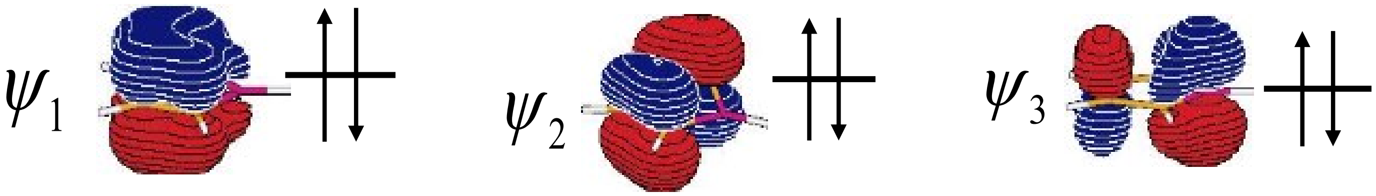
Atom-centered orbitals (χ_i)



Molecular orbital, ψ_a

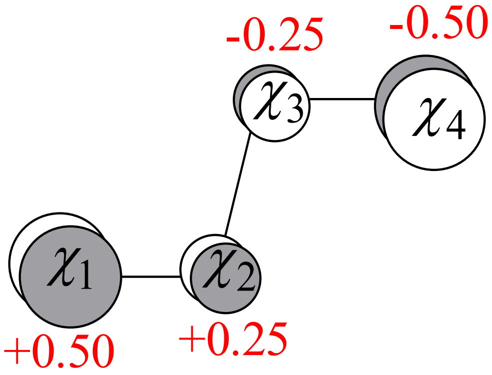
Coefficient for atomic orbital *i* in molecular orbital *a*.

$$\psi_a(\mathbf{r}) = \sum_{i=1}^{K} c_{ia}\chi_i(\mathbf{r})$$



K total number of atoms

$$\psi_a(\mathbf{r}) = \sum_{i=1}^{K} c_{ia}\chi_i(\mathbf{r})$$

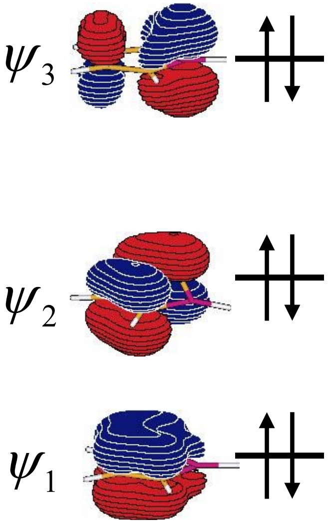


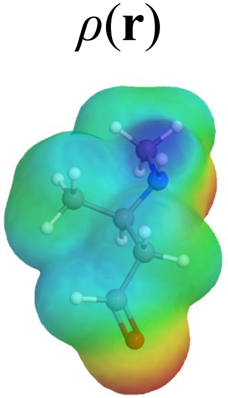
The total electron density, ρ, from a HF or DFT calculation is given as the sum of the square of each occupied molecular orbital (MO).

n total number of occupied MOs

$$\begin{aligned}\rho(\mathbf{r}) &= 2\sum_{a=1}^{n}\left|\psi_a(\mathbf{r})\right|^2\\ &= 2\sum_{a=1}^{n}\left(\sum_{i=1}^{K}c_{ia}\chi_i(\mathbf{r})\right)\left(\sum_{j=1}^{K}c_{ja}\chi_j(\mathbf{r})\right)\\ &= 2\sum_{a=1}^{n}\sum_{i=1}^{K}\sum_{j=1}^{K}c_{ia}c_{ja}\chi_i(\mathbf{r})\chi_j(\mathbf{r})\end{aligned}$$

composed of components from each atom





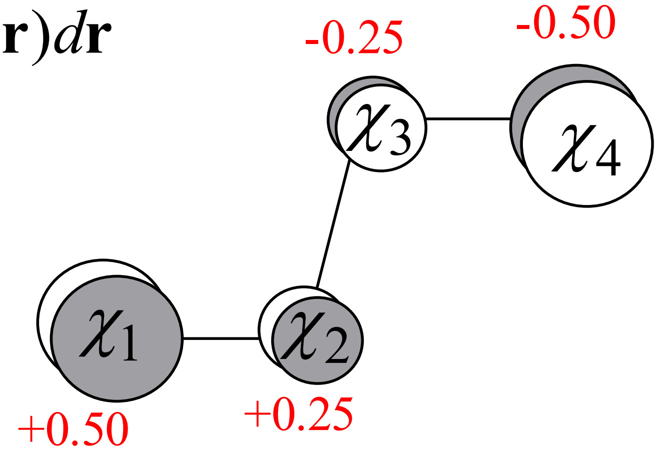
The integral of the total electron density over all space equals the total number of electrons in the molecule:

$$n_e = \int \rho(\mathbf{r})d\mathbf{r} = \int \left(2\sum_{a=1}^{n}\sum_{i=1}^{K}\sum_{j=1}^{K} c_{ia}c_{ja}\chi_i(\mathbf{r})\chi_j(\mathbf{r}) \right) d\mathbf{r}$$

In a Mulliken charge, the amount of charge assigned to an atom involves partitioning the above integral into portions that belong to each atomic center:

$$n_e^A = 2\sum_{a=1}^{n}\underbrace{\sum_{i=1}^{K}\sum_{j=1}^{K}}_{\text{One or both of } i \text{ and } j \text{ must be atom A}} c_{ia}c_{ja}\int \chi_i(\mathbf{r})\chi_j(\mathbf{r})d\mathbf{r}$$

