

Molecular and Materials Modelling

Lecture Summary
(Spring Semester 2015)

Giuseppe Accaputo
g@accaputo.ch

September 21, 2015

Disclaimer

This is a summary of the *Molecular and Materials Modelling* lecture [1] taught by Prof. VandeVondele during the spring semester 2015 at the ETH Zürich and was written by me as a preparation for the oral exam. All of the equations shown in this summary have been presented during the lecture; equations taken from other sources are appropriately referenced in the text.

Lecture 1: Introduction

- Potential energy surface (PES): interaction model that can be written as $V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$, a function of all coordinates \mathbf{r}_i of the atoms $1, \dots, N$
 - If the PES is known, the forces can be obtained easily by evaluating

$$F_i = \frac{-\partial V}{\partial r_i} \quad (1)$$

- Empirical methods: an empirical form is employed to describe the PES, e.g., Lennard-Jones potential
 - + Can compute the PES very efficiently; results in short time-to-solution, larger models ($10^3 - 10^6$) and longer time scales (ns, ms)
 - + Can be highly accurate if properly parametrized
 - - Might have poor accuracy where they have not been tested
 - - Might be time consuming to setup

- Does not easily give access to properties related to electronic structure
- First principles methods: the PES is obtained from an approximate solution of the Schrödinger equation
 - + Relatively accurate
 - + Wide range of systems can be studied *out of the box*
 - + Does give access to properties related to electronic structure
 - - Demanding to compute PES; results in smaller models (< 1000 atoms) and shorter time scales (< 100ps)

- The Lennard-Jones potential is defined as

$$V_{LJ} = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (2)$$

- Can be used to describe inert gases (gases which do not undergo chemical reactions), e.g., noble gases like He
- Describes long range van der Waals interactions ($(\cdot)^6$), i.e., the dispersion forces like dipole-dipole interactions and also describes the short range repulsion ($(\cdot)^{12}$)
- Often used in simulations when the objective is to model a general class of effects and the only requirement is to have a physically reasonable potential
- To try to interpret a multidimensional function like the PES it is recommended to use an integrated quantity like the free energy instead of trying to visualize it
- Energy minima ($F_i = \frac{-\partial V}{\partial r_i} = 0$) are stable states of the system
 - The lowest energy pathway connecting two minima passes through a transition state (saddle points on the PES), which corresponds to the reaction mechanism that converts the system from one to the other stable state
 - Number of minima is expected to grow as $O(\exp\{N\})$
- Locating the absolute minimum of a PES is called *global optimization* and is a rather difficult task to accomplish
- Energy minimization (also called geometry optimization): find the lowest value of the PES by changing the coordinates *little by little* to relax the structure to a nearby *local* energy minimum
- Locating minima by having gradients available can be done using:
 1. Steepest descent (not efficient, but simple)
 2. Non-linear conjugate gradient
 - Converges relatively quickly, is robust and needs only $O(N)$ storage

3. Quasi-Newton methods

a) Broyden-Fletcher-Goldfarb-Shanno (BFGS)

- Approximates the Hessian matrix $\underline{\mathbf{B}}$ by solving the equation

$$\underline{\mathbf{B}}_k \mathbf{p}_k = -\nabla f(\mathbf{x}_k) \quad (3)$$

b) L-BFGS

- Limited memory version of the BFGS, where the matrix $\underline{\mathbf{B}}$ is approximated by a few (~ 10) vectors only

- Locating transition states:

1. Constrained optimization: if a reaction coordinate is known, minimize the energy with the constraint that the reaction coordinate has a specific value

- Reaction coordinate: abstract variable that connects reactant and product of a given reaction, e.g., distance between the atoms
- During the optimization the constraint $RC = a$ is strictly fulfilled by using a Lagrangian multiplier for example
- The system is restrained by adding a penalty term to the total energy

2. Nudged Elastic Band (NEB): method that locates transition states and minimum energy paths between known reactants and products

- Minimization of a chain of states connected by springs; this full chain is minimized
- Climbing image NEB: drives the image with the highest energy up to the saddle point; this image does not see the spring forces along the band, making this version of NEB less dependent on the spring constant k
 - * The true force at this image along the tangent is inverted; the image thus tries to maximize its energy along the band, and minimize it in all other directions. When this image converges, it will be the exact saddle point

Lecture 2: Biomolecular Force Fields

Interaction Terms for Force Fields

- n -Body interaction terms found in common force fields:

1. $n = 1$

- Kinetic energy (all atoms)
- Interaction with an external electric field (charged atoms)
- Atomic position restraining (listed or all atoms; *unphysical*)

2. $n = 2$
 - Pairwise nonbonded interaction (all atom pairs)
 - Covalent bond (bonded atoms)
 - Distance restraining (listed atom pairs; *unphysical*)
3. $n = 3$
 - Covalent bond-angle bending (pairs of bond)
4. $n = 4$
 - Improper dihedral interaction (atoms in dihedrals)

Atomic Orbitals and Valence Bond Theory

Atomic Orbitals

- Orbit: a definite path
- Orbital: regions of space
- Heisenberg uncertainty principle: it is impossible to define with absolute precision, at the same time, both the position and the momentum of the electron; thus it is impossible to plot an orbit of an electron around a nucleus, since we do not know where the electron is and where it's going to be an instant later
- s orbitals
 - Orbitals that are spherically symmetric around the nucleus
 - $1s$ orbital: the 1 represents the fact that the orbital is in the energy level closest to the nucleus (first energy level)
 - $2s$ orbital: is similar to the $1s$ orbital except that the region where there is the greatest chance of finding the electron is further from the nucleus (second energy level)
- p orbitals
 - A p orbital is rather like 2 identical balloons tied together at the nucleus
 - At the first energy level, the only orbital available is the $1s$ orbital
 - p orbitals can point in a particular direction
- Each orbital can only hold 2 electrons
 - 1st energy level: $1s$
 - 2nd energy level: $2s, 2p_x, 2p_y, 2p_z$
 - 3rd energy level: $3s, 3p_x, 3p_y, 3p_z$, and 5 d orbitals
- $1s^2$ means that a $1s$ orbital is filled with 2 electrons

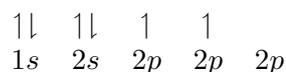
- Electrons are fitted starting from the orbitals closest to the nucleus, i.e., $1s, 2s, 2p, 3s, 3p, 4s, 3d, 4p$
 - The $n s$ orbitals come before the $n+1 d$ orbitals (see position of the $4s$ orbital), since in the energy diagram the $n s$ orbitals have a slightly lower energy than the $n + 1 d$ orbitals

Covalent Bonding

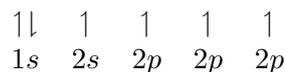
- Every atom wants to look like noble gases, i.e., having a fully occupied outer shell
- With ionic bonds, two atoms actually exchange electrons; with covalent bonds, the atoms *share* the electrons, resulting in an increased number of electrons for both atoms (example: two oxygen atoms sharing 2 electrons, resulting in a second outer shell fully occupied with 8 electrons)
 - With a covalent bond, the electrons that are part of this bond can actually move between the probability distribution clouds of both atoms, giving both atoms the *feeling* of having additional electrons in the outer shell
 - Atoms with the same electronegativity can build covalent bonds, that is, atoms that are rather similar, e.g., two oxygen atoms. For H_2O we also have covalent bonds between the H atoms and the O atom, but it's called a *polar covalent bond*, since the shared electrons will spend a bit more time in the probability distribution cloud of the O atom because of the difference in electronegativity, but are still shared between both atoms

Hybridization

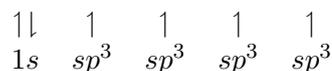
- Hybridization is the concept of mixing atomic orbitals ($1s, 2s, 2p$, and so on) into new hybrid orbitals suitable for the pairing of electrons to form chemical bonds in valence bond theory
- sp^3 hybridization: hybridization is shown by explaining the covalent bonds of the CH_4 (methane) molecule; methane is called a sp^3 -hybridized molecule
 - C has atomic number 6 and thus the following ground state configuration:



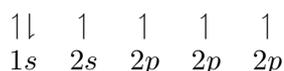
- * In this configuration the carbon atom can utilize its two singly occupied p orbitals to form two covalent bonds with two hydrogen atoms
- By an excitation of an electron from the doubly occupied $2s$ orbital to the empty $2p$ orbital, the carbon atom can also bond to four hydrogen atoms:



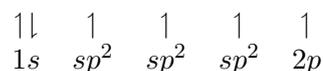
Quantum mechanically, the lowest energy is obtained if the four bonds are equivalent, which requires that they be formed from equivalent orbitals (same energy level) on the carbon; thus, the sp^3 hybrid is introduced, which is obtained from the linear combination of the valence-shell s and p wave functions:



- In CH_4 , four sp^3 hybrid orbitals are overlapped by hydrogen $1s$ orbitals, thus resulting in four σ bonds (four single covalent bonds) of equal length and strength
- sp^2 hybridization: hybridization is shown by explaining the covalent bonds of the C_2H_4 (ethylene) molecule; ethylene is called a sp^2 -hybridized molecule
 - Hund's rule states that all orbitals among the same energy levels have to be filled with at least one electron before being paired up again; thus, the following configuration results:



- For ethylene, supporting evidence shows that the carbon is sp^2 hybridized, meaning that one s orbital is being mixed with two p orbitals. Each carbon is joining to two H and another C atom, meaning compared to methane we only need 3 bonds, so 3 new hybrid orbitals (the sp^2 orbitals) are formed to fulfill this requirement:



- The bonds between the atoms are constructed as follows (see Fig. 1):
 - * One sp^2 orbital of each carbon atom forms a σ bond between the carbon atoms by overlapping
 - σ bonds are made from hybridized orbitals
 - * The remaining two sp^2 orbitals overlap with $1s$ orbital of the two hydrogen atoms, producing 4 σ bonds
 - * The remaining *non-hybridized* $2p$ orbital (it's the $2p_z$ orbital to be more precise) forms a π bond between the two carbons by parallel overlapping
 - π bonds are made from leftover p orbitals

AMBER Force Field

- The classical Hamiltonian: the AMBER force field (see Fig. 2)

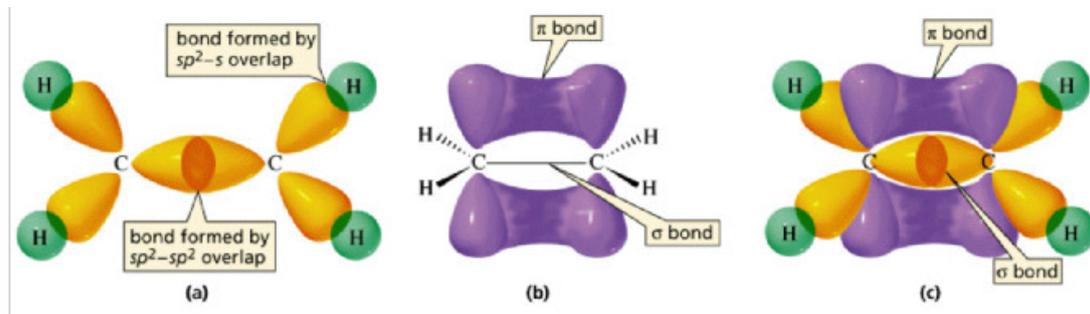


Figure 1: Hybridization of the ethylene molecule

$$\begin{aligned}
 V(r) &= E_{\text{bonded}} + E_{\text{nonbonded}} \\
 &= \sum_{\text{bonds}} K_b (b - b_0)^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_0)^2 \\
 &+ \sum_{\text{dihedrals}} (V_n/2)(1 + \cos[n\phi - \delta]) \\
 &+ \sum_{\text{nonbij}} (A_{ij}/r_{ij}^{12}) - (B_{ij}/r_{ij}^6) + (q_i q_j / r_{ij})
 \end{aligned}$$

Figure 2: The AMBER force field

- Bond energy terms (covalent interaction, intramolecular):

$$K(r - r_0)^2 \quad , \quad (4)$$

where r_0 is the reference bond length

- Angular terms:

$$K(\theta - \theta_0)^2 \quad , \quad (5)$$

where θ_0 is the reference angle

- Dihedral terms:

$$K[1 + \cos(n\phi - \phi_0)] \quad , \quad (6)$$

where ϕ_0 is the reference torsion angle

* The approximation keeps into account the angular periodicity

- Coulomb interactions (noncovalent interaction, intermolecular):

$$\sum_i^N \sum_{j>i}^N \frac{1}{4\pi\epsilon_0\epsilon_1} \frac{q_i q_j}{r_{ij}} \quad (7)$$

* This sum is not suitable for large and periodic systems, due to the long range of the interactions

- Van der Waals terms (noncovalent interaction, intermolecular):

$$\sum_i^N \sum_{j>i}^N 4\epsilon \left[\left(\frac{r_{ij}}{\sigma} \right)^{-12} - \left(\frac{r_{ij}}{\sigma} \right)^{-6} \right] \quad (8)$$

