Force Fields Crash Course

SFB 677

Bernd Hartke

Theoretical Chemistry
Institute for Physical Chemistry
Christian-Albrechts-University
Olshausenstraße 40
D-24098 Kiel, Germany

`http://ravel.pctc.uni-kiel.de/`

`e-mail: hartke@pctc.uni-kiel.de`

Contents

- the force field (FF) concept
- various FF types:
  - simple (bio)chemistry FFs
  - FFs for semiconductors, metals, bulk, surfaces,…
  - polarizability
  - "reactive" FFs
- brief glimpse on applications
- which software to use

NOT contained:

- detailed theory & algorithms for
  - geometry optimization (local or global)
  - molecular dynamics (classical or quantum)
- calculation of observables from MD, e.g.: diffusion coefficients, pair/angle distribution functions, IR/Raman spectra, heat capacity, entropy, free energy,…

recommended literature

books:

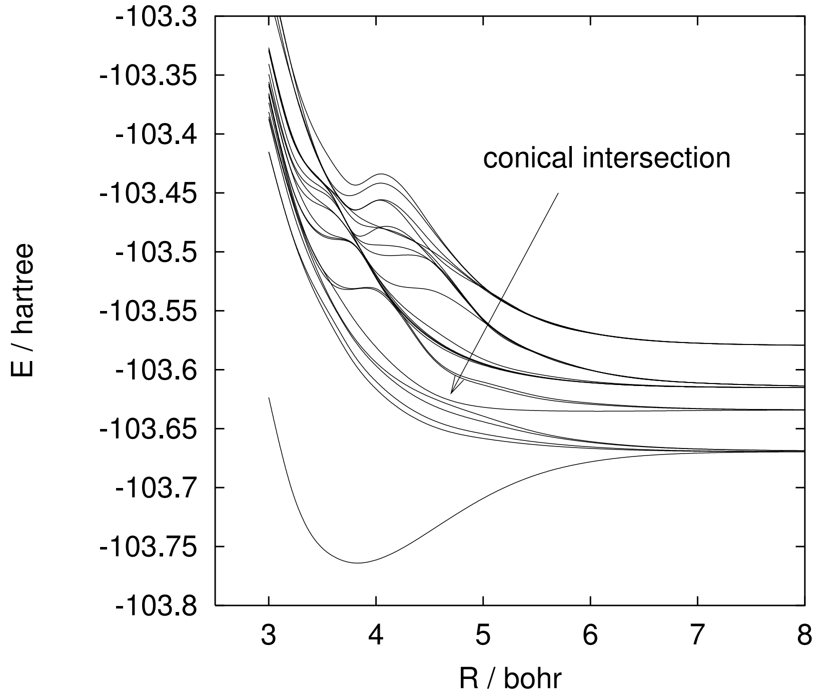
- A. R. Leach: "Molecular Modeling – Principles and Applications", 2nd ed., Pearson Education, 2001.
- C. C. Cramer: "Computational Chemistry – Theories and Models", 2nd ed., Wiley, 2004.
- F. Jensen: "Introduction to Computational Chemistry", 2nd ed., Wiley, 2007.

review papers:

- P. Ballone, Entropy 16 (2014) 322; DOI: 10.3390/e16010322:
  general FF introduction-like review
- A. D. MacKerell, Jr., J. Comput. Chem. 25 (2004) 1584; DOI: 10.1002/jcc.20082:
  biochemistry force fields
- B. L. Foley, M. B. Tessier and R. J. Woods, WIREs Comput. Mol. Sci. 2 (2012) 652; DOI: 10.1002/wcms.89:
  carbohydrate force fields
- C. M. Baker, WIREs Comput. Mol. Sci. 5 (2015) 241; DOI: 10.1002/wcms.1215:
  polarizable force fields
- H. Heinz, T.-J. Lin, R. K. Mishra and F. S. Emami, Langmuir 29 (2013) 1754; DOI: 10.1021/la3038846:
  the interface force field
- K. Farah, F. Müller-Plathe and M. C. Böhm, ChemPhysChem 13 (2012) 1127; DOI: 10.1002/cphc.201100681:
  reactive force fields