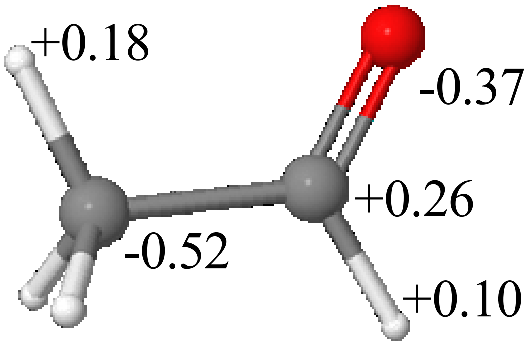
net electronic charge on atom A



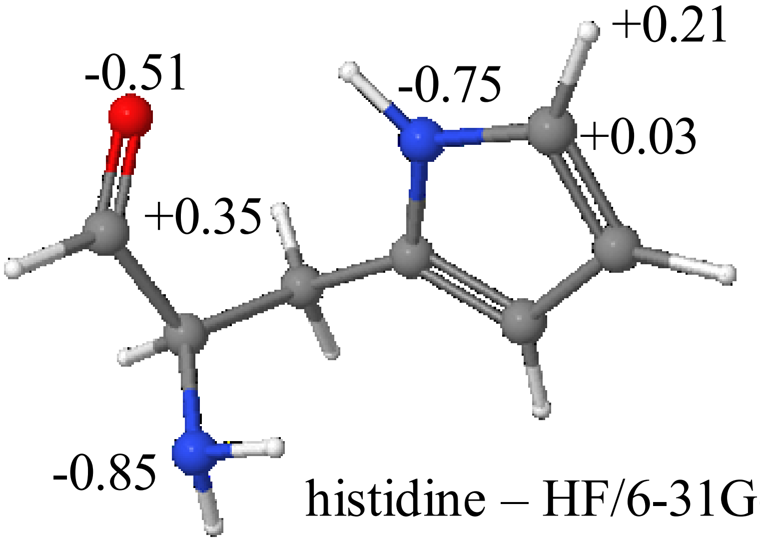
The final Mulliken charge is the sum of the above net electronic charge on the atom plus the nuclear charge:

$$q^A = n_e^A + Z^A$$

Mulliken Charges



acetaldehyde – B3LYP/6-31G(d)



histidine – HF/6-31G(d)

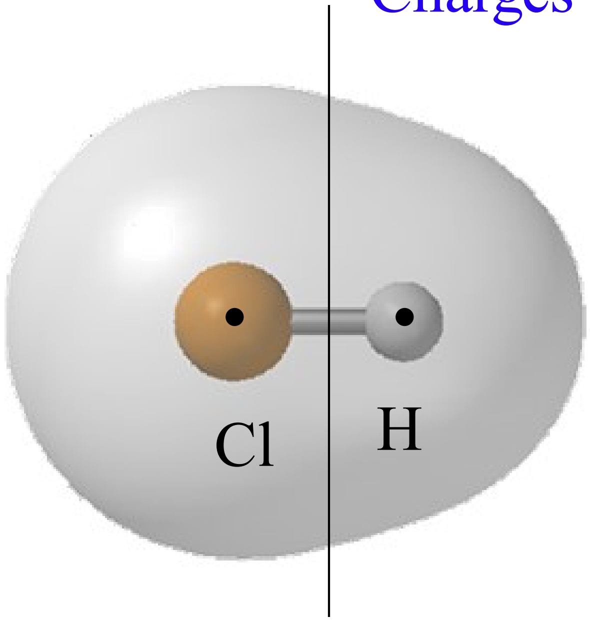
From a chemical point of view,

- Mulliken charges are generally reasonable;
- They assign orbital contributions equally to both atoms in a bond;
- One problem Mulliken charges have is that they are strongly basis set dependent.

acetaldehyde
HF calculation

	Mulliken Charge (e)		
basis set	C(Me)	C(CO)	O
STO-3G	-0.21	0.15	-0.21
3-21G	-0.52	0.34	-0.73
6-311G(d,p)	-0.34	0.29	-0.49

Charges Based on the Electron Density



The simplest way to partition the total charge is to assign the charge to the nucleus that is closest to it.

Consider the case of HCl to the left, what might be considered "wrong" with this partitioning?

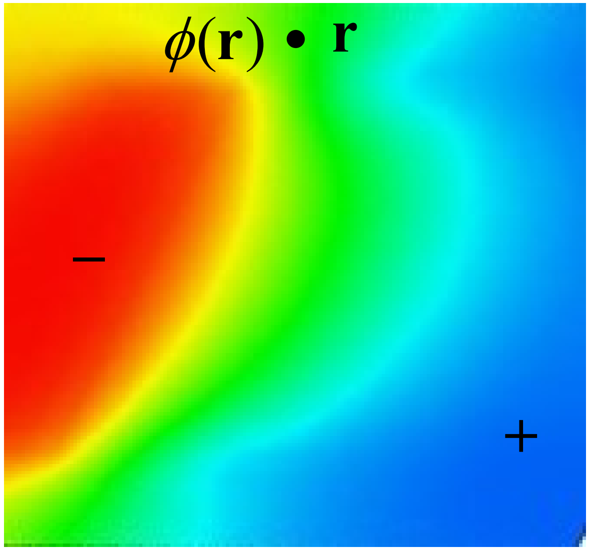
The so-called Bader charges based on the theory of "atoms in molecules" (AIM) involves partitioning of the electron density in a systematic manner.

Electrostatic Potential (ESP) Fitted Charges

The most appropriate charges to be used for molecular simulation are so-called electrostatic potential fitted charges.

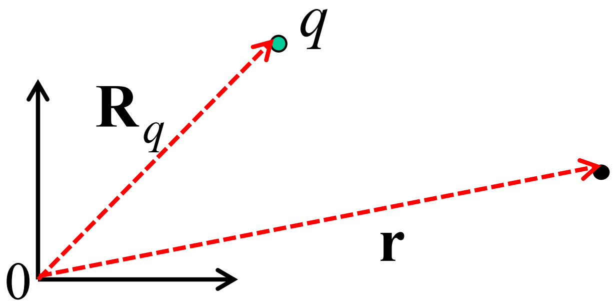
What is the electrostatic potential?

The electrostatic potential at a point in space **r**, $\phi(\mathbf{r})$, is defined as the electrostatic potential energy per unit of charge at that point.



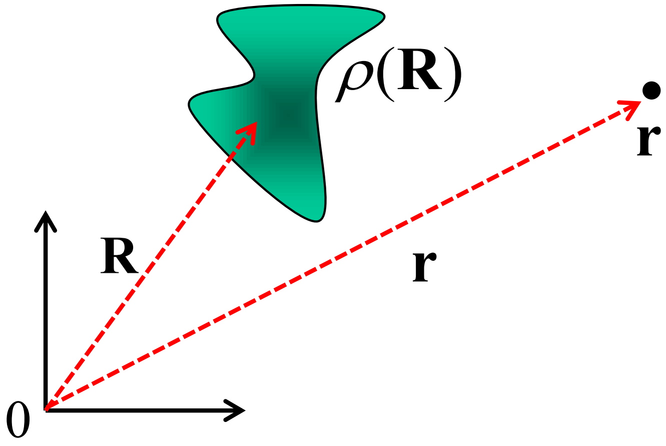
i.e., the energy required to bring a unit charge (+1) from an infinite distance away to the point **r**.