* Electronegativity: the power of an atom in a molecule to attract electrons to itself; a high electronegativity means a high power to attract electrons

* Dipole moment μ : if a molecule is polar, than it's said to have a dipole moment, i.e., $\mu > 0$; it is the measure of net molecular polarity, which is the magnitude of the charge Q at either end of the molecular dipole times the distance r between the charges

Lecture 3: Surfaces and Potentials for Metals

Surface Free Energy

- Work which has to be expended in order to increase the size of the surface of a phase; it's referenced as *surface free energy* when used for solid surfaces, and it's called *surface tension* when concerned with a liquid phase

- The internal energy of a bulk system (infinite solid) in equilibrium is given by

$$U = TS - PV + \mu N \quad , \quad (9)$$

where S is the entropy, V the volume, P the pressure, and N the particle number

- If we cleave the bulk, a surface of area A is created and the total energy increases by an amount proportional to the area A ,

$$U = TS - PV + \mu N + \gamma A \quad (\text{Euler equation}) \quad , \quad (10)$$

where γ is called the surface tension (or surface free energy)

Wulff Theorem and the Wulff Construction

- Gibbs stated that a crystal will arrange itself such that the surface free energy is minimized by assuming a shape of low surface energy.
- Gibbs defined the quantity

$$\Delta G_i = \sum_j \gamma_j A_j \quad , \quad (11)$$

with γ_j being the surface free energy per unit of the j th crystal face and A_j the area of said face

- ΔG_i represents the difference in energy between a real crystal composed of i molecules with a surface and a similar configuration of i molecules inside a infinitely large crystal, and is therefore the energy associated with the surface
- The equilibrium shape of the crystal will then be that which minimizes the value of ΔG_i

Wulff Theorem

- The length of the vector drawn normal to the crystal face h_j will be proportional to its surface energy γ_j , i.e.,

$$h_j = \lambda \gamma_j \quad . \quad (12)$$

- The vector h_j is the *height* of the j th face, drawn from the center of the crystal to the face; for a spherical crystal this is simply the radius

Wulff Construction

- The crystal shape in equilibrium is constructed as the envelope of the family of perpendicular lines

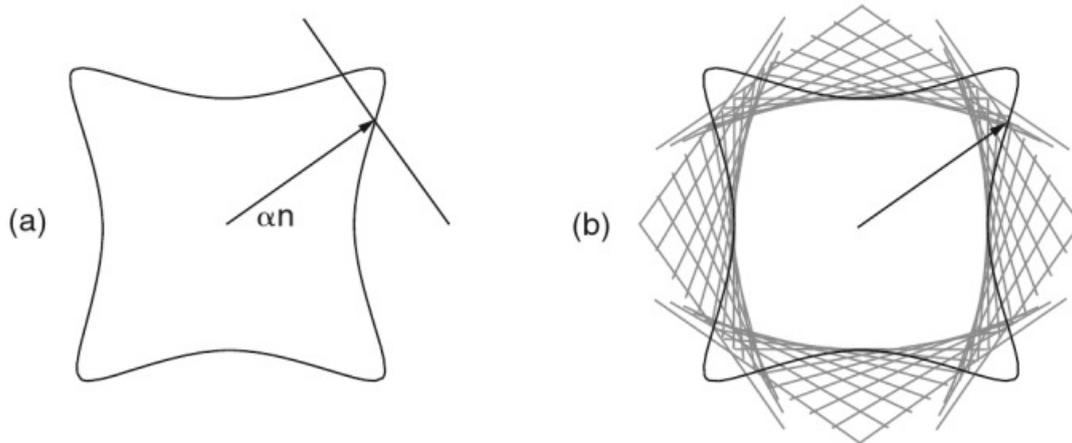


Figure 3: (a) Polar plot of the surface tension α as a function of the normal \mathbf{n} to a crystal face. (b) Wulff's construction of the corresponding crystal surface as the envelope of the family of perpendicular lines; the resulting equilibrium shape has pointy corners and curved edges

Miller Indices

- A family of lattice planes is determined by three integers $h, k,$ and l ; (h, k, l) are the Miller indices:

$$h\mathbf{n}_1 + k\mathbf{n}_2 + l\mathbf{n}_3 = p \quad (13)$$

- The Miller indices $(001), (110), (111)$ denote high-symmetry faces in a cubic system that usually have a low surface free energy and are often used in the Wulff construction

Potentials for Metals

Metallic Bond

- Metallic bond is the bonding between molecules within metals
 - Sharing of a sea of delocalised electrons (electrons are not attached to one particular ion) amongst a lattice of positive ions, where the electrons act as a “glue” giving the substance a definite structure
- Electrons and positive ions have a strong attractive force between them; this results in high melting or boiling points

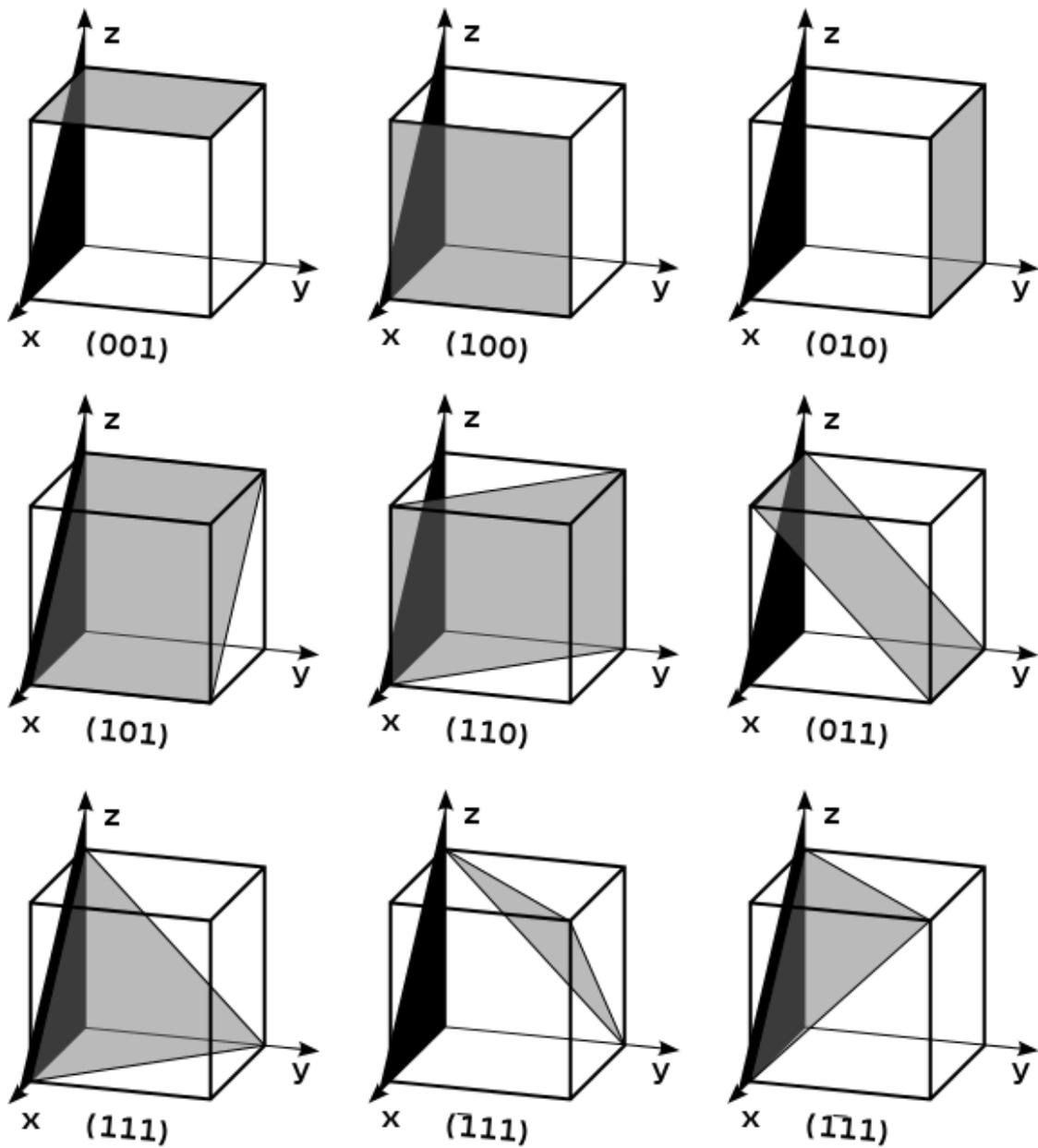
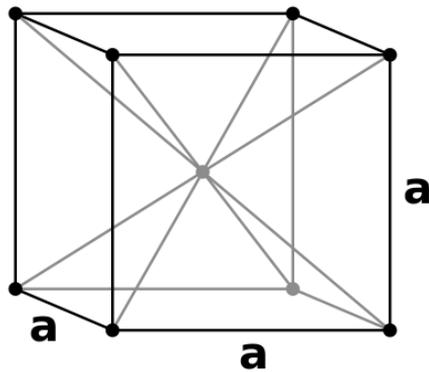
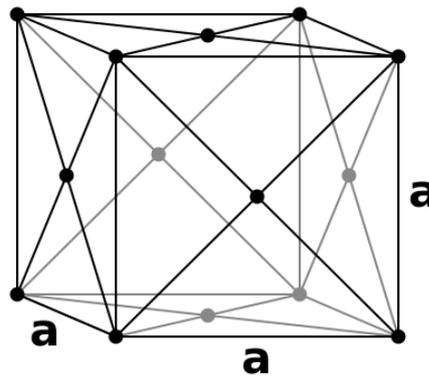


Figure 4: Example of crystallographic planes and Miller indices for a cubic structure



(a) Unit cell of a body-centered cubic material



(b) Unit cell of a face centered cubic

Limitations of Pair Potentials for Metals

- Total potential energy of the system of N atoms interacting via pair potential is

$$V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \sum_i \sum_{j>i} U_2(|\mathbf{r}_j - \mathbf{r}_i|) \quad (14)$$

- Lennard-Jones: van der Waals interaction (intermolecular force)

$$V_{LJ} = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6 \right] \quad (15)$$

- Motions which tend to change the coordination appreciably are greatly discouraged by their high energy cost; such motions easily occur in the proximity of surfaces and defects (the extreme case consists of pulling an atom out of the system)
 - Two-body schemes cannot reproduce this mechanism, since they imply a linear dependence on the energy of an atom upon its coordination
 - Vacancy formation leads to a change in coordination of twelve neighbouring atoms from twelve to eleven for the given atom in an f.c.c. bulk, which is a rather small change; this leads to a low vacancy formation energy
 - Extracting a single atom from the system implies changing its coordination from twelve neighbouring atoms to zero (no neighbours are available anymore, since it has been removed from the bulk), leading to a rather high vacancy formation energy because of the dramatic change
- In the paper “Simulation of gold in the glue model”, Ercolessi found some indicators of many-body effects in f.c.c. crystals (see Fig. 6)
- Limitations of pair potentials for metals:

Material	C_{12}/C_{44}	E_v^F/E_c	$E_c/k_B T_m$	S_m
LJ	1.0	1.00	12.8	2.16
Ar	1.1	0.95	11.1	1.69
Kr	1.0	0.66	11.5	1.69
Al	2.0	0.23	41.5	1.39
Pb	2.3	0.24	39.4	0.96
Ni	1.2	0.31	29.8	1.22
Pt	3.3	0.26	33.3	1.16
Cu	1.5	0.33	29.9	1.16
Ag	1.9	0.36	27.8	1.10
Au	3.7	0.25	32.8	1.13

Figure 6: Some indicators of many-body effects in f.c.c. crystals. Several metals are compared with a Lennard-Jones solid (LJ) (Ladd and Woodcock 1978, Jacucci 1984), taken as a typical two-body system, and the rare-gas solids Ar and Kr

1. Pair potentials do not have environmental dependence, e.g., atom in the bulk is too similar to the atom on the surface (surface-specific properties are basically ignored)
2. Pair potentials do not account for directional nature of the covalent bond
 - Atoms bonded covalently prefer specific orientations in space relative to one another. As a result, molecules in which atoms are bonded covalently have definite shapes. Ionic bonding (where one or more atoms lose electrons and other gain them in order to produce a noble gas electron configuration) for example has no directional nature.
 - Transition metals are those elements having a partially filled d orbital. Covalent contributions of these metals cannot be described. Pair potentials work better for metals in which cohesion is provided by s and p electrons
3. The vacancy formation energy E_v is significantly overestimated by pair potentials
 - $E_v \sim E_c$ with pair potentials, where E_c is the cohesive energy, i.e., the energy required for separating the condensed material into isolated free atoms; experimentally we get $E_v = 0.25 E_c$ for gold and $E_v = 0.33 E_c$ for copper. If $E_v \sim E_c$ would hold, than pairwise bonds would be breaking
4. The ratio between the cohesive energy E_c and the melting temperature T_m is