The FF concept (1/4)

Standard quantum-chemistry workflow:



Electrons are > 2000 times heavier than nuclei
⇒ *Born-Oppenheimer separation:*

1. *electronic part:* repeat for many i:
    - fix nuclei in space, at positions $\vec{R}_i$
    - calculate electronic wavefunctions and energies E_{el}

    Result: pointwise approximation to $E_{el}(\vec{R})$
2. *nuclear part:* calculate nuclear dynamics, with potential energy $V(\vec{R}) = E_{el}(\vec{R})$

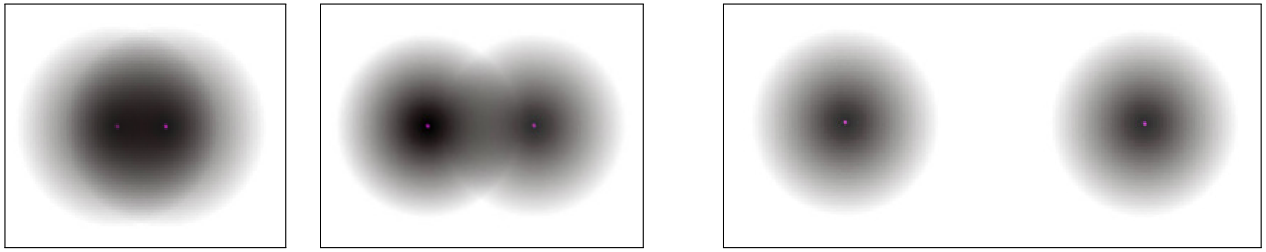
Note: This simplifies the calculations, but is *not* an approximation if many coupled $V(\vec{R})$ are used.

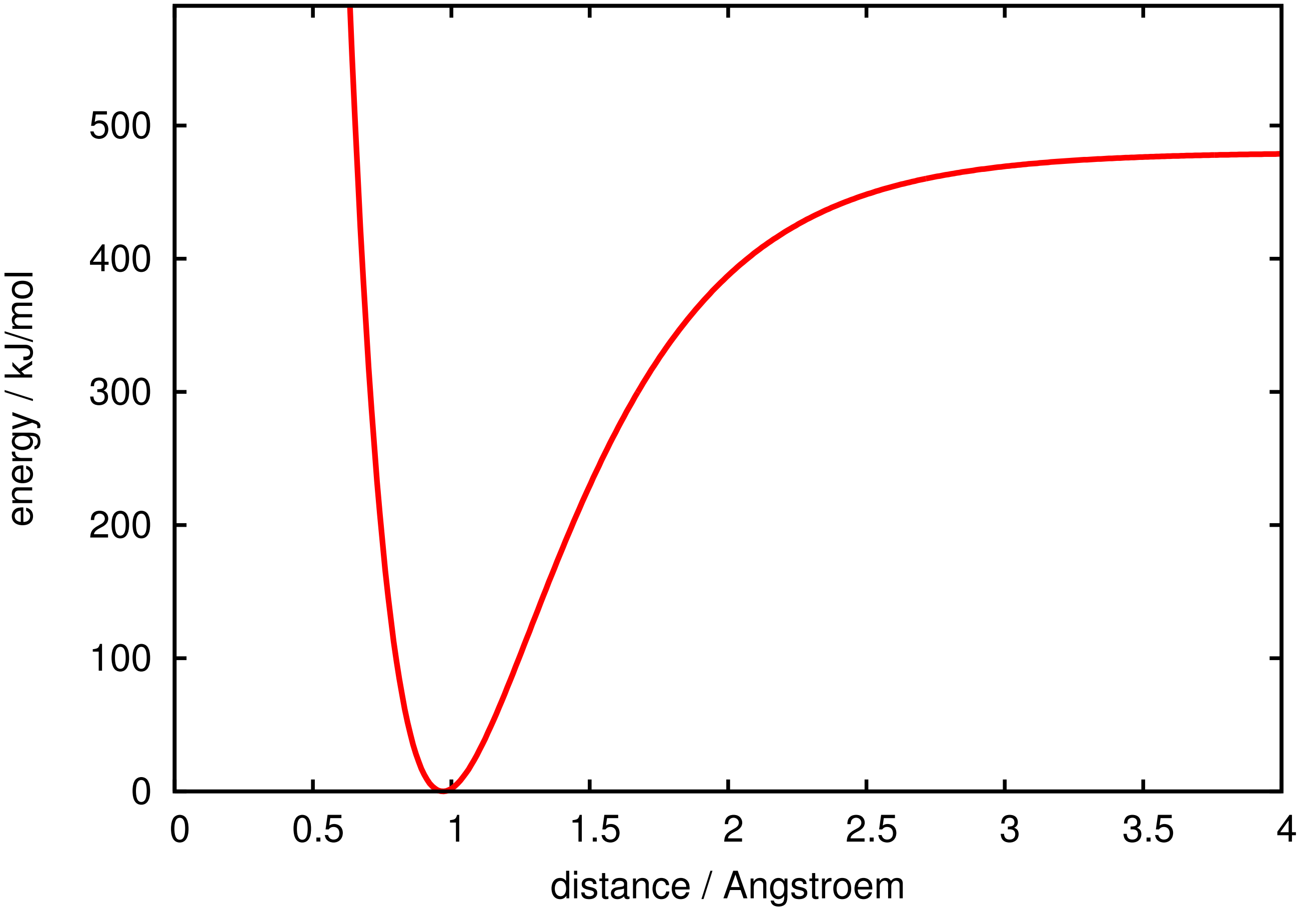
FF idea: step 1 is expensive ⇒ skip it! Obtain function(s) $V(\vec{R})$ from e.g.:

- educated guesses
- interpolation between / fitting to a few computed $E_{el}(\vec{R})$ values

The FF concept (2/4)

“coarse-graining away” the electrons





The FF concept (3/4)

- **things to learn:**
  - there is nothing classical-mechanical about FFs; they simply are a short-cut to solutions of the electronic Schrödinger equation
  - eigenvalues of the electronic Schrödinger equation are unique, deterministic, and well-behaved $\Rightarrow$ can be modelled by simple mathematical functions
  - $V(\vec{R}) = E_{el}(\vec{R})$ is *fully* determined by which nuclei are in which positions $\vec{R}$ (this includes information on how many electrons there are), i.e., by the usual *input* to a quantum-chemical calculation

$\Rightarrow$ in principle, FFs can handle *everything* that is accessible to quantum chemistry! (in principle, at any level, including fullCI/CBS, not just HF or DFT)

For example, FFs can (and have been) used for

- complicated transition-metal coordination complexes
- metal-organic frameworks (MOFs)
- isolated molecules, bulk, surfaces
- liquids, solids, interfaces, films, layers, clusters
- isolators, semiconductors, metals

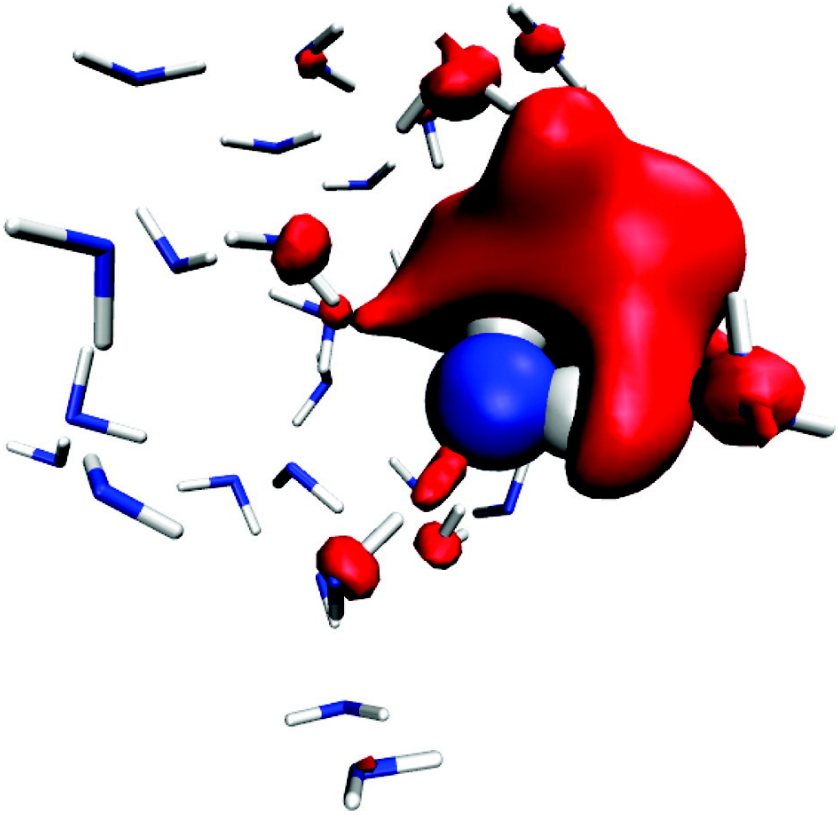
The FF concept (4/4)

- **FF advantages:**
  - instead of nuclei and electrons only complete atoms: fewer particles!
  - electrons *have to* be treated quantum-mechanically, whole atoms *can be* treated classically (or quantum-mechanically)
  - evaluation of the simple function $V(\vec{R})$ is **MUCH** faster than solving the electronic Schrödinger equation: see below

- **FF disadvantages:**
  - properties depending *explicitly* on the electrons are difficult or impossible to represent with FFs (but there are few properties of this kind; and tricks can be used: see below)
  - no FF is known that is both universal and sufficiently accurate
  - $\Rightarrow$ often, for new systems, a new FF has to be generated
  - transferability of a FF to other systems is unclear

In contrast, ab-initio quantum chemistry is universal by construction.
So, why FFs?

Why FFs? The speed advantage



ab-initio MD:
nuclei: classical-mechanical dynamics
electrons: quantum-mechanical forces on the nuclei

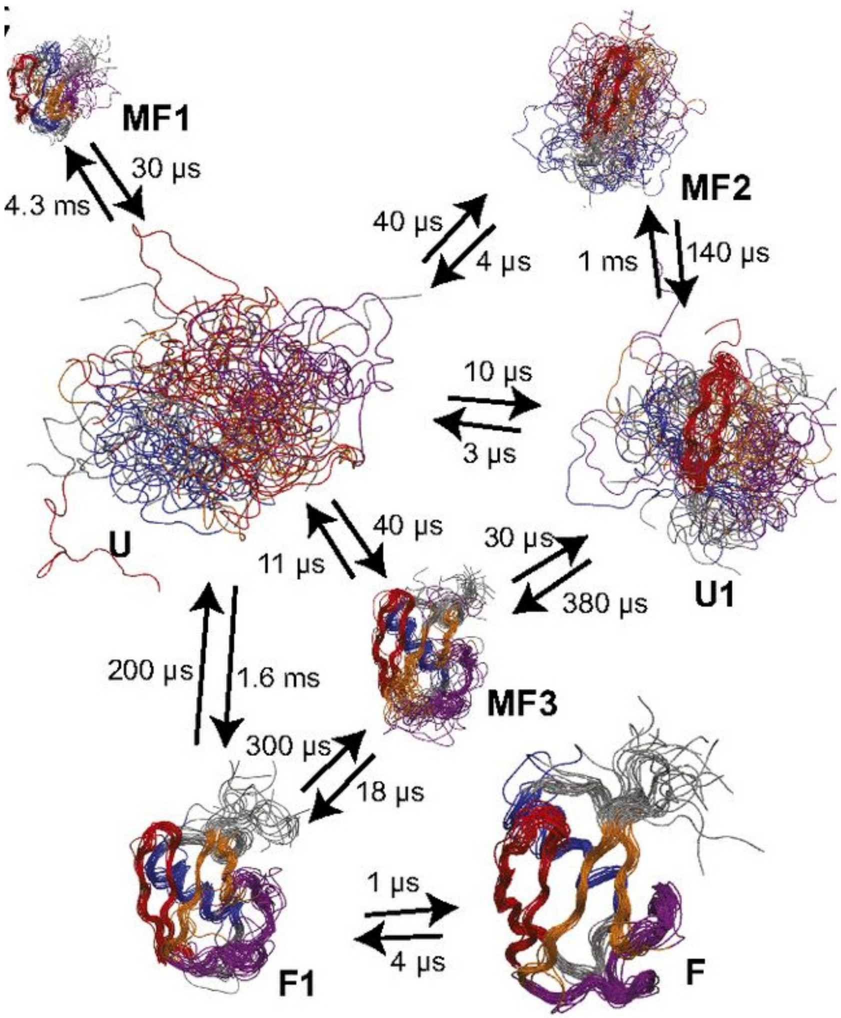
Accessible Systems: [1]