The electrostatic potential at point 'r' due to a point charge 'q' is defined as:



$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_o}\frac{q}{\left|\mathbf{r}-\mathbf{R}_q\right|}$$

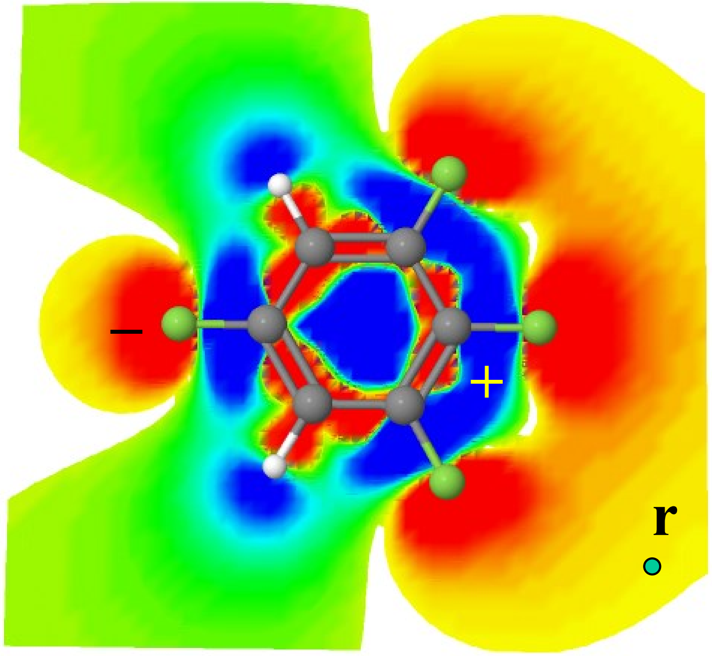
For continuous charge distribution, ρ, the electrostatic potential at the point '$\mathbf{r}$' due to this charge distribution is defined as:

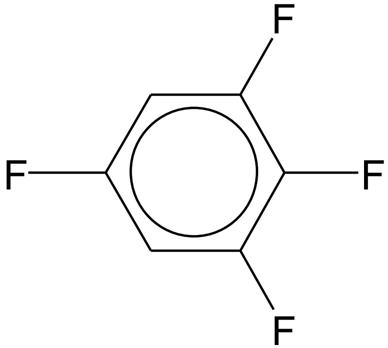


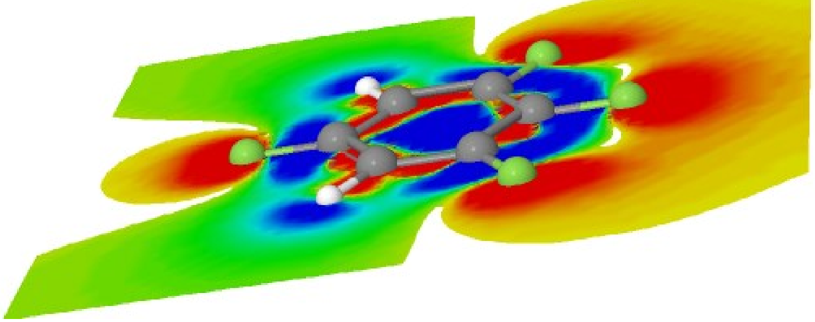
$$\phi(\mathbf{r}) = \frac{1}{4\pi\varepsilon_o}\int\frac{\rho(\mathbf{R})}{\left|\mathbf{r}-\mathbf{R}\right|}d\mathbf{R}$$

integration over the space of the charge distribution

Shown below is the electrostatic potential of fluorobenzene ($C_6F_4H_2$) on the plane of the ring.







The electrostatic potential at a point due to a molecule is given by (in atomic units):

$$\phi(\mathbf{r}) = -\int \frac{\rho(\mathbf{r}')d\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|} + \sum_{I=1}^{N} \frac{Z_I}{|\mathbf{r}-\mathbf{R}_I|}$$

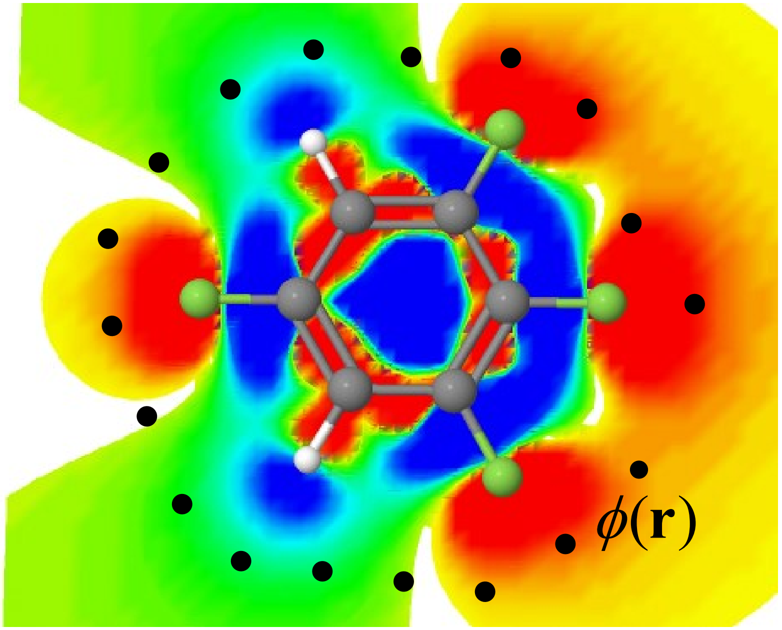
ρ is the electron density of the molecule