- underestimated as much as 2 times
5. Pair potentials do not describe the deviations from the Cauchy relation for elastic constants in cubic crystals ($C_{12} = C_{44}$)
 6. In a two-body system, the melting temperature T_m is usually near $0.1E_c/k_B$; in Au we have $T_m \sim 0.03E_c/K_b$
 7. The low $E_c/k_B T_m$ ratio causes an unrealistically high evaporation rate to take place at two-body surfaces near the melting point

Embedded Atom Method (EAM)

- Approximation describing the energy between atoms in a metal (interatomic potential)
- Approach to the description of bonding in metallic systems is based on the concept of local density, which accounts for the dependence of the strength of individual bonds on the local environment (especially important for the simulation of surfaces and defects)
- The total energy of a system of N atoms consists of a sum of embedding energies and a pairwise potential; the total energy is thus given by

$$E_{\text{tot}} = \sum_i F_\alpha \left(\underbrace{\sum_{i \neq j} \rho_\beta(r_{ij})}_{\text{embedding energy}} \right) + \frac{1}{2} \sum_{i \neq j} \underbrace{\Phi_{\alpha\beta}(r_{ij})}_{\text{pairwise potential}} \quad , \quad (16)$$

where r_{ij} is the distance between atoms i and j , $\Phi_{\alpha\beta}$ is a pairwise potential function, ρ_β is the contribution of the electron charge density from atom j of type β at the location of atom i , and F is an embedding function that represents the energy required to place an atom i of type α into the electron cloud (the embedding energy)

- The sum over ρ_β is a measure of the local electron density ρ_i
- A consequence of Pauli's principle (i.e., repulsion between electrons) is that the strength of individual bonds decreases as the local environment becomes more crowded; the term representing the local electron density ρ_i accounts for this dependence on the environment of an atom
 - * With two-body potentials the strength of individual bonds does not depend on the environment, but only on the distance between the two atoms involved: the position of all other atoms are not relevant
- EAM uses first-principle considerations in Eq. (16); for example, ρ_i is taken from Hartree-Fock calculations
- The general form of the potential can be considered as a generalization of the basic idea of the Density Functional Theory, i.e., the local electron density can be used to calculate the energy

Glue Potential

- Has the same analytical form as EAM (Eq. (16)), i.e., the total energy consists of a sum of pairwise potentials and a sum accounting for the many-body effects, which is called *the glue* in this case:

$$E_{\text{tot}} = \sum_i U_\alpha \left(\underbrace{\sum_{i \neq j} \rho_\beta(r_{ij})}_{\text{glue potential}} \right) + \frac{1}{2} \sum_{i \neq j} \underbrace{\Phi_{\alpha\beta}(r_{ij})}_{\text{pairwise potential}} \quad , \quad (17)$$

- The philosophy of the glue model implementations is that of trying to construct potentials that work reasonably well in a wide range of geometries, with special attention to surfaces and melting point
 - With EAM, potentials for the same element can exhibit different properties, particularly in geometries not considered during fitting (e.g., surfaces, liquid, etc.)
 - Fits are done by considering all functions constituting the potential presented in Eq. (17) as free to be determined, i.e., the different functions in the glue Hamiltonian in Eq. (17) are completely determined from empirical fits; the sole goal is to maximize the accuracy of the model
 - * In comparison, the EAM Hamiltonian (Eq. (16)) uses first-principle methods for the fit
 - Rather successful potentials were constructed for gold and lead
 - Fully empirical fits have severe limitations but are *easily* available, whereas ab initio methods produce very accurate interactions but are very expensive to compute \implies Force-Matching Method uses a fit that is not only driven by experimental data, but also by a large quantity of numerical data about the interactions obtained by first-principles

Lecture 4: Molecular Dynamics

Deriving the Verlet Algorithm from a Lagrangian

- Hamilton's principle: "Of all the possible paths along which a dynamical system may move from one point to another within a specified time interval (consistent with any constraints), the actual path followed is that which minimizes the time integral of the difference between the kinetic and potential energies."

$$\delta \int_{t_1}^{t_2} T(\mathbf{x}, \dot{\mathbf{x}}, t) - U(\mathbf{x}, t) dt = \delta \int_{t_1}^{t_2} L(\mathbf{x}, \dot{\mathbf{x}}, t) dt = 0 \quad , \quad (18)$$

where $T(\cdot)$ is the kinetic energy, $U(\cdot)$ the potential energy and the quantity $T(\mathbf{x}, \dot{\mathbf{x}}, t) - U(\mathbf{x}, t)$ is called the Lagrangian $L(\mathbf{x}, \dot{\mathbf{x}}, t)$

- The Lagrangian L has to satisfy the following relation:

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = 0 \quad , \quad i = 1, \dots, n \quad . \quad (19)$$

These equations are called The Lagrangian equations of motion and work with generalized coordinates

- The theory required to solve problems of the type presented in Eq. (18) is based on Euler's equations:

1. Let $y(x)$ be an extremum path from position 1 to 2, i.e., the minimized path integral of a function f (see Fig. 7)

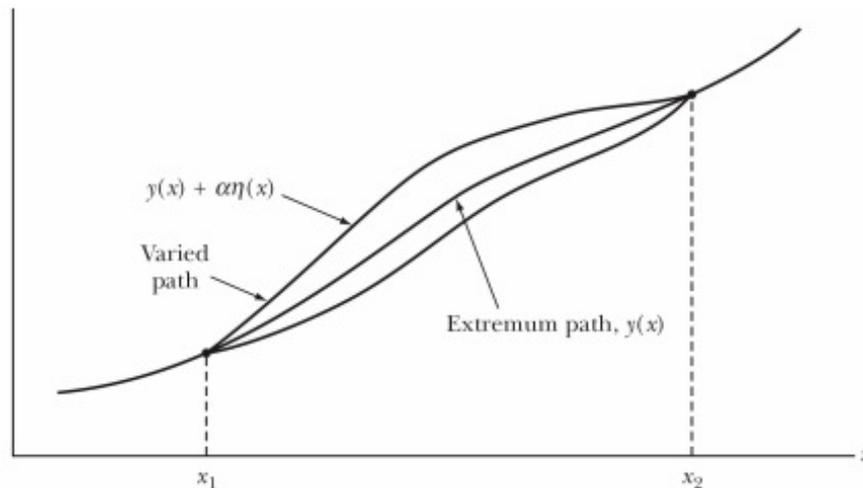


Figure 7: Path $y(x)$ that is used to move from position 1 to position 2

2. If we vary the path slightly by adding a second function $\eta(x)$, the path integral is expected to increase:

$$y(\alpha, x) = y(0, x) + \alpha\eta(x) \quad (20)$$

* $\eta(x)$ is required to have a continuous first derivative, i.e., $\eta(x) \in C^1(\mathbb{R})$, and it is required to vanish at the end points, i.e., $\eta(x_1) = \eta(x_2) = 0$

3. The path integral f is an extreme if

$$\int_{x_1}^{x_2} \left(\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial \dot{y}} \right) \eta(x) dx \right) \Big|_{\alpha=0} = 0 \quad (21)$$

4. Since $\eta(x)$ is an arbitrary function, Eq. (21) is only satisfied if

$$\frac{\partial f}{\partial y} - \frac{d}{dx} \left(\frac{\partial f}{\partial \dot{y}} \right) = 0 \quad (22)$$

This equation is known as Euler's equation

- Lagrange equations of motion can be written in rectangular coordinates as (Newton's law; Newtonian mechanics)

$$-\nabla_i V = \mathbf{F}_i = m_i \mathbf{a}_i \quad , \quad (23)$$

where \mathbf{a}_i is the acceleration, \mathbf{F}_i the force and V the potential energy

- The equations of motions are integrated using a time integration algorithm (e.g., Verlet algorithm)
- Derivation of the Verlet algorithm:
 1. Start with the discretization of the action integral:

$$\begin{aligned} I &= \int_{t_1}^{t_2} L dt = \Delta \sum_{n=0}^{N-1} T(q^{(n)}, \dot{q}^{(n)}) - V(q^{(n)}) \\ &\simeq \Delta \sum_{n=0}^{N-1} \frac{1}{2} m \left(\frac{q^{(n+1)} - q^{(n)}}{\Delta} \right)^2 - V(q^{(n)}) \end{aligned} \quad (24)$$

2. The variation of I becomes a standard derivation in this case:

$$0 = \frac{\partial I}{\partial q^{(k)}} = \frac{m}{\Delta^2} (q^{(k)} - q^{(k-1)}) - \frac{m}{\Delta^2} (q^{(k+1)} - q^{(k)}) - \frac{\partial V(q^{(k)})}{\partial q^{(k)}} \quad (25)$$

$$\iff q^{(k+1)} = 2q^{(k)} - q^{(k-1)} + F(q^{(k)}) \frac{\Delta^2}{m} \quad (26)$$

- The algorithm has an error of $O(\Delta^4)$ and does not include the explicit calculation of the velocities

- The velocity Verlet algorithm:

$$q^{(k+1)} = q^{(k)} + v^{(k)} \Delta + \frac{1}{2m} F(q^{(k)}) \Delta^2 \quad (27)$$

$$v^{(k+1/2)} = v^{(k)} + \frac{1}{2m} F(q^{(k)}) \Delta \quad (28)$$

$$v^{(k+1)} = v^{(k+1/2)} + \frac{1}{2m} F(q^{(k+1)}) \Delta \quad (29)$$

$$(30)$$

- Both Verlet forms have the following properties:

1. Conservation of total energy $E = T + U$, i.e., $\frac{dE}{dt} \stackrel{!}{=} 0$:

$$\frac{dE}{dt} \stackrel{!}{=} 0 \quad (31)$$

$$\Leftrightarrow \frac{dE}{dt} = \frac{d}{dt}(T + U) = \frac{dT}{dt} + \frac{dU}{dt} \quad (32)$$

$$= \frac{d}{dt} \left(\frac{1}{2} m \dot{x}^2 \right) - m \ddot{x} = m \ddot{x} - m \ddot{x} = 0 \quad (33)$$

2. Conservation of the phase space area
3. Are time reversible
4. Satisfy the symplectic condition

Treatment of Periodic Boundary Conditions in Molecular Dynamics

- Minimum image convention: only closest image is candidate to interact
- Use Verlet list to keep track of the neighbouring atoms (uses cell lists to identify the neighbouring atoms, resulting in a scaling time of $O(N)$ instead of $O(N^2)$)

Main Quantities Extracted from Trajectories

- Instantaneous value of a quantity A :

$$A(t) = f(\mathbf{r}_1(t), \dots, \mathbf{r}_N(t), \dot{\mathbf{r}}_1(t), \dots, \dot{\mathbf{r}}_N(t)) \quad (34)$$

- Average value of a quantity A :

$$\langle A \rangle = \frac{1}{n} \sum_{t=1}^n A(t) \quad (35)$$

- Mean square displacement (MSD):

$$MSD = \langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \rangle \quad (36)$$

- The MSD is a measure of the average distance a molecule travels
- MSD can be used to calculate the diffusion coefficient in three dimensions:

$$D = \lim_{t \rightarrow \infty} \frac{1}{6t} MSD \quad (37)$$

- Estimation of the pressure using MD (virial equation):

$$PV = Nk_B T + \frac{1}{D} \left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \mathbf{F}_i \right\rangle \quad (38)$$

- Real space correlations:

$$\langle A(\mathbf{r})A(0) \rangle \quad (39)$$

- The radial distribution function $g(r)$ is pair correlation function that describes how, on average, the atoms in a system are radially packed around each other:

$$\rho g(r) = \frac{1}{N} \left\langle \sum_{i=1}^N \sum_{i \neq j}^N \delta(r - r_{ij}) \right\rangle \quad (40)$$

- * $g(r)$ gives the probability of finding a particle i in the distance r_{ij} of a particle j (see Fig. 8 for the radial distribution function of solid and liquid Argon)
- * Is a particularly effective way of describing the average structure of disordered molecular systems such as liquids, where there is continual movement of the atoms and it is extremely useful to be able to deal with the average structure
- * $g(r)$ can be deduced experimentally from x-ray or neutron diffraction studies, thus providing a direct comparison between experiment and simulation