- on massively parallel HPC hardware
- 100–1000 aomts
- 2–200 ps simulated time in 1 week real time
- time step $\approx$ 0.2 fs
  $\Rightarrow 10^4$–10^6 time steps in 1 week real time

---

[1] J. Hutter, WIREs Comput. Mol. Sci. 2 (2012) 604.

Why FFs? The speed advantage



FF-MD:
nuclei: classical-mechanical dynamics
force field: forces on the nuclei

Accessible Systems:

- on massively parallel HPC hardware
- long times: folding of ubiquitin^2, total simulated time 8 ms (time step 5 fs $\Rightarrow 1.6 \times 10^{12}$ steps)
- huge systems: 1.34×10^{11} atoms in simulations^3 of material deformation after hypersonic impact

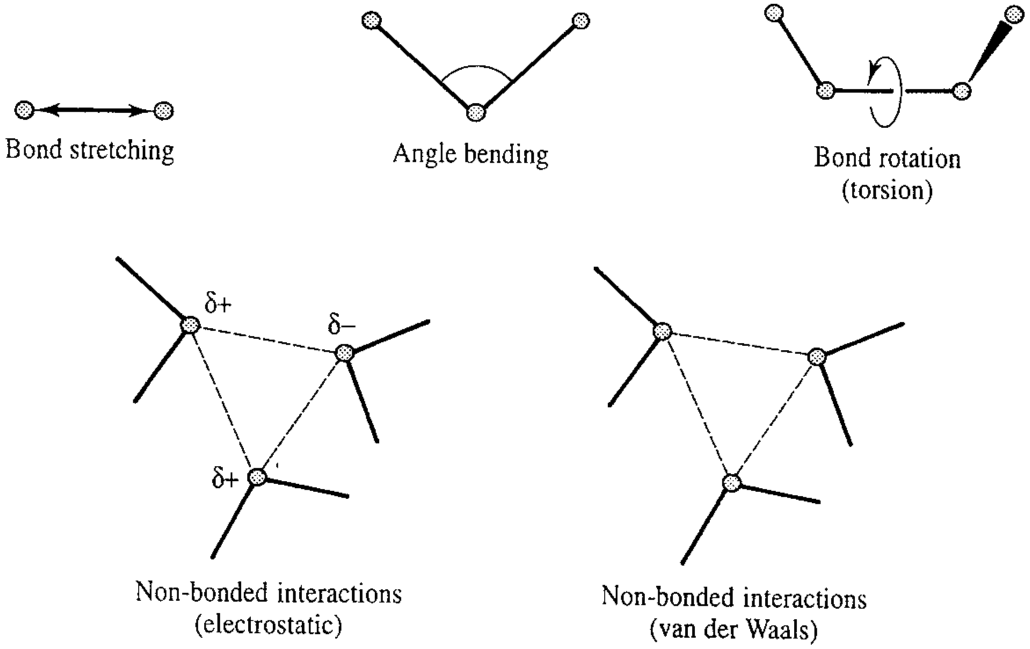
---

[2] S. Piana, K. Lindorff-Larsen and D. E. Shaw, Proc. Natl. Acad. Sci. USA 110 (2013) 5915.
[2] A. Nakano *et al.*, Int. J. High Perf. Comput. Appl. 22 (2008) 113.