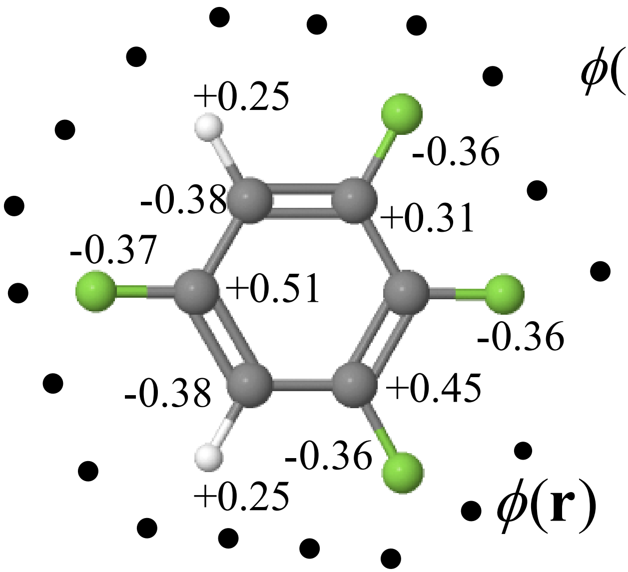
$\mathbf{R}_I$ is the position of the I^{th} nucleus

Z_I is the charge of the I^{th} nucleus

Electrostatic Potential (ESP) Fitted Charges

Consider the electrostatic potential at a point, **r**, due to a molecule:



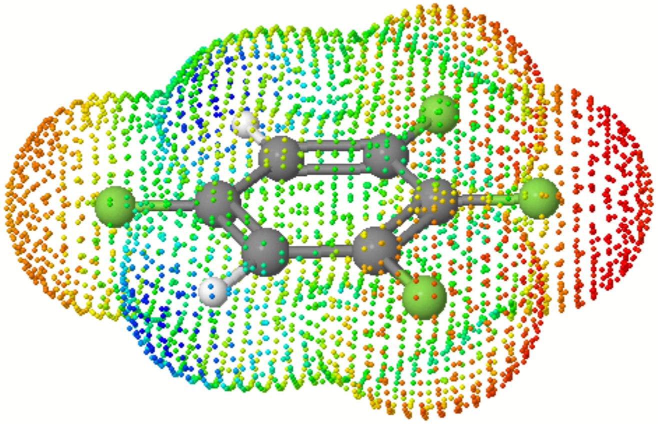


$$\phi(\mathbf{r}) = \sum_{I=1}^{N} \frac{\delta_I}{|\mathbf{r} - \mathbf{R}_I|}$$

How do we find a set of partial atomic charges at the nuclear positions, δ_I, that reproduces the electrostatic potential of the wave function:

- Choose a grid of points around the molecule in a systematic manner;
- Assign unknown charges as variables δ_I at the nuclear positions;
- Apply a least squares fit for the values of the point charges such that they reproduce the electrostatic potential of the wave function on the grid of points;
- Apply constraints during the fit process (net charge is zero, constrain charges to less than 1, etc.).

Details of the Charge Fitting Process

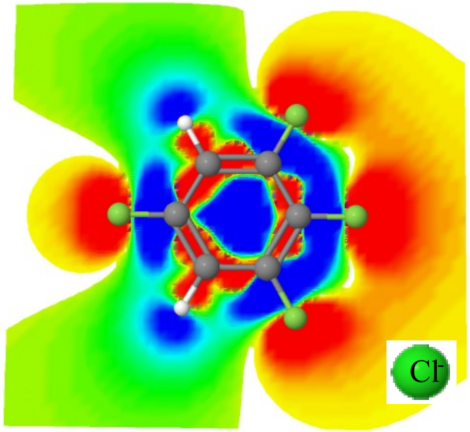


1. Usually the fit points of the electrostatic potential (ESP) are taken on the van der Waals surface of the molecule and further because this is the closest distance at which molecules interact with each other during a simulation.

2. The atomic point charges are constrained to add up to the total net charge of the molecule.

3. The point charges should reproduce the molecule’s dipole moment.

4. Other constraints may be applied to keep the point charges > 1*e*



Within the fixed point charge model, ESP point charges are generally the most appropriate charges to use for molecular simulation.

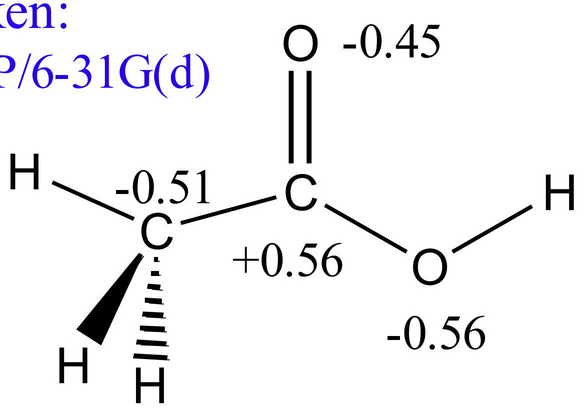
Why?

Many ESP charge fitting schemes have been developed

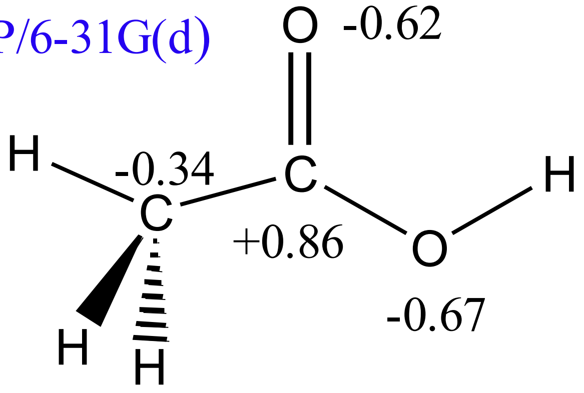
Some common methods of determining ESP charges are:

- **M**erz-Singh-**K**ollman (MK) *J. Comp. Chem.* **1990**, *11,* 431.
- **CHelpG** methods. (*J. Comp. Chem.* **1990**, *11,* 361)
- **RESP** (*J. Phys. Chem.* **1993**, *97*, 10269)
- **REPEAT** (*J. Chem. Theor.* Comp. **2009**, *5*, 2866) - for solids (Dr. Woo)
- **DDEC/c3** (*J. Chem. Theor.* Comp. **2012**, *8*, 2844; **2014**, *10*, 5377) (T. Manz)

Mulliken:
B3LYP/6-31G(d)