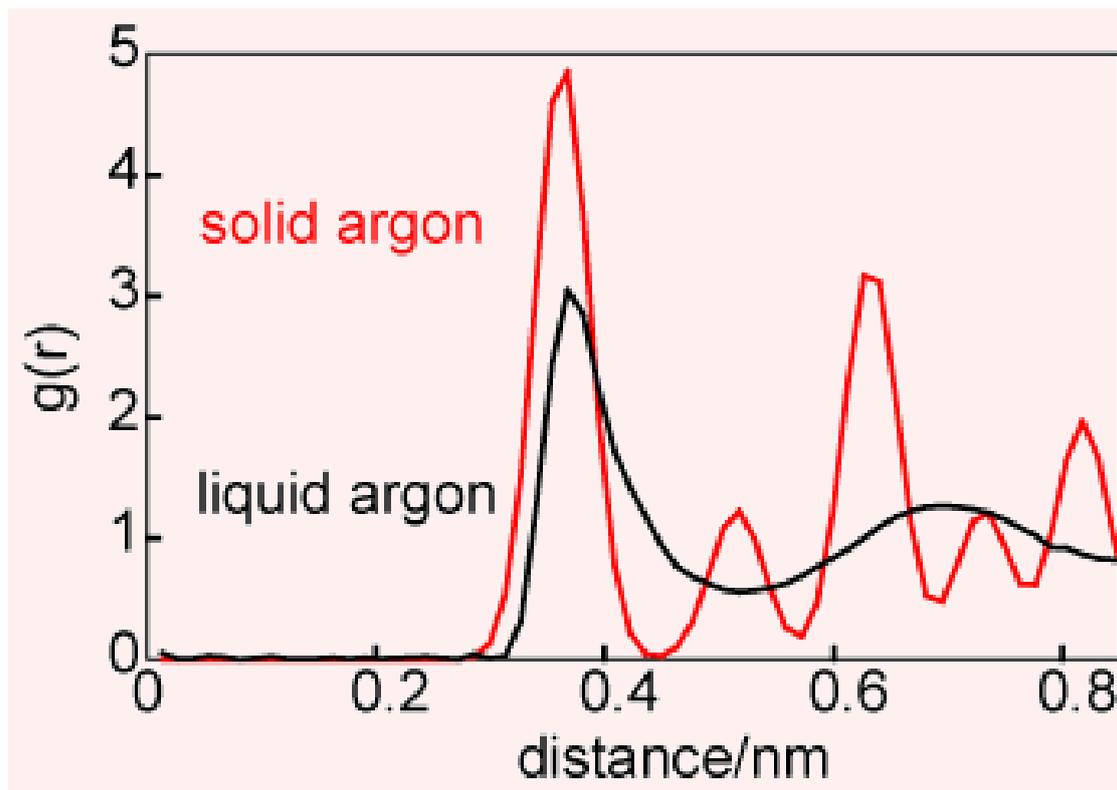


Figure 8: At short separations (small r), $g(r) = 0$, indicating the width of the atoms; the first peak in $g(r)$ is typically the strongest feature, since it is due to relatively strong chemical bonding and repulsion effects felt between atoms in the first shell; the attenuation of the peaks at increasing radial distances for the liquid Argon from the center indicates the decreasing degree of order, which indicates the absence of a long-range order in liquids and gases; for the solid Argon the long-range order is visible by the existence of peaks at increasing radial distances

Ways to Control the Temperature

Velocity Rescaling

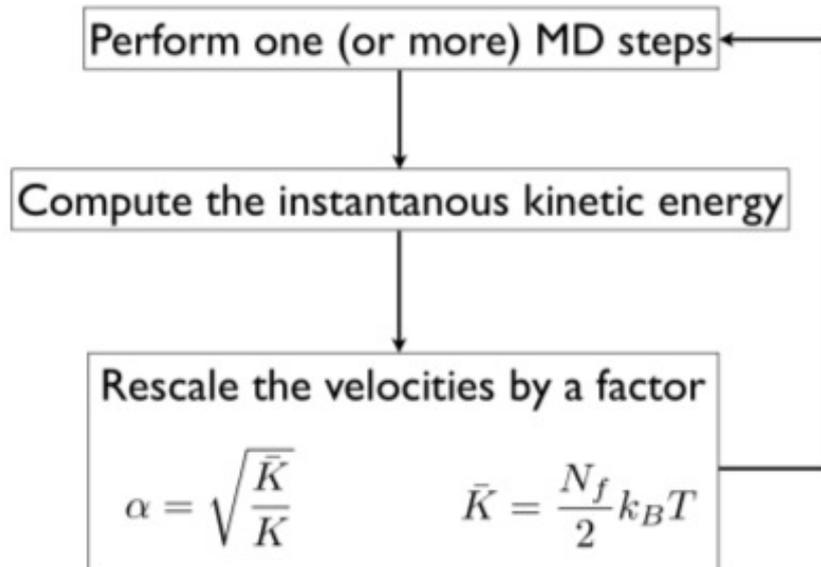


Figure 9: Velocity rescaling

Nosé-Hoover Thermostat

- Add a new degree of freedom s that describes the heat bath; the kinetic energy T and the potential energy U are then defined as

$$T(s) = \frac{1}{2}Q\dot{s}^2$$

$$V(s) = (m + 1)k_B T \log s$$

- The new Hamiltonian is given by

$$H_{\text{new}} = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m_i s^2} + \frac{1}{2}Q\dot{s}^2 + V(\vec{x}_1, \dots, \vec{x}_N) + V(s) \quad (41)$$

- The equation of motions are defined as

$$Q\ddot{s} = \sum_{i=1}^N m_i s \dot{x}_i^2 - \frac{1}{s}(m + 1)k_B T \quad (42)$$

- Choice of the mass Q affects the convergence toward the canonical distribution
 - Q is too large \implies equilibration is too slow
 - $Q \rightarrow \infty \implies$ recovers microcanonical molecular dynamics, but the goal is to simulate a canonical ensemble
 - Q is too small \implies temperature exhibits spurious oscillations. Use $\overline{\Delta T} = \sqrt{2/(Nd)} \overline{T}$ as the width of the temperature distribution (d is the dimension, N is the number of particles)
- Hoover proved in 1985 that the Nosé-Hoover thermostat is the only method with a single friction parameter that gives the canonical distribution

Lecture 5: Ewald Sums and Determination of Melting Temperatures

How to Fit Charges in a Classical Potential

- Strength of non-bonded forces (strongest first):
 1. Ionic interactions
 2. Hydrogen bonds
 - Describes the attraction between a slightly positive hydrogen on one molecule and a slightly negative atom on another molecules
 3. Dipole-dipole interactions
 - Describes the attraction between two polar molecules. H_2O is a polar molecule, since it has a slight positive charge on one side and a slight negative charge on the other (see Fig. 10)

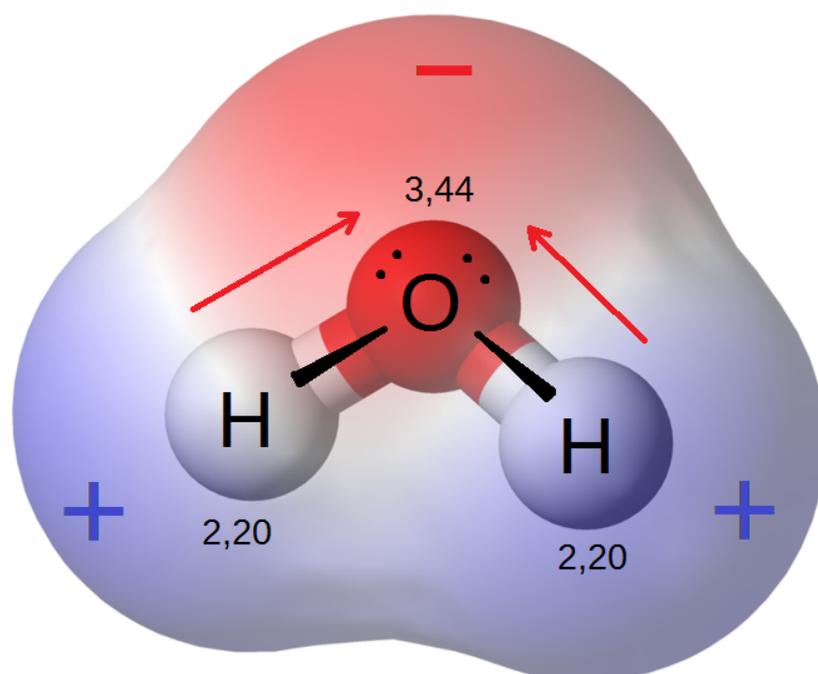


Figure 10: The water molecule is made up of oxygen and hydrogen, with respective electronegativities of 3.44 and 2.20. The dipoles from each of the two bonds (red arrows) add together to make the overall molecule polar.

Multipole Expansion

- Multipoles are used in computational chemistry to predict the electric potential field due to a complex molecule
- The potential in general is given by

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \rho(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (43)$$

- By considering an arbitrary charge distribution $\rho(\mathbf{r}')$, we wish to find the electrostatic potential due to this charge distribution at a given point \mathbf{r} , which is at a large distance from the given charge distribution, i.e., $\mathbf{r} \gg \mathbf{r}'$
- Multipoles are used to expand the $1/|\mathbf{r} - \mathbf{r}'|$ term for $\mathbf{r} \gg \mathbf{r}'$ and are based on Legendre polynomials:

$$\frac{1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{r} \frac{1}{\sqrt{1 + r'^2/r^2 - 2r'/r \cos \theta'}} = \sum_{n=0}^{\infty} P_n(\cos \theta') \left(\frac{r'}{r}\right)^n \quad (44)$$

- The potential in Eq. (43) can be rewritten using the multipole expansion presented in Eq. (44):

$$V(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_{n=0}^{\infty} \frac{1}{r^{n+1}} \int \rho(\mathbf{r}') P_n(\cos \theta') r'^n d^3\mathbf{r}' \quad (45)$$

- $n = 0$ term is the monopole moment (point charge)
- $n = 1$ term is the dipole moment
- $n = 2$ term is the quadrupole moment
- Classes of multipole expansion:
 1. Central multipole expansion: interaction centers are located at the center of mass
 2. Distributed multipole expansion: point charges, dipoles, quadrupoles, etc. are distributed throughout the molecules

Ewald Summation Method for Long Range Forces

- The potential energy is defined as

$$V_{\text{Coulomb}} = \frac{1}{2} \sum_{\mathbf{n}}' \left(\sum_{i=1}^N \sum_{j=1}^N \frac{z_i z_j}{|\mathbf{r}_{ij} + \mathbf{n}|} \right), \quad (46)$$

where the sum runs over all lattice points \mathbf{n} and the prime indicates that we omit $i = j$ for $\mathbf{n} = \mathbf{0}$

- Eq. (46) is conditionally convergent, i.e., it depends on the order by which the sum over the images is performed
- Improve convergence of Eq. (46) by splitting $1/r$:

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1-f(r)}{r} \quad (47)$$

– Choose f such as:

1. $\frac{f(r)}{r}$ is short-ranged, i.e., it is negligible beyond a certain cutoff
 2. $\frac{1-f(r)}{r}$ is a long-ranged and slowly varying contribution that can be Fourier transformed
- The total Coulomb energy consists of the following components:
 1. + Compensating cloud
 2. - Self interacting term
 3. + Short range contribution
 4. + Long range contribution
 5. - Charged system

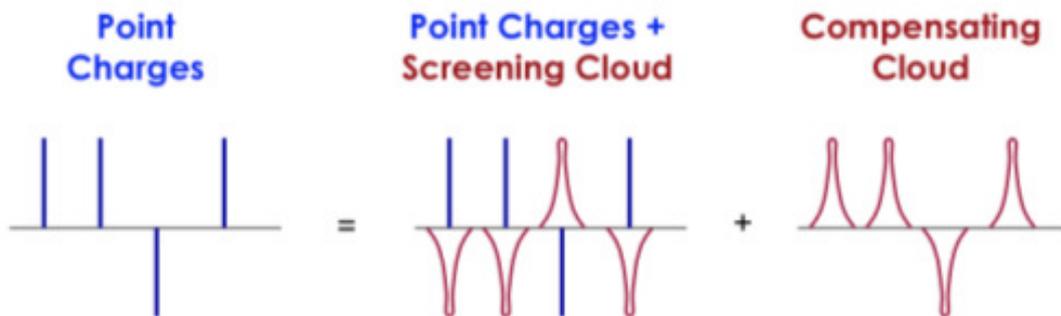


Figure 11: For the point charges plus screening cloud, the splitting of the potential results in a singular and short-ranged potential; using the compensating cloud, the splitting of the potential results in a smooth and long-ranged function which can be efficiently calculated using a Fourier transform. The screened charge is a Gaussian of width α ; with the presence of α , the electrostatic potential decays rapidly.