Simple (bio)chemistry FFs

Functional form of a “class-I” force field (AMBER/GAFF, CHARMM, OPLS, MM2/3/4, . . .):

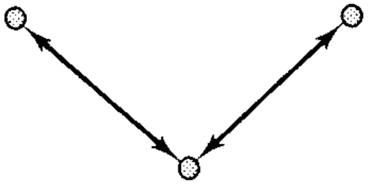
$$
\begin{aligned}
V(\vec{R}) \;=\; & \sum_{\text{bonds}} \frac{k_i}{2} \left(\ell_i - \ell_{i,0}\right)^2 && (1)\\
& + \sum_{\text{angles}} \frac{k_i}{2} \left(\vartheta_i - \vartheta_{i,0}\right)^2 && (2)\\
& + \sum_{\text{torsions}} \frac{k_n}{2} \left(1 + \cos(n\omega - \gamma)\right) && (3)\\
& + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{q_i q_j}{4\pi\epsilon_0 \, r_{ij}} && (4)\\
& + \sum_{i=1}^{N} \sum_{j>i}^{N} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right] && (5)
\end{aligned}
$$



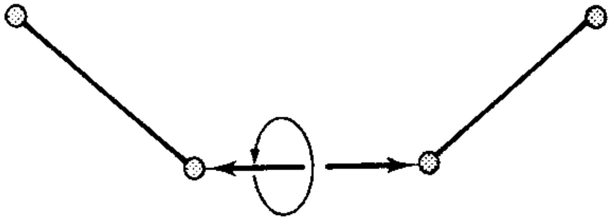
- several different contributions, with “physical” functional forms
- several parameters in each term
- obviously, each parameter in each term should also depend on which atoms are involved
- in contrast to ab-initio quantum chemistry, FF users also need to specify bonds between atoms!

Simple (bio)chemistry FFs

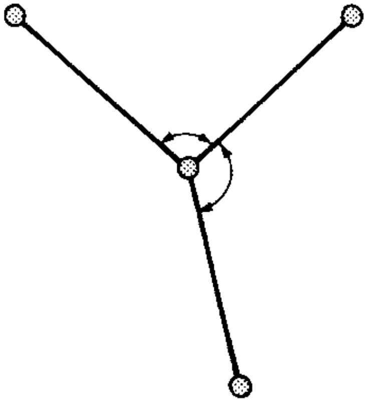
Obviously, the class-I functional form is too simple; we should also consider class-II, class-III, . . . , terms:



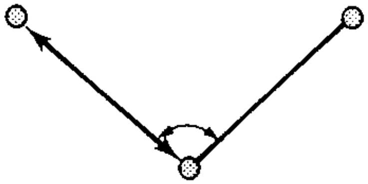
Stretch – stretch



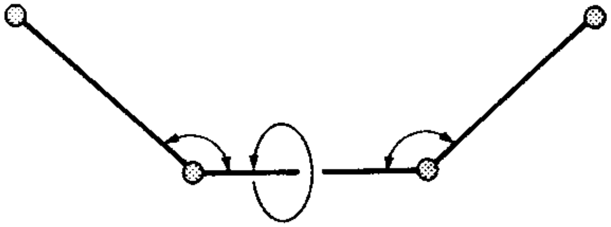
Stretch – torsion



Bend – bend



Stretch – bend

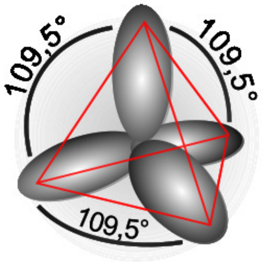


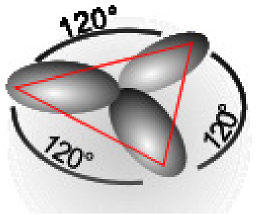
Bend – torsion

(cf. e.g. Leach)