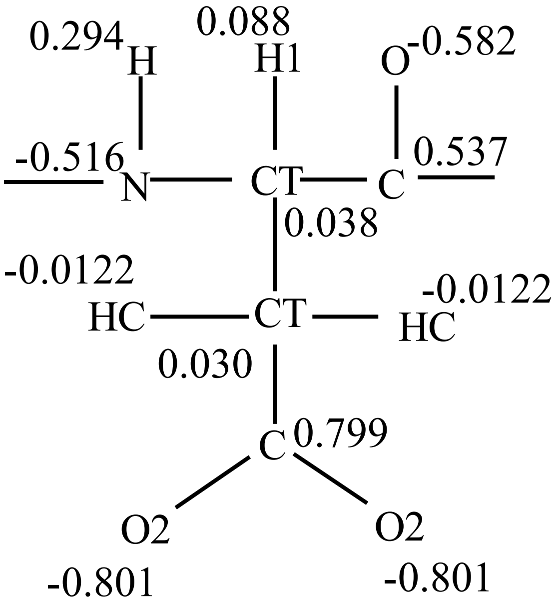


CHELPG:
B3LYP/6-31G(d)

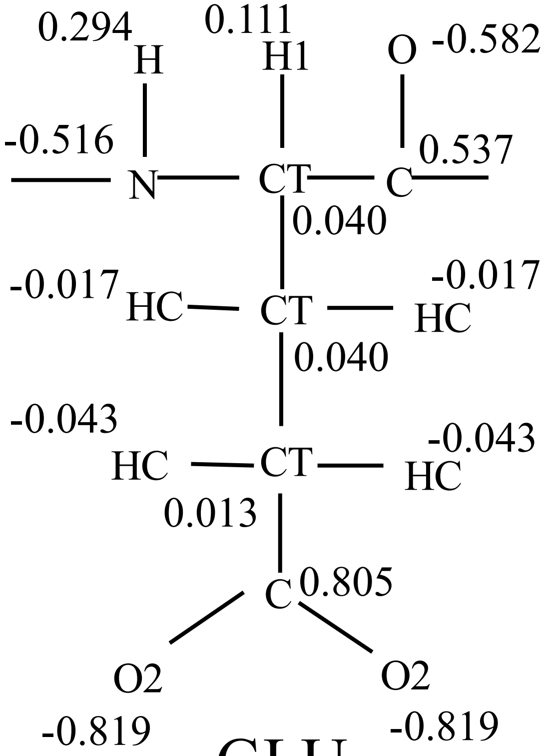


ESP charges are much less basis set dependent compared to Mulliken charges.

AMBER, a popular force field for simulating proteins and nucleic acids, uses charges derived from HF/6-31G* calculations of the individual nucleotides and amino acids.



ASP



GLU

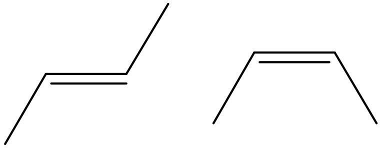
AMBER95 force field: Kollman *JACS*, **1995**, *117*, 5179.

As with the van der Waals pair potentials, intra-molecular point charge interactions are only included if they are 1-5 interactions (or 1-4) or greater.

Limitations of the Point Charge Model

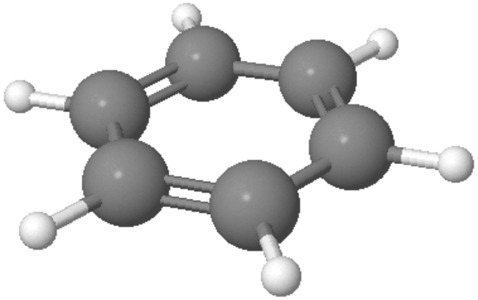
The fixed point charge model is the most common way of modeling electrostatic interactions between molecules, within molecular mechanics methods.
There are, as one may expect, limitations:

- Polarization of the charge density is not captured
- Conformational charges may change the charge distribution



- Charge transfer effects cannot be accounted for
- Sometimes, charges centered on the nuclei alone cannot reproduce the electrostatic multipole moments, e.g., in the case of benzene. However, dummy atoms (additional centers of charge not associated with nuclei) can be added

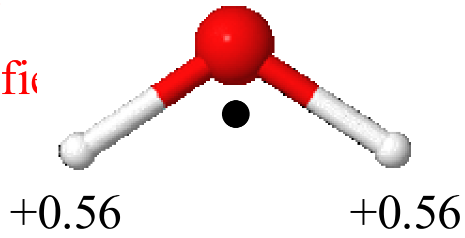
Benzene has a quadrupole moment perpendicular to the ring



Alternative Electrostatic Models

Some force fields place point charges off of the nuclear centers to form "dummy atoms"

TIP4P/2005
water force fi

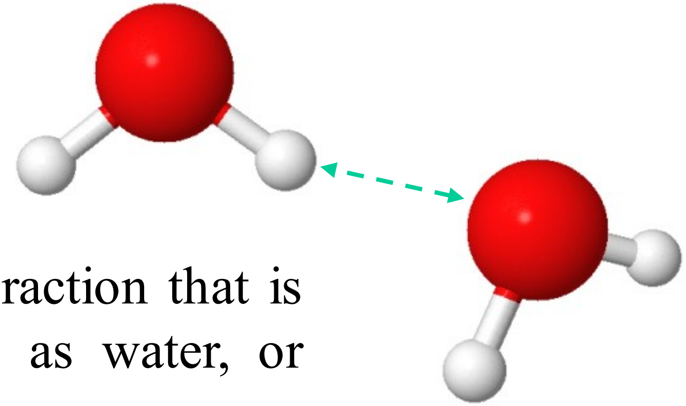


There have been numerous attempts to develop 'polarizable' force fields where the charge distribution can 'react' to its environment.

- Polarizable force fields for specific systems have been successful, i.e. polarizable water force fields.
- This is an intense area of modern research, particularly for the simulation of biological systems, but no good general approach has been developed.

Nov. 2007 issue of J. of Chemical Theory and Computation:
"Special Issue on Polarization"

Treatment of Hydrogen Bonds



Hydrogen bonds are a type of non-bonded interaction that is particularly important for some systems such as water, or biological molecules.

Many force fields do not treat hydrogen bonds explicitly. In other words, they are treated during the parameterization of the regular van der Waals + point charges.

e.g. AMBER for nucleic and amino acids

The pair potentials used in force fields are not orientation dependent, but hydrogen bonds are.