Determination of the Melting Temperature from MD Simulations

- Set up large sample consisting of approx. 50% solid and 50% liquid

- Melt half of the system (liquid) by running a NPT simulation coupled to a Nosé-Hoover thermostat and keep the other half of the system fixed (solid)
- Goal: Try to establish an equilibrium state where solid and liquid can coexist. If achieved, the temperature of the state has to be T_m by definition
- Let T_0 be the initial temperature of the system. Run a constant energy MD simulation (NPE is suggested by Ercolessi, since the melting temperature depends on the pressure of the system) and observe the change in kinetic energy
 1. Assume $T_0 > T_m$. The liquid is favored over the solid, i.e., some material melts. The corresponding latent heat of melting (the introduction of heat cannot be observed as a temperature change, as the temperature remains constant during the process, therefore it is *latent*) is absorbed by the system; the latent heat is the provided potential energy needed to break the bonds (the latent heat has to be greater than the bond energy). Since the total energy is conserved, the absorption of latent heat automatically implies a decrease of the kinetic energy, considering that the potential energy has to increase for the breaking of the bonds. The decrease of kinetic energy results in a decrease of the temperature, which then converges exponentially to T_m from above
 2. Assume $T_0 < T_m$. The inverse reasoning applies, i.e., latent heat is released and converted into kinetic energy, and the temperature converges exponentially to T_m from below
- In the exercise, we ran a NPT MD simulation for the melting of the system and multiple NPE MD simulations with different energy values to determine the energy E_m believed to be in the proximity of the melting jump in the caloric curve. Both simulations run with constant pressure and a variable volume. In this case, the volume of the box automatically changes as material melts or crystalizes, to accomodate the difference in density between two phases at constant pressure

Lecture 6: Ensembles Beyond NVE and Free Energy Differences

Ensembles

- In a thermodynamical equilibrium, the properties of a system do not change with time and the equilibrium state is *history-independent*, i.e., the properties of the system in equilibrium do not depend on how it was prepared but rather on the current conditions of the state (e.g., temperature, pressure, and density)
- Microcanonical ensemble: corresponds to an isolated system at equilibrium with constant E, V , and N . In this ensemble, each microstate that has the same energy as the specified energy E appears with equal probability

- The partition function expresses the density of states. For the microcanonical ensemble this is accomplished by running through every possible microstate n at a given V and N , and add 1 for each microstate that has energy equal to the specified energy:

$$Z(E, V, N) = \sum_{\text{all microstates } n \text{ at } V, N} \delta_{E_n=E} \quad (48)$$

- Microcanonical ensemble can be simulated using Molecular Dynamics, since integrating Newton's equations of motions results in energy conservation
- Canonical ensemble: occurs when a system with fixed V and N is at constant temperature (connected to a heat bath). In the canonical ensemble, the probability of each microstate n is proportional to $\exp\{-\beta E_n\}$; the microstates are distributed following Boltzmann's law
 - The canonical partition function is given by

$$Z(T, V, N) = \sum_{\text{all microstates } n \text{ at } V, N} \exp\{-\beta E_n\} \quad , \quad (49)$$

where $\beta = \frac{1}{k_B T}$

- Microstate probabilities are defined as

$$p_m = \frac{\exp\{-\beta E_m\}}{Z(T, V, N)} \quad (50)$$

- Canonical ensemble average:

$$\langle A \rangle = \frac{1}{Z(T, V, N)} \int A(\mathbf{r}^N, \mathbf{p}^N) \exp\{-\beta H(\mathbf{r}^N, \mathbf{p}^N)\} d\mathbf{r}^N d\mathbf{p}^N \quad (51)$$

- * $6N$ -dimensional integral that can be efficiently evaluated by using Monte Carlo simulations
- * The Metropolis method generates samples according to a given probability distribution. In the case of the canonical ensemble, we want to generate samples that follow the Boltzmann distribution
- Canonical ensemble can be simulated using Metropolis Monte Carlo (MMC), since MMC generates samples that follow the Boltzmann distribution

Metropolis Monte Carlo

1. Establish an initial microstate
2. Make a random trial change in the microstate

3. Compute $\Delta E = E_{\text{trial}} - E_{\text{old}}$
4. If $\Delta E \leq 0$, accept the new microstate
5. If $\Delta E > 0$:
 - a) Compute $w = \exp\{-\beta\Delta E\}$
 - b) Generate a random number r in the unit interval
 - c) If $r \geq w$, accept the new microstate; else retain the previous microstate
6. Determine the value of the desired physical quantities
7. Repeat steps (2) through (6) to obtain a sufficient number of microstates
8. Periodically compute averages over microstates
 - To keep a population that is in equilibrium still in equilibrium, *detailed balance* is enforced:

$$P_{\text{eq}}(Y)W(Y \rightarrow X) = P_{\text{eq}}(X)W(X \rightarrow Y) \quad (52)$$

This condition also ensures that the algorithm generates a sequence of states distributed according to the Boltzmann distribution

- The transition probability W is given by

$$W(X \rightarrow Y) = T(X \rightarrow Y) \cdot A(X \rightarrow Y) \quad (53)$$

* Properties of the transition probability:

1. Ergodicity: $\forall X, Y : W(X \rightarrow Y) > 0$, i.e., every configuration can be sampled
2. Normality: $\sum_Y W(X \rightarrow Y) = 1$
3. Homogeneity: $\sum_Y p_{\text{st}}(Y)W(Y \rightarrow X) = p_{\text{st}}(X)$

- The acceptance probability A is defined as

$$A(X \rightarrow Y) = \min\left(1, \frac{P_{\text{eq}}(Y)}{P_{\text{eq}}(X)}\right) \quad (54)$$

* Boltzmann: $P_{\text{eq}}(Y) = 1/Z(T, V, N) \exp\{-\beta E(X)\}$

Free Energy and Free Energy Differences

- The free energy is a thermodynamic state quantity, i.e., it is a property of a system that depends only on the current state of the system, and not on the way in which the system acquired that state (e.g., trajectories)
 - A state function describes the equilibrium state of a system
- The Helmholtz free energy A is defined as

$$A = -\frac{1}{\beta} \ln Z(T, V, N) \quad (55)$$

- The Helmholtz free energy is the maximum amount of work a system can do at constant volume and constant temperature, i.e., in the canonical ensemble
- The Gibbs free energy is the maximum amount of work a system can do at constant pressure and constant temperature, i.e., in the isothermal-isobaric ensemble
- The Helmholtz free energy in Eq. (55) can be rewritten to

$$A = -\frac{1}{\beta} \ln \langle \exp\{-\beta H(\mathbf{x}, \mathbf{p})\} \rangle \quad , \quad (56)$$

where H is the Hamiltonian describing a system with N interacting particles. This expression is rather difficult to evaluate, since it is a $6N$ -dimensional integral (can be calculated for very simple systems for which an analytical expression for the partition function can be obtained)

- For practical applications, it generally suffices to calculate relative free energies

Free Energy Perturbation

- The free energy difference for going from state A to state B is obtained from the Zwanzig equation:

$$\Delta A_{A \rightarrow B} = A_B - A_A = -\frac{1}{\beta} \ln \langle -\beta [H_B(\mathbf{r}, \mathbf{p}) - H_A(\mathbf{r}, \mathbf{p})] \rangle_A \quad , \quad (57)$$

where $H_A(\mathbf{x}, \mathbf{p})$ and $H_B(\mathbf{x}, \mathbf{p})$ are the Hamiltonians of states A and B , respectively, and $\langle \cdot \rangle_A$ denotes an ensemble average over a simulation run for state A

- Free energy perturbation calculations only converge properly when the difference between the two states is small enough

Thermodynamic Integration

- Thermodynamic integration is one of the most common methods for calculating free energy differences
- Calculation of the free energy difference:

1. Introduce the coupling parameter $\lambda \in [0, 1]$ and define the following Hamiltonian:

$$H(\mathbf{r}, \mathbf{p}; \lambda) = \begin{cases} H_A(\mathbf{r}, \mathbf{p}) & \text{if } \lambda = 0 \\ H_B(\mathbf{r}, \mathbf{p}) & \text{if } \lambda = 1 \end{cases} \quad (58)$$

- Example: $H_1 = H_0 + \lambda(H_1 - H_0)$

2. The free energy difference can now be defined as

$$\begin{aligned}\Delta A_{A \rightarrow B} &= A(\lambda = 1) - A(\lambda = 0) \\ &= \int_0^1 \frac{dA(\lambda)}{d\lambda} d\lambda = \int_0^1 \left\langle \frac{\partial H(\mathbf{r}, \mathbf{p}; \lambda)}{\partial \lambda} \right\rangle_{\lambda} d\lambda\end{aligned}\quad (59)$$

3. The last integral in Eq. (59) is evaluated using numerical quadrature.

- Although TI is very accurate, it has its disadvantages:
 1. Many simulations are needed for the calculation of the integral in Eq. (59), since every quadrature point requires the performance of a simulation
 2. For every quadrature point $\lambda_i \in (0, 1)$, the resulting state is unphysical and irrelevant for anything else but the free-energy calculation
 3. For each new pair of states A, B a new set of simulations is required
 4. Simulations at each quadrature point need some equilibration time (very time consuming)

Potential of Mean Force

- Reaction coordinate is an abstract one-dimensional coordinate which represents progress along a reaction pathway. It is usually a geometric parameter that changes during the conversion of one or more molecular entities, e.g., bond length and bond angle, or for more complex reactions the bond order
- Reaction coordinates are often plotted against free energy to demonstrate in some schematic form the potential energy profile
- The potential of mean force (PMF) is the potential that gives the average force over all the configurations of all the $n + 1, \dots, N$ particles acting on a particle j at any fixed configuration keeping fixed a set of particles $1, \dots, n$
- Find the likelihood that the system is in a configuration that yields a particular value of the order parameter. Taking the logarithm of this yields the free energy profile, which is known as the PMF
 - The order parameter is some observable physical quantity that is able to distinguish between two distinct phases, e.g., the net magnetization in a ferromagnetic system undergoing a phase transition
- The PMF $W(\xi)$ is defined from a Boltzmann weighted average $\langle p(\xi) \rangle$,

$$W(\xi) = W(\xi^*) - k_B T \ln \left[\frac{\langle p(\xi) \rangle}{\langle p(\xi^*) \rangle} \right], \quad (60)$$

with

$$\langle p(\xi) \rangle = \frac{\int d\mathbf{R} \delta(\xi'[\mathbf{R}] - \xi) \exp\{-\beta U(\mathbf{R})\}}{\int d\mathbf{R} \exp\{-\beta U(\mathbf{R})\}} \quad , \quad (61)$$

where $U(\mathbf{R})$ represents the total energy of the system and $\xi'[\mathbf{R}]$ is a function depending on a few or several degrees of freedom (e.g., $\xi'[\mathbf{R}]$ may be an angle, a distance, etc.) and ξ being the reaction coordinate

- Generate a series of configurations along the reaction coordinate ξ using Umbrella sampling, conducting individual simulations at chosen intervals
 - Umbrella sampling has the advantage that it helps bridging energy gaps in situations where regions are separated by free energy barriers (e.g., a solid at its melting point. The order parameter Q describes the state of the system, i.e., liquid has a low Q and solid a high Q ; both phases are low in energy, but are separated by a free energy barrier at intermediate values of Q)

Lecture 7: Including Electrons Part 1 (Hartree-Fock)

Electronic Structure Problem

- The exact non-relativistic, time independent molecular Hamiltonian is given by

$$H = H_e + T_n \quad , \quad (62)$$

with

$$H_e = - \sum_i \frac{1}{2} \nabla_i^2 - \sum_{i,A} \frac{Z_A}{r_{iA}} + \sum_{i>j} \frac{1}{r_{ij}} + \sum_{A>B} \frac{Z_A Z_B}{R_{AB}} \quad (63)$$

and

$$T_n = - \sum_A \frac{1}{2M_A} \nabla_A^2 \quad , \quad (64)$$

where \mathbf{r} are the position vectors of the electrons and \mathbf{R} the position vectors of the nuclei

- The general solution (wave function) is $(3n + 3N)$ -dimensional and can only be solved for very few particles, e.g., H_3^+
- Electronic structure problem is one of the two steps to study quantum mechanical motion of a molecular system and arises from the Born-Oppenheimer approximation: the nuclei of atoms are so much heavier than the electrons that we can view them as classical particles and can consider them as stationary for the purpose of

calculating the properties of electrons; the Hamiltonian operator for the electrons thus becomes

$$H_{\text{Electrons}} = \sum_{i=1}^N \left(-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}_i) \right) + \sum_{i < j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \quad , \quad (65)$$

where the potential of the M atomic nuclei with charges $Z_i e$ at the locations \mathbf{R}_i is given by

$$V(\mathbf{r}) = -e^2 \sum_{i=1}^M \frac{Z_i}{|\mathbf{R}_i - \mathbf{r}|} \quad (66)$$

- The Born-Oppenheimer approximation is the separation of electronic and nuclear motion. The wave function of the molecule thus becomes

$$\Psi_{\text{Molecule}} = \Psi_{\text{Electron}} \cdot \Psi_{\text{Nuclei}} \quad . \quad (67)$$

We solve for the electronic ground-state first, and then calculate the energy of the system in that configuration and solve for the nuclear motion

- The natural state of a system is to use the lowest energy possible (principle of minimum energy), thus we are looking for the wave function Ψ of the ground state, i.e., the lowest energy wave function Ψ
- The average energy of a system in state Ψ is given by

$$E[\Psi] = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \quad , \quad (68)$$

where

$$\langle \Psi | H | \Psi \rangle = \int \Psi^* H \Psi d\mathbf{x} \quad . \quad (69)$$

- Full minimization of the functional $E[\Psi]$ in Eq. (68) will give the true ground state wave function Ψ_0 and energy $E[\Psi_0] = E_0$. The variational principle states that the energy computed from the trial state Ψ is an upper bound to the true ground-state energy E_0 , i.e.,

$$E_0 \leq E[\Psi] \quad (70)$$

Slater Determinant

- Describes the wave function of a multi-fermionic system
- Antisymmetrized and normalized product of N single-particle wave functions Ψ_i (also called *molecular orbitals*) is given by the Slater determinant

$$\Psi_{SD}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \Psi_1(\mathbf{x}_1) & \dots & \Psi_N(\mathbf{x}_1) \\ \vdots & & \vdots \\ \Psi_1(\mathbf{x}_N) & \dots & \Psi_N(\mathbf{x}_N) \end{vmatrix} \quad (71)$$