Simple (bio)chemistry FFs

Most biochemistry force fields have *atom types*:
not only for different bond angle preferences,





sp^3 *vs* sp^2

but also for much finer differentiation:
Small(!) part of the OPLS-AA oxygen atom definitions:

atom	16	7	OH	OH Alcohol	8	15.999	2
atom	24	10	OH	OH Trifluoroethanol	8	15.999	2
atom	29	10	OH	OH Phenol	8	15.999	2
atom	31	10	OH	OH Diols	8	15.999	2
atom	33	12	OS	O Ether	8	15.999	2
atom	39	13	OS	O Acetal	8	15.999	2
atom	40	10	OH	OH Hemiacetal	8	15.999	2
atom	83	22	O	NC=O Amide	8	15.999	1
atom	95	22	O	O Urea	8	15.999	1
atom	99	22	O	O=CNHC=O Imide	8	15.999	1
atom	108	26	OH	OH Carboxylic Acid	8	15.999	2
atom	109	22	O	C=O Carboxylic Acid	8	15.999	1
atom	112	28	O2	COO- Carboxylate	8	15.999	1
atom	118	22	O	HC=O Aldehyde	8	15.999	1

atom	121	22	O	C=O Ketone	8	15.999	1
atom	164	43	OW	O Water (TIP3P)	8	15.999	2
atom	166	45	OW	O Water (SPC)	8	15.999	2
atom	170	47	O-	O- Hydroxide Ion	8	15.999	1
atom	187	4	O	Urea C=O	8	15.999	1
atom	192	4	O	Imide C=O	8	15.999	1
atom	217	4	O	Aldyhyde/Acyl Halide C=O	8	15.999	1
atom	266	4	O	Uracil & Thymine O2	8	15.999	1
atom	268	4	O	Uracil & Thymine O4	8	15.999	1
atom	280	4	O	Cytosine O2	8	15.999	1
atom	310	4	O	Guanine O6	8	15.999	1
atom	324	4	O	CytosineH+ O2	8	15.999	1
atom	359	5	OH	Methoxide CH3O-	8	15.999	1
atom	373	5	OH	Hydroxide Ion OH-	8	15.999	1
atom	379	42	O2	DiMe Phosphate O=P-O	8	15.999	1
atom	380	29	OS	DiMe Phosphate CH3-O	8	15.999	1
atom	407	4	O	Ester C=O	8	15.999	1
atom	408	20	OS	Ester CO-O-R	8	15.999	2
atom	414	20	OS	Phenyl Ester -OPh	8	15.999	2
atom	416	23	OY	Sulfonamide -SO2N	8	15.999	1
atom	435	23	OY	Sulfone R-SO2-R	8	15.999	1
atom	438	23	OY	Sulfoxide R-SO-R	8	15.999	1
atom	507	20	OS	Furan O	8	15.999	2
atom	512	20	OS	Oxazole O	8	15.999	2

Simple (bio)chemistry FFs

- **atom types, advantages:** one atom type = one very specific situation
  - far better accuracy of the FF
  - much clearer which reference data are needed (and for which parameters)
  - different atom types constitute *independent* parts of the FF
- **atom types, disadvantages:**
  - the user has to assign an atom type to every atom by hand (painful to impossible for $\gg 100$ atoms)
  - realistic chance that the atom type you need is missing
  - *serious* proliferation of parameters: (for MM3(91), with 71 atom types; from Leach)

term	estimated #params	actual #params
vdW	142	142
bond	900	290
angle	27000	824
torsion	1215000	2466

$\Rightarrow$ bad consequences:

- $*$ fitting all possible parameters would require $\approx 10^7$ independent reference data items: impossible!
- $*$ for your non-trivial molecule, it is almost certain that some torsional parameters are missing!
  $\Rightarrow$ Use your chemical judgement to borrow values from similar, existing torsions.