For this reason, some force fields include explicit directional hydrogen bond potentials, but this is not particularly common.

Putting it in Context: Steps in Running an MD Simulation

1. Preparation and setup

define geometry of system

set up force field

adjust MD settings

2. Equilibration

start MD simulation

look at output

monitor properties

3. Production run

continue MD simulation

acquire data

4. Analysis

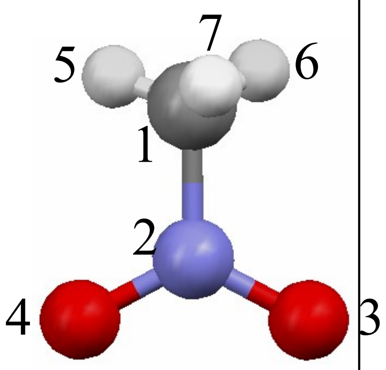
property averaging

data analysis

Every MD Software Program has a Different Force Field Set Up

- Many programs have the force field set up automatically from the input geometry
- In DL_POLY all interactions are specified in the FIELD file

```
480 molecule nitromethane surface slab
UNITS  kJ
molecular types 1
nitromethane
NUMMOLS 480
ATOMS  7
C      12.0110  -0.3053     1   0   1
N      14.0067   0.8206     1   0   1
O      15.9994  -0.4774     1   0   1
O      15.9994  -0.4774     1   0   1
H       1.0079   0.1465     1   0   1
H       1.0079   0.1465     1   0   1
H       1.0079   0.1465     1   0   1
BONDS      6
mors    1   2  251.04     1.50      2.01
mors    1   5  426.77     1.09      1.89
mors    1   6  426.77     1.09      1.89
mors    1   7  426.77     1.09      1.89
mors    2   3  390.37     1.23      2.46
mors    2   4  390.37     1.23      2.46
ANGLES      9
harm     1    2    3  589.04      117.04
harm     1    2    4  589.04      117.04
harm     2    1    5  449.63      107.56
harm     2    1    6  449.63      107.56
```



Parameterization of Force Fields

A force field can contain a very large number of parameters, even for a limited class of molecules.

> e.g. AMBER95 force field has about 400 angle bending parameters – this is only to treat nucleic and amino acids.

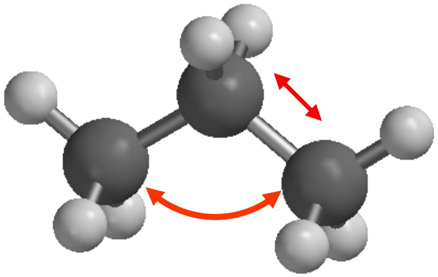
Parameterization and force field development is an Art and Science.

Even development of very specific force fields (i.e. to treat liquid water only) is not straightforward.

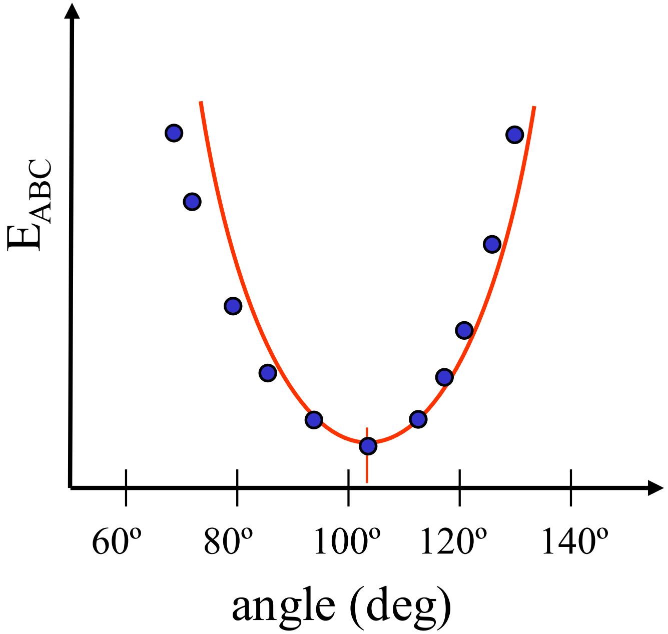
Development of general force fields, e.g. AMBER, has taken decades of research.

Parameterization Strategies

- Geometric parameters and force constants can be fit to high level quantum chemical (QC) calculations (see Assignment 2).



$$E_{ABC} = k_{ABC}(\theta_{ABC} - \theta^{eq}_{ABC})^2$$

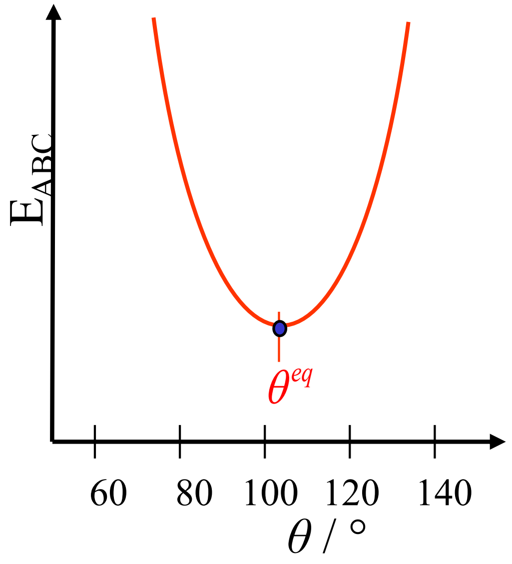


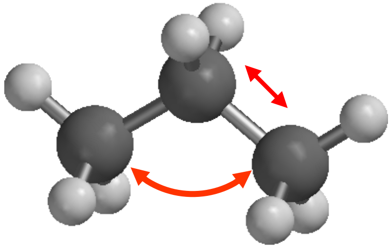
- Charges are typically fit to QC calculations.