- The fermionic many-body wave function can be
 1. an infinite sum of Slater determinants (full configuration interaction)
 2. a single Slater determinant (Hartree Fock theory)

Basis Functions

Electron Gas

- The electron-electron interactions are completely neglected, i.e., $V(\mathbf{r}_i) = 0 \forall i$ in Eq. (66)
- Ideal choice for basis functions are plane waves

$$\psi_{\mathbf{k}}(\mathbf{r}) = \exp\{-i\mathbf{k}\mathbf{r}\} \quad (72)$$

Atoms and Molecules

- Slater-Type-Orbitals (STO): consist of a product between a radial ($R(r)$) and angular wave function ($Y_{l,m}(\theta, \phi)$)
 - Advantage: these wave functions have the correct asymptotic radial dependence and the correct angular dependence
 - Disadvantage: the matrix elements of the Hamiltonian do not have closed form solutions
- Gauss-Type-Orbitals (GTO): use Gaussian functions
 - Advantage: the matrix elements of the Hamiltonian have closed form solutions, since Gaussians can be easily integrated
 - Disadvantage: Non-orthogonal

Electrons in Solids

- Use linear augmented plane waves (LAPW): smoothly cross over from localized wave function behaviour near the nuclei to plane waves in the region between the atoms

The Hartree-Fock Theory

- The Hartree-Fock approximation is based on the assumption of independent electrons, i.e., multi-fermionic system
- Ansatz: Suppose that the N -particle ground state wave function Ψ_0 is approximated as an antisymmetrized product of N single-particle wave functions Ψ_i (also

known as *molecular orbitals*) called a Slater determinant:

$$\Psi_0 \approx \Psi_{HF}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \Psi_1(\mathbf{x}_1) & \dots & \Psi_N(\mathbf{x}_1) \\ \vdots & & \vdots \\ \Psi_1(\mathbf{x}_N) & \dots & \Psi_N(\mathbf{x}_N) \end{vmatrix} \quad (73)$$

– The spin orbitals Ψ_i in Eq. (73) are expanded using basis functions Φ_j :

$$\Psi_i = \sum_j^M c_{ij} \Phi_j \quad . \quad (74)$$

- The Hartree-Fock approximation is the method whereby the molecular orbitals Ψ_i are found that minimize the energy for the Slater determinant Ψ_{HF} shown in Eq. (73), i.e.,

$$E_{HF} = \min_{\Psi_{HF}} E[\Psi_{HF}] \quad , \quad (75)$$

where the functional E is defined in Eq. (68)

- The functional shown in Eq. (75) is minimized with respect to the expansion coefficients, c_{ij} (Eq. (74))
- The exact solutions of the Hartree-Fock equations correspond to the *exact* Hartree-Fock spin orbitals and can only be calculated for atoms
- The introduction of a basis set (Eq. (74)) transforms the Hartree-Fock equations to a set of matrix equations called Hartree-Fock-Roothan equations (Eq. (78)). Only as the basis set approaches completeness, i.e., as one approaches the Hartree-Fock limit, will the spin orbitals that one obtains approach the exact Hartree-Fock spin orbitals

The Hartree-Fock Method

1. Hartree-Fock wave function (Slater determinant) can be written in second quantized form

$$|\Psi\rangle = \prod_{\mu,\sigma} c_{\mu\sigma}^\dagger |0\rangle \quad , \quad (76)$$

where $c_{\mu\sigma}^\dagger$ is orthogonal and creates an electron in the orbital $\phi_\mu(\mathbf{r}, \sigma)$

2. Expand $c_{\mu\sigma}^\dagger$ in terms of creation operators $\hat{a}_{n\sigma}^\dagger$, which are not necessarily orthonormal:

$$c_{\mu\sigma}^\dagger = \sum_{n=1}^L d_{\mu n} \hat{a}_{n\sigma}^\dagger \quad (77)$$

where \hat{a}^\dagger can be a STO or GTO, with σ being the spin and n being the orbital

3. Minimize the energy E_0 by changing the $d_{\mu\sigma}$ under the condition that the $|\Phi\rangle_\mu$ are normalized; use Lagrange multipliers to enforce the normalization constraint
4. The result are the Hartree-Fock-Roothan equations, which have the form of a generalized eigenvalue problem, i.e.,

$$\underline{\mathbf{A}}[\underline{\mathbf{C}}] \underline{\mathbf{C}} = \underline{\mathbf{B}} \underline{\mathbf{C}} \underline{\boldsymbol{\lambda}} \quad , \quad (78)$$

where $\underline{\mathbf{A}}$ is the Fock matrix, $\underline{\mathbf{B}}$ the overlap matrix, $\underline{\mathbf{C}}$ the matrix containing the molecular orbital coefficients c_{ij} , and $\underline{\boldsymbol{\lambda}}$ a diagonal matrix of the orbital energies λ_i ; $\underline{\mathbf{A}}$ depends on the solution $\underline{\mathbf{C}}$ (self-consistent field theory)

- The self-consistent procedure to solve Eq. (78) is given as:
 - a) Assume initial set of molecular orbitals $\underline{\mathbf{C}}$
 - b) Construct $\underline{\mathbf{A}}$ from $\underline{\mathbf{C}}$
 - c) Diagonalize $\underline{\mathbf{A}}$ to get new $\underline{\mathbf{C}}$
 - d) Go back to (b) if needed
- The equations presented in Eq. (78) are called the *Hartree-Fock-Roothan* equations and are solved iteratively until convergence to a fixed point is achieved
- Computational complexity of the self-consistent field method is $O(M^3)$, where M is the size of the basis set
 - For the matrices $\underline{\mathbf{A}}, \underline{\mathbf{B}} \in \mathbb{K}^{M \times M}$, the full diagonalization takes $O(M^3)$ time
 - For the matrix $\underline{\mathbf{C}} \in \mathbb{K}^{M \times N}$, verifying the orthonormality takes $O(M^2 N)$ or $O(MN^2)$ time

Lecture 8: Including Electrons Part 2 (Density Functional Theory)

Density Functional Theory

- Many-body wave function living in \mathbb{R}^{3N} is replaced by the electron density, which lives in \mathbb{R}^3
- Theoretically, DFT is an exact theory, if the exchange-correlation functional $E_{xc}[\rho]$ would be known

Hohenberg-Kohn Theorems

1. The ground-state energy E_0 of an electronic system in an external potential V is a functional of the electron density $\rho(\mathbf{r})$:

$$E_0 = E[\rho] = F[\rho] + \int d\mathbf{r} V(\mathbf{r})\rho(\mathbf{r}) \quad , \quad (79)$$

with an unknown universal functional $F[\rho]$

- The ground-state energy can be obtained variationally: the density that minimises the total energy is the exact ground-state density, i.e. the density of the ground state wave function minimizes the functional E_0 shown in Eq. (79). It also confirms the existence of the universal functional $F[\rho]$

The DFT Method

- Start with Ansatz

$$F[\rho] = \underbrace{E_h[\rho]}_{\text{Hartree term}} + \underbrace{E_k[\rho]}_{\text{Kinetic term}} + \underbrace{E_{xc}[\rho]}_{\text{Exchange-correlation term (unknown)}} \quad (80)$$

- E_h is given by the Coulomb repulsion between two electrons (or Hartree energy)

$$E_h = \int d\mathbf{r} \int d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \quad (81)$$

- E_k is the kinetic energy of a non-interacting electron gas with the same density ρ
 - E_{xc} is the unknown contribution
- To calculate the ground state density we have to minimize the energy given in Eq. (79):

$$\partial E[\rho] = 0 \quad (82)$$

Local Density Approximation (LDA)

- Approximate the potential v_{xc} arising from the functional $E_{xc}[\rho]$ by replacing v_{xc} with the potential of a uniform gas with the same density. In this case, we ignore the fact that the functional E_{xc} may depend on $\rho, \nabla\rho, \nabla^2\rho, \dots$ and assume that it takes the local density, i.e., it does not take into account the non-homogeneity of the true electron density:

$$E_{xc}^{\text{LDA}}[\rho](\mathbf{r}) = \int \rho(\mathbf{r}) \epsilon_{xc}(\rho(\mathbf{r})) d\mathbf{r} \quad , \quad (83)$$

where ϵ_{xc} is the exchange-correlation energy per particle of a uniform electron gas of density $\rho(\mathbf{r})$

- The exchange-correlation energy can be written as

$$\epsilon_{xc}(\rho(\vec{r})) = \epsilon_x(\rho(\vec{r})) + \epsilon_c(\rho(\vec{r})) \quad (84)$$

* ϵ_x was derived by Slater and ϵ_c was proposed by Wigner

Generalized Gradient Approximation (GGA)

- GGA not only uses the local density, but also takes into account the non-homogeneity of the true electron density by supplementing the density with information about the gradient of the charge density, $\nabla\rho(\mathbf{r})$:

$$E_{xc}^{\text{GGA}}[\rho](\mathbf{r}) = \int f(\rho_{\uparrow}, \rho_{\downarrow}, \nabla\rho_{\uparrow}, \nabla\rho_{\downarrow}) d\mathbf{r} \quad , \quad (85)$$

Lecture 9: Including Electrons Part 3

- Short-comings of GGA:
 1. Missing van der Waals interactions
 2. Self-interaction error
 3. Strong correlations
- One of the fundamental assumptions of quantum chemistry is that an electron does not interact with itself. For DFT, this leads to a simple condition on the exact (and unknown) exchange-correlation functional: for any one-electron density distribution, the exchange-correlation energy must identically cancel the Coulomb self-interaction energy of the electron cloud
 - In GGA DFT the cancellation between both energies is imperfect, i.e. $E_{\text{Coulomb}} - E_{xc} \neq 0$
- A band gap is an energy range in a solid where no electron states can exist
 - GGA DFT underestimates the band gaps of materials, whereas Hartree-Fock overestimates them
- Since the exchange-correlation energy E_{xc} can be written as a sum of the exchange energy E_x and the correlation energy E_c (see Eq. (84)), a hybrid exchange-correlation functional is usually constructed as a linear combination of the Hartree-Fock exact exchange functional E_x^{HF} and any number of exchange and correlation explicit density functionals

$$E_{xc}^{\text{Hybrid}} = (1 - a_x)E_x^{\text{HF}} + a_x E_x^{\text{DFT}} + E_c^{\text{DFT}} \quad (86)$$

- The popular B3LYP (Becke, three parameter, Lee-Yang-Parr) exchange-correlation functional is

$$E_{xc}^{\text{B3LYP}} = E_x^{\text{LDA}} + a_0(E_x^{\text{HF}} - E_x^{\text{LDA}}) + a_x(E_x^{\text{GGA}} - E_x^{\text{LDA}}) + E_c^{\text{LDA}} + a_c(E_c^{\text{GGA}} - E_c^{\text{LDA}}), \quad (87)$$

- Conventional GGA density functionals do not contain the physics of the van der Waals interaction. The exchange part alone can yield anything from severe overbinding to severe over-repulsion depending on the choice of functional

- Although van der Waals interactions have a simple form and are completely missing, adding a term like C_6/r^6 is not suitable, since C_6 depends for example on the chemical environment of atoms
- To work around this problem, i.e., accounting for the van der Waals interactions, is to add a *force-field like correction* to the conventional Kohn-Sham DFT energy E_{DFT} :

$$E_{DFT-Disp} = E_{DFT} + E_{Disp} \quad (88)$$

- Grimme released a scheme called DFT-D (DFT-D2 and DFT-D3, depending on the level) that is a good approximation for the correction term E_{Disp} in Eq. (88) and provides parameters for all atom types (adapted to their chemical environment), has parameters optimized for a large set of density functionals and has three-body terms
 - * DFT-D, DFT-D2: Consider all pairs of atoms for the correction term, resulting in

$$E_{Disp} = -\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{L}}' \frac{C_{6ij}}{r_{ij,L}^6} f_{d,6}(r_{ij,L}) \quad , \quad (89)$$

where the summations are over all N atoms and all translations of the unit cell \mathbf{L} , and f is a damping function whose role is to scale the force field such as to minimize contributions from interactions within typical bonding distance

- * DFT-D3: is asymptotically exact for a gas of weakly interacting neutral atoms, considers three-body terms, provides a consistent description of all chemically relevant elements of the periodic system, uses geometry dependent dispersion coefficients and the approach is correct with all density functionals for finite systems (molecules) or nonmetallic infinite systems
- The Hartree-Fock method approximates the antisymmetric wave function by a single Slater determinant, but this approximation does not take into account the Coulomb correlation, leading to a total electronic energy different from the exact solution of the non-relativistic Schrödinger equation within the Born–Oppenheimer approximation; thus, we get the following order of energies:

$$E_{HF} > \underbrace{E_{HF \text{ limit}} > E_{\text{Post-HF methods}} > E_{\text{Exact solution of Schrödinger eq.}}}_{\text{Electron correlation energy}} > E_{\text{Relativistic energy}} \quad (90)$$