From fixed point charges to polarizability

!still 2b done!

- TIP4P as example for a fixed-point-charge FF: nice h-bonds!
- rule of thumb: w/o polarizability fine for neutrals and cations, not good for anions
- Drude model for polarizability
- polarizable FFs: TTM2/3-F, AMOEBA, etc.

!still 2b done!

FFs for (inorganic) materials: dielectrica, semiconductors, metals

Without directional bonding (angle, torsion) terms, many FFs are sums of pairwise interactions. However, pure pairwise potentials fail for certain bulk or surface properties, for several general reasons:

- ratio between melting temperature T_m and cohesive energy E_c (energy cost of removing an atom from the solid):

$$\text{observed for metals:} \quad \frac{E_c}{k_B T_m} = 30 \,, \qquad \text{pair potentials:} \quad \frac{E_c}{k_B T_m} = 10 \tag{6}$$

- ratio between E_c and vacancy formation energy E_v:

$$\text{observed for metals:} \quad \frac{E_v}{E_c} = \frac{1}{3} \text{ to } \frac{1}{4} \,, \qquad \text{pair potentials:} \quad \frac{E_v}{E_c} = 1 \quad \text{(w/o relaxation)} \tag{7}$$

- in cubic crystals, there are only three elastic constants (C_{11}, C_{12} and C_{44}) that do not have to be identical due to symmetry. Pair potentials further restrict this to $C_{12} = C_{44}$, although in many materials (metals, oxides, gold, …) these two values are different.
- Real metal surfaces tend to relax inwards, but surfaces of pair-potential models tend to relax outwards.

Intuitive reason: Pair potentials fail to describe changes in pair contributions depending on 3rd, 4th, … atoms in the immediate neighborhood. E.g., on surfaces, there typically are fewer but stronger bonds than in the bulk.

Therefore, starting from the ≈1980ies onwards, many different potentials with suitable 3-body, 4-body, … contributions were developed. The following attempts a unified view:

FFs for (inorganic) materials: dielectrica, semiconductors, metals

Embedded-atom method (EAM):

$$V(\vec{R}) = \sum_{j<i} V_2(r_{ij}) + \sum_i V_1(\rho_i) \qquad \text{with} \tag{8}$$

- a pairwise repulsive potential V_2, depending on atom-atom distances r_{ij}
- an "embedding function" V_1, representing the energy to put atom i into its position, within the electron density ρ_i at this location, given by a simple superposition of atomic densities from the neighboring atoms.

Originally (Daw/Baskes, Phys.Rev.B 29 (1984) 6443; Foiles/Baskes/Daw, Phys.Rev.B 33 (1986) 7983),

- atomic densities were taken from Hartree-Fock calculations for single atoms (via tabulated values),
- and functional forms for V_1 and V_2 were adjusted to fit known bulk properties, since formal derivation of these functions yielded only qualitatively correct results at best (Foiles/Baskes/Daw).

Advantages:

“reactive” FFs

- reaxFF, COMB, and others; including relation to Tersoff, BOPs, etc.
- methods for coupling non-reactive FFs, including EVB-QMDFF

even weirder stuff

- coarse-grained FFs
- photochemistry with FFs
- Bill Goddard's "electron FF"

discussion: is it advisable/possible to build your own FF?

FFs w/o physical terms

- neural net potentials and others
- machine learning stuff

some application cases

a brief glimpse on software

(difficult!)

- tinker
- gromacs, amber, namd, et al.
- LAMMPS
- ASE
- Avogadro, packmol, and friends

? WW of 2 molecules in the gas phase

(not sure, this would have go all the way towards global structure optimization)

? adsorption at surfaces

brief discussion of advantages/disadvantages of

- 2D/3D-periodic surface calculations, vs.
- non-periodic cluster models for surfaces