- Most force fields use QC calculations for some of the parameterization, but few use exclusively QC calculations.
  - ab initio water force field - intermolecular *(PNAS*, **2006**, *103*, 8613)

Fitting to Experimental Molecular Properties

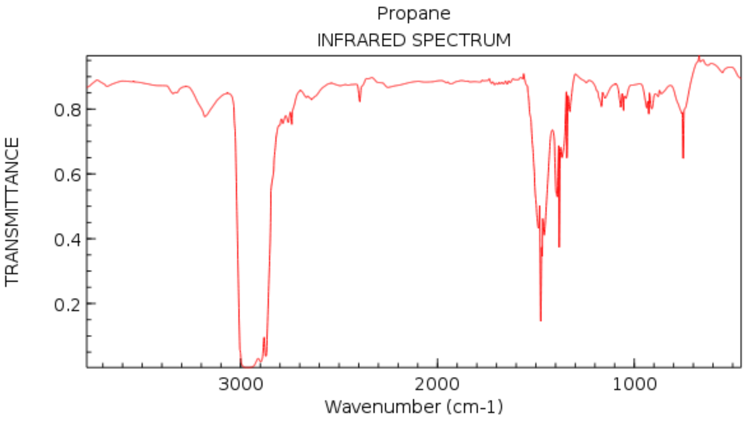
- Some force field parameters can be directly related to experimental measurements. For example, equilibrium bond distance and angle parameters can be extracted from X-ray crystallography data.





$$E_{ABC} = k_{ABC}(\theta_{ABC} - \theta_{ABC}^{eq})^2$$

- Some force constants can also be fit to vibration spectra, but this can be rather complicated.



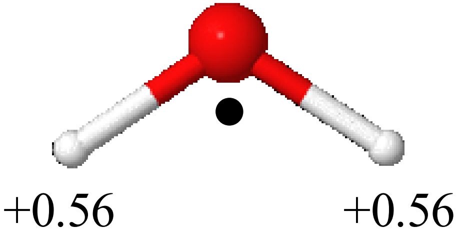
NIST Chemistry WebBook (http://webbook.nist.gov/chemistry)

Fitting to Experimental Bulk Properties

Often times, force field parameters are fit to reproduce bulk properties, such as the density of a liquid, or the melting temperature.

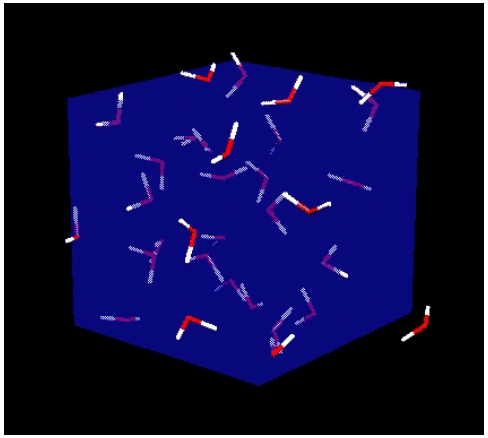
Example: TIP4P/2005 water force field

J. Chem. Phys. **2005**, *123*, 234505



- H-O distance: 0.9572 Å (fixed to expt. value)
- H-O-H angle: 104.52° (fixed to expt. value)
- Single Lennard-Jones vdW potential on O (none on H)
- + charges on H, negative charge below the O center

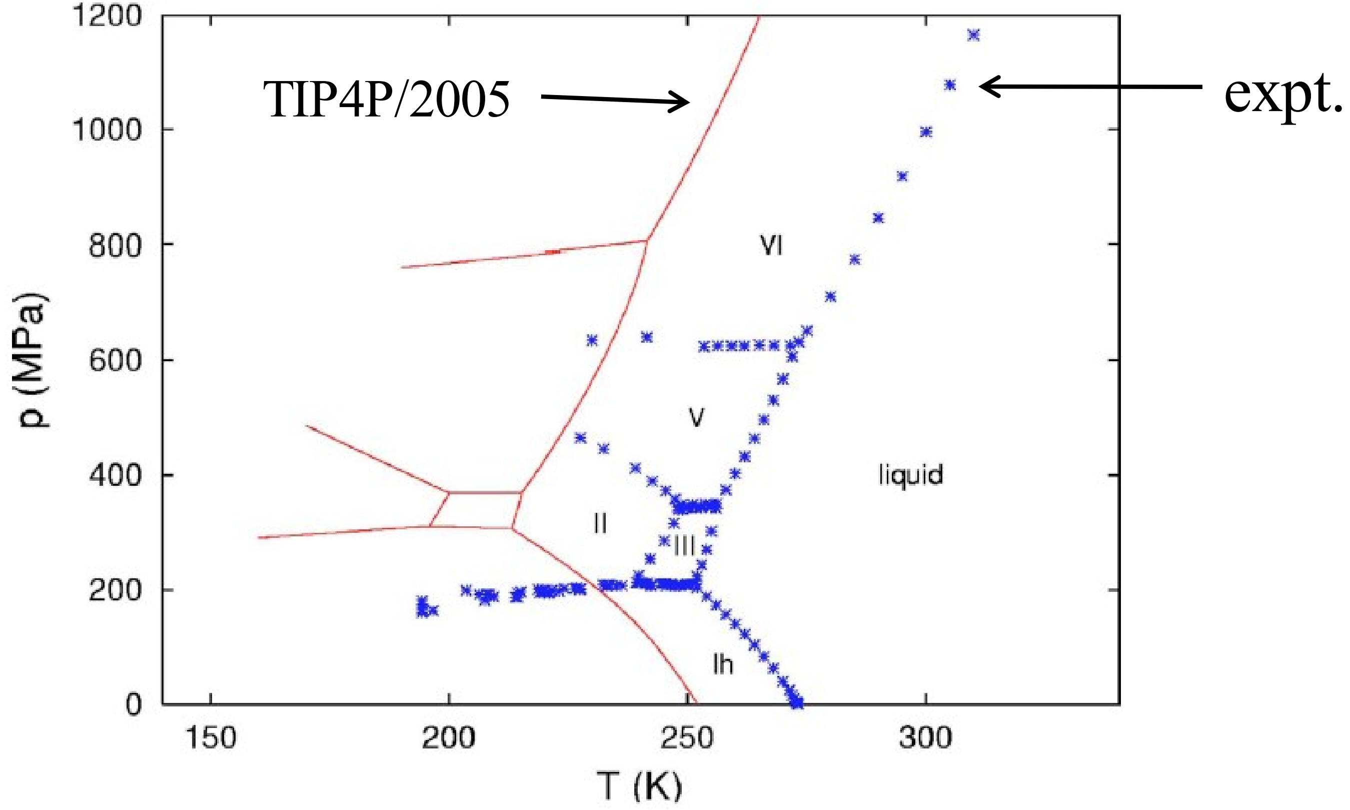
Property at room T (298 K), 1 atm	**TIP4P/2005**	**Experiment**
Density (g·cm^{-3})	0.9979	0.9971
Isothermal compressibility (10^5 MPa^{-1})	46.5	45.8
Thermal expansion coefficient (K^{-1})	28	25.6
Heat capacity at constant P ($\text{cal·mol}^{-1}\text{·K}^{-1}$)	18.9	18.0
Heat of vaporization (kcal·mol^{-1})	10.89	10.52
Self diffusion coefficient ($\text{m}^2\text{·s}^{-1}$)	2.08	2.27



property	TIP4P/2005	Experiment
Melting temperature (K)	252.1	273.15
Liquid density at melting T (g·cm^{-3})	0.993	0.999
Solid density at melting T (g·cm^{-3})	0.921	0.917