- Møller–Plesset perturbation theory (MP) improves the Hartree-Fock method by adding electron correlation effects by means of Rayleigh–Schrödinger perturbation theory (MP2: 2nd degree)

* Problems:

1. Computationally very demanding ($O(N^5)$)
 2. Requires much larger basis sets than GGA DFT calculations
 3. Tends to overestimate van der Waals interactions
- Double-Hybrid DFT: Use the MP2 energy expression to provide part of the correlation in an exchange-correlation functional, i.e.,

$$E^{\text{D-Hyb}_{xc}} = (1 - a_x)E_x^{\text{HF}} + a_x E_x^{\text{DFT}} + (1 - a_c)E_c^{\text{PT2}} + a_c E_c^{\text{DFT}} \quad (91)$$

- The MP2 expression can be computed in $O(N^4)$ time
- Double-Hybrid DFT outperforms MP2 (with reduced computational costs, too)

Lecture 10: Interaction with Matter and Electronic Spectroscopies

The Scanning Tunnelling Microscope

- Scanning tunnelling microscope: instrument for imaging surfaces at the atomic level (by probing the electron density of the sample) and is based on the concept of quantum tunneling

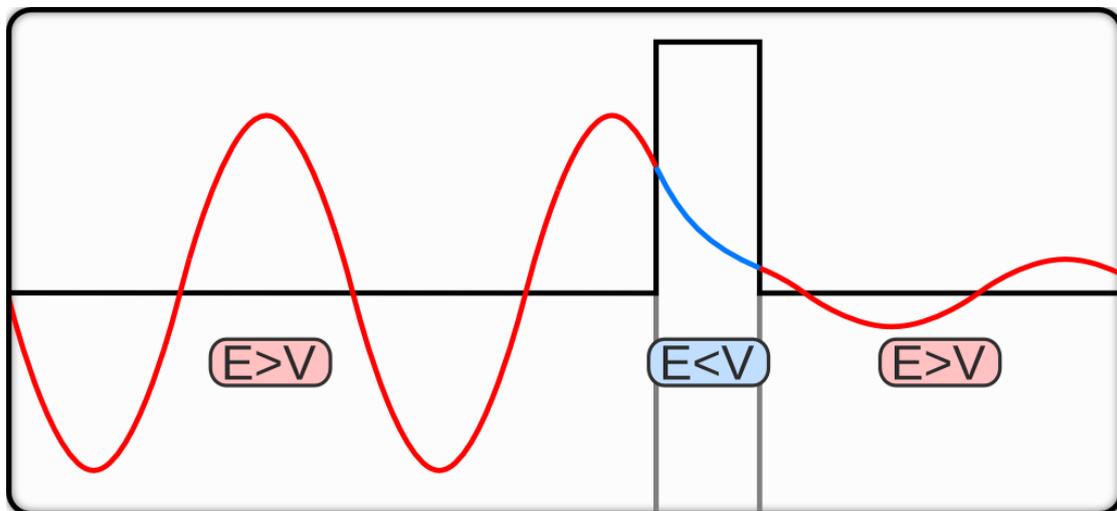


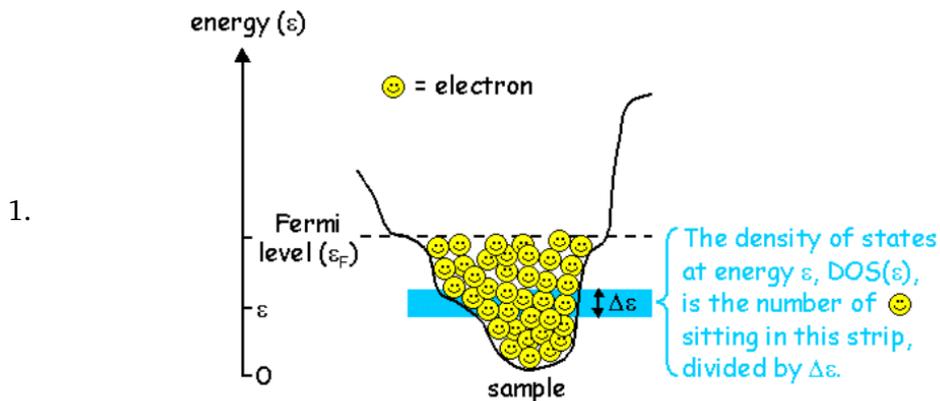
Figure 12: Quantum tunnelling through a barrier. The energy of the tunnelled particle is the same but the amplitude is decreased.

Procedure

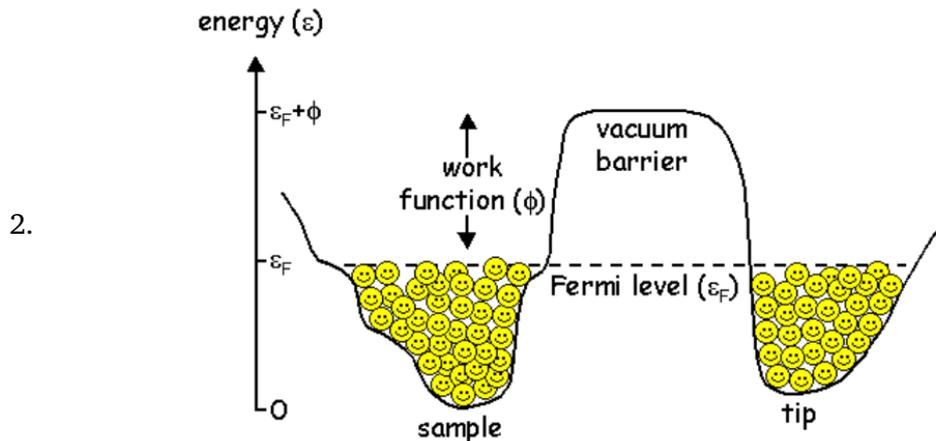
1. Voltage bias is applied and tip is brought close to the sample
 - Tip-sample separation W is typically in the $4 - 7\text{\AA}$ range, which is the equilibrium position between attractive ($3 < W < 10\text{\AA}$) and repulsive ($W < 3\text{\AA}$) interactions
2. The voltage bias will cause electrons to tunnel between the tip and sample, creating a current that can be measured
 - Data is obtained from the resulting changes in current
3. There exist two modes for imaging the surface:
 - a) Constant current mode: feedback electronics adjust the height by a voltage to the piezoelectric height control mechanism. This leads to a height variation and thus the image comes from the tip topography across the sample and gives a constant charge density surface
 - b) Constant height mode: the voltage and height are both held constant while the current changes to keep the voltage from changing; this leads to an image made of current changes over the surface, which can be related to charge density

Tunneling and Available States

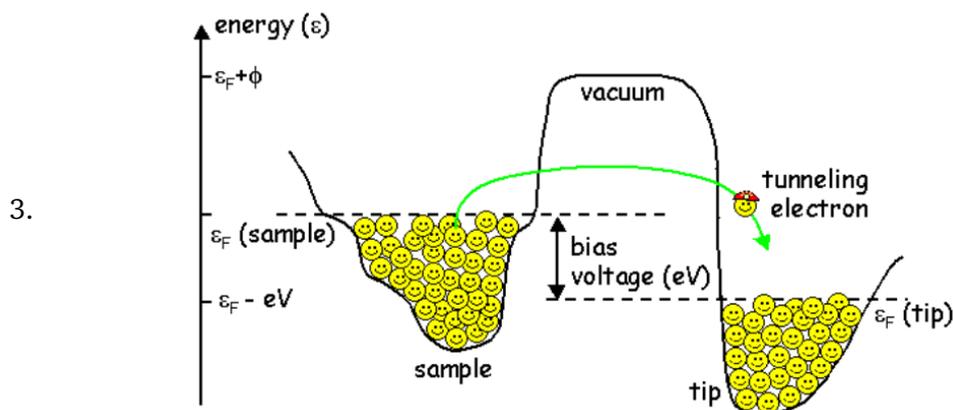
- The Fermi energy is the maximum energy occupied by an electron at $0K$. By the Pauli exclusion principle, we know that the electrons will fill all available energy levels, and the top of that Fermi sea of electrons is called the Fermi energy or Fermi level.



The electrons fill up the energy valley in the sample until there are no more electrons. The top energy level is called the Fermi level, ϵ_f



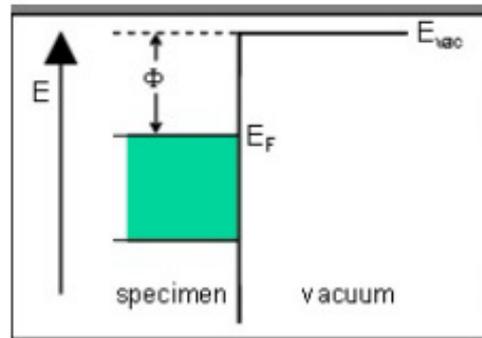
The electrons in the tip and the sample are separated by the vacuum barrier. The height of the vacuum barrier is called the work function Φ . For an electron to climb the vacuum barrier we would need to supply a very large amount of energy. In this case we do not need to supply the energy, since quantum mechanics tells us that the electron can tunnel right through the barrier. For the tunneling to happen, we would need empty states available, but as long as the tip and the sample are held at the same electrical potential, their Fermi levels line up exactly, resulting in no empty states



By applying a bias voltage to the sample with no respect to the tip, we effectively raise the Fermi level of the sample with respect to the tip. Now there are empty states available for tunneling into

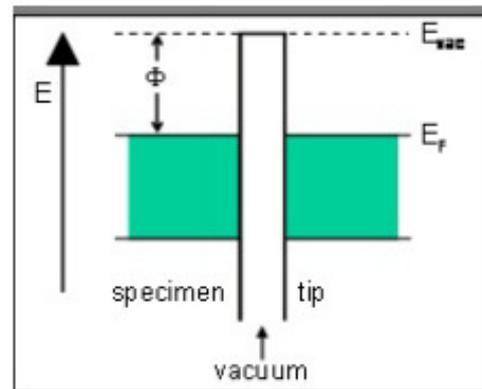
Tunneling Current

1.



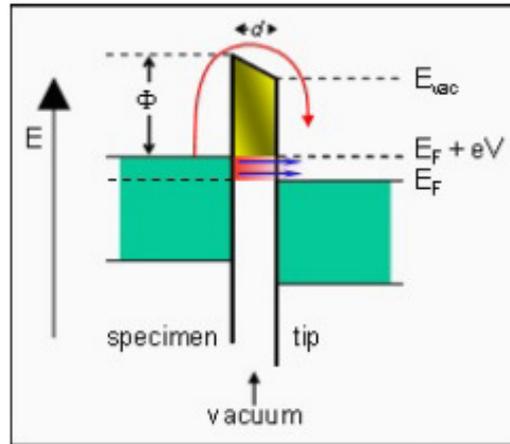
In a metal, the energy levels of the electrons are filled up to a particular energy, known as the Fermi energy E_F . In order for an electron to leave the metal, it needs an additional amount of energy Φ , the so-called work function

2.



When the specimen and the tip are brought close to each other, there is only a narrow region of empty space left between them. On either side, the electrons are present up to the Fermi energy. They need to overcome a barrier Φ to travel from tip to specimen or vice versa

3.