The TIP4P/2005 can reproduce the experimental phase diagram of water / ice fairly accurately

Phase Diagram of Water



Parameterization of Force Fields

Fitting to bulk properties can result in 'unphysical' atomic parameters

e.g. a vdW radius that is much larger than experiment

! However, parameterization can 'patch-up' the deficiencies in our force field.

Recall we are effectively replacing solving the Schrödinger equation to obtain the system energy with a simple sum of molecular mechanics energy terms:

$$\int \Psi^* \hat{H} \Psi d\tau = E(\mathbf{R}) \approx E_{bonds} + E_{angles} + \ldots + E_{non-bonded}$$

We can also patch up quantum mechanical effects with additional parameters.

In other words, fitting a force field to bulk experimental parameters can make up for shortcoming in the fundamental approximations in molecular mechanics.

It also means that caution should be used when taking parameters from one force field for use in another force field.

Transferability and Accuracy

A transferable force field is one which can be used for many different molecules and mixtures.

In general, the more transferable a force field for general use, the less accurate it is for any specific material.

Specialized force fields can be very accurate. e.g. TIP4P/2005 for water.

- But will this the water force field perform just as good if we add Na^+ and Cl^- ions?

For many classes of important compounds, reasonably accurate force fields have been developed.

'AMBER' for bio-macromolecules (peptides and nucleic acids)

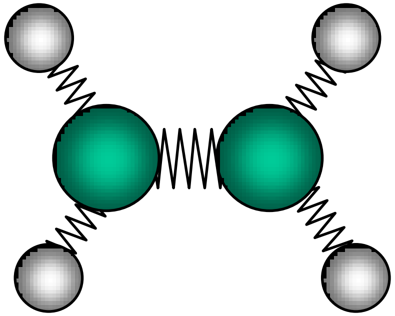
Force fields are only reliable for molecules for which they are parameterized.

'AMBER' does not perform well for carbohydrates

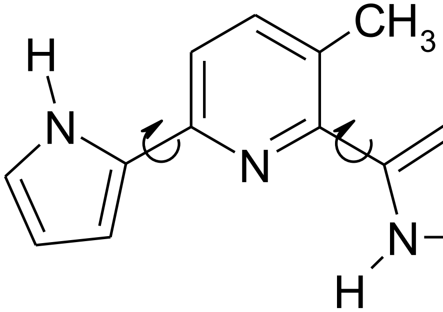
There currently is not a single force field that is robust or accurate over a large range of **inorganic** complexes. Transition metals are problematic.

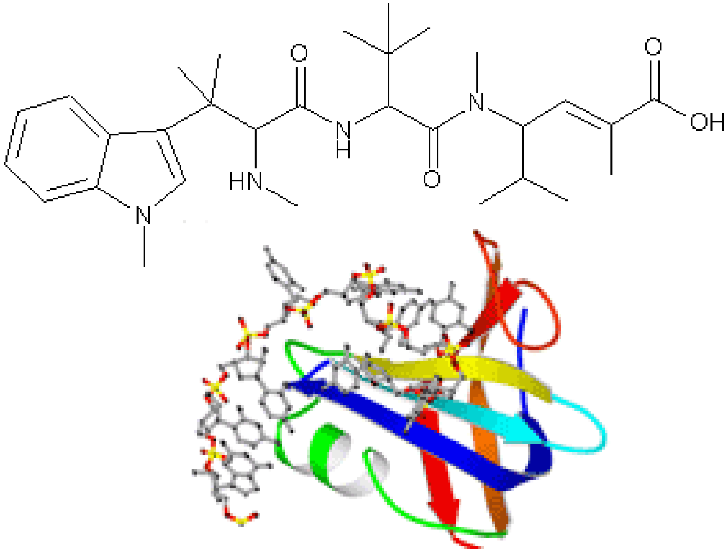
Molecular Mechanics - Limits

- MM calculations assume a particular bond connectivity.
- In general chemical, **reactions cannot be treated**. (covalent bond breaking and forming processes)

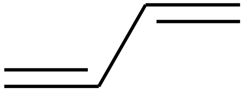


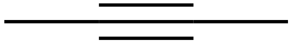
- Neither the reaction thermodynamics nor reaction barriers can be treated reliably.
- In general, with MM methods only **conformations** can be compared in energy.

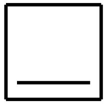


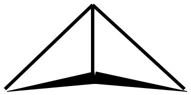


Can we use molecular mechanics to compare the energies of C_4H_6?


butadiene


2-butyne


cyclobutene


bicyclo[1.1.0]butane

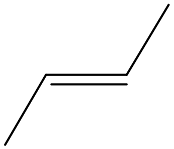
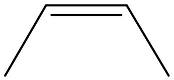
	Relative energy (kcal/mol)		
Isomer	**Experiment**	**B3LYP/6-31G(d)**	**MM (UFF)**
butadiene	0	0	0.0
2-butyne	9	8.4	-4.2
cyclobutene	11	11.7	88.8
bicyclo[1.1.0]butane	26	27.5	562.8

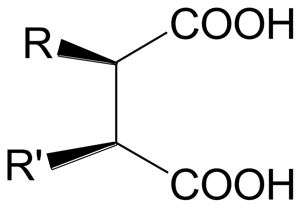
why is the order of the stability switched?

why so large?

What is the fundamental reason why we cannot use molecular mechanics to compare the energies of these isomers?

Can we use molecular mechanics to compare the energies of these isomers?


vs



vs