If the distance d between specimen and tip is small enough, electrons can *tunnel* through the vacuum barrier. When a voltage V is applied between specimen and tip, the tunneling effect results in a net electron current, in this example from specimen to tip. This is the tunneling current

- The tunneling current I depends on the tip-surface distance d , on the voltage V , and on the height of the barrier Φ (the work function), i.e.,

$$I = \text{const.} \cdot eV \exp\left\{-2\frac{\sqrt{2m\Phi}}{\hbar}d\right\} \quad , \quad (92)$$

where $\frac{\sqrt{2m\Phi}}{\hbar}$ is the decay length, e the electron charge and m the electron mass

Calculation of the Tunneling Current

- Fermi's golden rule is a result from first-order time-dependent perturbation theory, and it states that the transmission rate from an initial state $|i\rangle$ to a final state $|f\rangle$ is given by

$$R_{i \rightarrow f} = \frac{2\pi}{\hbar} |M_{f,i}|^2 \delta(E_i - E_f) \quad , \quad (93)$$

where $M_{f,i}$ is the matrix element of the perturbation potential between both states $|i\rangle$ and $|f\rangle$, and the δ -function ensures energy conservation

- The current is first given by

$$I = \frac{2\pi e}{\hbar} \sum_{\mu\nu} f(\epsilon_\mu) [1 - f(\epsilon_\nu + eV_b)] |M_{\mu,\nu}|^2 \delta(\epsilon_\mu - \epsilon_\nu) \quad , \quad (94)$$

where f is the Fermi-Dirac function, ϵ_μ the energy of the state of the tip Ψ_μ , and $M_{\mu,\nu}$ the tunneling matrix element between the tip-state χ_μ and the surface-state

Ψ_ν . For small temperatures and small V_b , the Fermi-Dirac distribution *disappears* and the Fermi-Dirac function f can thus be omitted, resulting in

$$I = \frac{2\pi e^2 V_b}{\hbar} \sum_{\mu\nu} |M_{\mu,\nu}|^2 \delta(\epsilon_\mu - E_F) \delta(\epsilon_\nu - E_F) \quad , \quad (95)$$

where E_F is the Fermi energy and the tunneling matrix elements $M_{\mu,\nu}$ describing the transition probability between tip χ_μ and surface states Ψ_ν are given by

$$M_{\mu,\nu} = \frac{\hbar^2}{2m} \int d\mathbf{S} (\chi_\nu^* \nabla \Psi_\mu u - \Psi_\mu u \nabla \chi_\nu^*) \quad (96)$$

Bardeen's Approach

- Bardeen's approach defines the form of the tunneling matrix elements $M_{\mu,\nu}$, which is presented in Eq. (96)

Tersoff-Hamann Theory

- The calculation of the transfer matrix element $M_{\mu,\nu}$ can be simplified by approximating the tip wave function χ_ν as a single s -orbital (spherical)
- Corrugation amplitudes of closed packed metal surfaces predicted by the Tersoff-Hamann approach (s wave tip) are too small to explain the measured atomic corrugations (which are much higher)
 - Julian Chen pointed out that this is partially due to the restriction to s -like tip orbitals. For example, a more directed tip state like a d_z -state pointing towards the sample might lead to higher values of the corrugation amplitude
 - Chen extended the Tersoff-Hamann model by introducing a derivative rule for other tip states

Modes of Operation

Constant Current Mode (CCM)

- By using a feedback loop the tip is vertically adjusted in such a way that the current always stays constant
- Current is proportional to the local density of states (LDOS); tip follows a contour of constant density of states during scanning, resulting in a topographic image of the surface

Constant Height Mode (CHM)

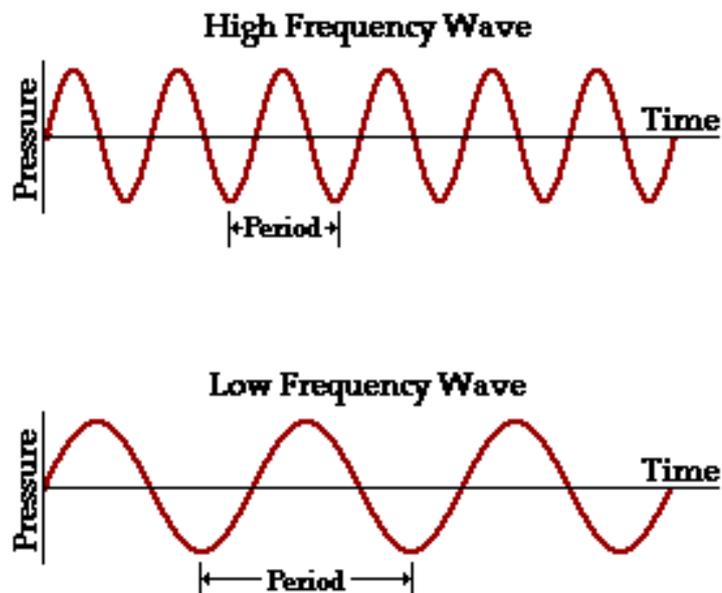
- The vertical position of the tip is not changed in this mode
- Since the adjusting of the surface height is not needed a higher scan speed can be obtained
- CHM can only be applied if the sample surface is very flat; surface corrugations higher than $5 - 10 \text{ \AA}$ will cause the tip to crash

Infrared Spectroscopy and Vibrational Modes

Lecture 3

- Spectroscopy is the study of interaction between matter and electromagnetic radiation. It uses electromagnetic radiation (light) to study matter
- Infrared (IR) spectroscopy is used for the identification, estimation and structure determination of compounds
 - IR spectrum of each molecule is unique and can thus be used as a signature or fingerprint to identify the molecule
- A spectrum is a range of possible frequencies
 - The electromagnetic spectrum is the range of all possible frequencies of electromagnetic radiation
- Wavelength λ [cm]: length of one complete wave cycle; also known as the period T
- Frequency ν $\left[\frac{\text{cycle}}{\text{sec.}} \right]$, [Hz]: number of wave cycles that pass through a point in one second
 - The frequency ν is inversely proportional to the period T :

$$\nu = 1/T \quad (97)$$



- For electromagnetic waves, the equation that relates wavelength λ and frequency ν is given by

$$\lambda\nu = c \quad , \quad (98)$$

where c is the speed of light

- Wavenumber $\bar{\nu}$ [cm^{-1}]:

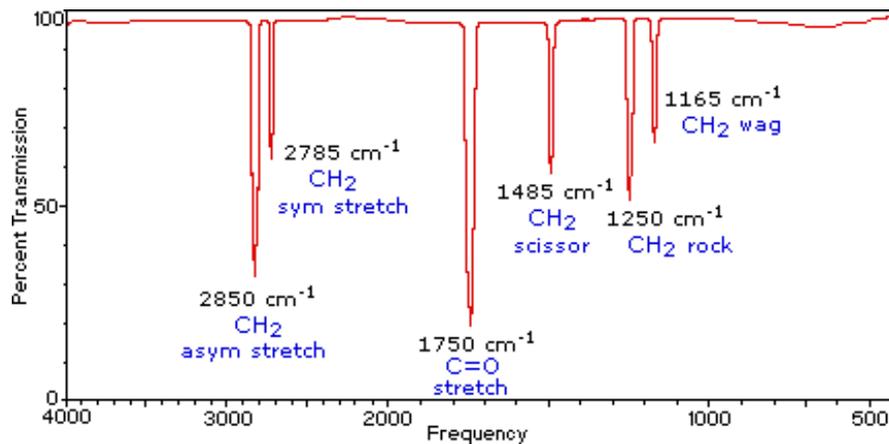
$$\bar{\nu} = \frac{1}{\lambda} \quad (99)$$

- Energy [J]:

$$E = h\nu = \frac{hc}{\lambda} = hc\bar{\nu} \quad (100)$$

- An increase in wavenumber corresponds to an increase in energy
- Infrared light imposed on a molecule does contain enough energy to interact with a molecule causing vibrational (infrared radiation, low energy) and rotational (microwave radiation, higher energy; not interesting) changes
- There are two types of molecular vibrations: stretching and bending
 - A molecule consisting of n atoms has:
 - * $3n - 6$ degrees of freedom (or fundamental vibrations / normal modes) if it is *nonlinear*
 - * $3n - 5$ degrees of freedom (or fundamental vibrations / normal modes) if it is *linear*

- Every fundamental vibration results in an absorption band (or range of wavenumbers) in the IR spectrum
- Molecular vibration is excited when the molecule absorbs a quantum energy E corresponding to the vibration's frequency ν according to the relation $E = h\nu$, i.e., absorption occurs when the radiant energy matches the energy of a specific molecular vibration (stretch or bend); the absorption is then measured and plotted as transmission probability (e.g., zero transmittance means total absorption) against the wavenumber, showing different absorption bands for the given molecular vibrations (see Fig. 13)

Figure 13: Gas phase infrared spectrum of formaldehyde ($\text{H}_2\text{C} = \text{O}$)

- The absorption wavenumber for a stretching vibration is given as

$$\bar{\nu} = \frac{1}{2\pi c} \sqrt{k \left(\frac{m_1 + m_2}{m_1 m_2} \right)} \quad (101)$$

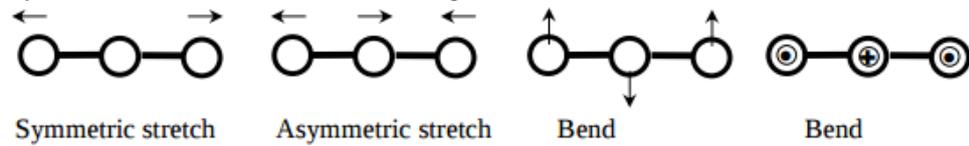
1. As the bond strength increases, the wavenumber increases
2. As the mass of one of the two atoms in the bond increases, the wavenumber decreases

Lecture 10

Theoretical Determination of Vibrational Spectrum: Classical Normal Mode Analysis and the Harmonic Approximation

- Vibrations of a molecule are given by its normal modes. Each absorption in a vibrational spectrum corresponds to a normal mode

- The four normal modes for carbon dioxide are the symmetric stretch, the asymmetric stretch and two bending modes:



- Linear molecules have $3N - 5$ normal modes, where N is the number of atoms; for non-linear molecules, the number of normal modes is $3N - 6$
- Each normal mode acts like a simple harmonic oscillator
 - The potential energy for Hooke's Law is obtained by integrating

$$F = -\frac{dV}{dx} = -kx \quad (102)$$

- Since the force constant k is unknown in molecular mechanics, it can be calculated from the second derivative of the potential energy:

$$k = \frac{d^2V}{dx^2} \quad (103)$$

- Hooke's Law force is substituted into Newton's Law, resulting in

$$m \frac{d^2x}{dt^2} = -kx \quad , \quad (104)$$

with the solution

$$x(t) = A \sin(2\pi\nu t) \quad , \quad (105)$$

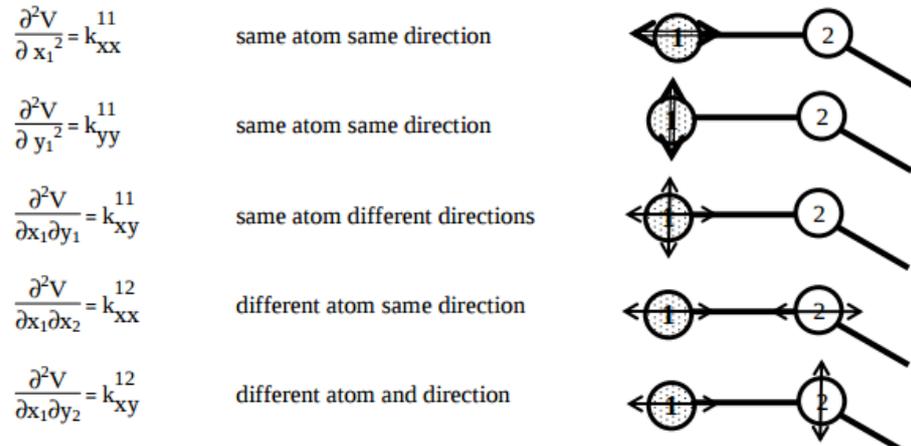
where A is the amplitude of the vibration and ν is the fundamental vibration frequency

- * MD and normal mode analysis (NMA) are quite similar, e.g., both include the kinetic and potential energy for the molecule and both calculate the Hessian and then integrate Newton's Laws of motion
- * The difference between MD and NMA is that the equations of motion are integrated numerically in MD simulations, but sinusoidal solutions are assumed for NMA
- Normal mode analysis:
 - Extensions of the atom coordinates are the differences in the positions and the equilibrium positions for that atom
 - * To get the equilibrium positions an energy minimization has to be run

- Use MD to find the potential energy V of the molecule as a function of the position of each atom

- * The second derivative of the potential V can be used to calculate the force constants

- * The various types of force constants are shown in the following figure:



- The complete list of these force constants is called the Hessian matrix, which is a $3N \times 3N$ matrix

IR Spectroscopy

- In an IR spectroscopy experiment we introduce light (an electromagnetic wave) which interacts with a molecule by disturbing the local electric field E_F
 - The energy of the photon must match the difference in energy between two vibrational states, i.e.,

$$E_{ph} = \left(n + 1 + \frac{1}{2}\right) h\nu - \left(n + \frac{1}{2}\right) h\nu = h\nu \quad , \quad (106)$$

where ν is the frequency of the photon (Planck-Einstein relation)

- In order to be IR active, i.e. a stretch or bend which can be absorbed in IR, a vibration must cause a change in the dipole moment of the molecule
- Static calculations of the vibrational normal modes in harmonic approximation in condensed phase are problematic, since they need to be started from an optimized minimum of the PES (molecules in the gas)
- All vibrational spectra calculated from MD simulations are based on the Fourier transform of certain autorrelation functions, e.g, power spectra is a Fourier transform of the particle velocity autocorrelation function and the IR spectra is a Fourier transform of the molecular dipole moment autocorrelation function

- Correlation functions: Given two time dependent properties of a system $A(t)$ and $B(t)$, find out if there is a cause-and-effect relationship operating between them, i.e., the value of one of them at a particular time t influences the value of the other at some time later. The correlation function $C(t)$ is defined as

$$C(t) = \langle A(0)B(t) \rangle \quad (107)$$

- The autocorrelation function of a time dependent property $A(t)$ is defined as

$$C(t) = \langle A(0)A(t) \rangle \quad , \quad (108)$$

or since the starting time is arbitrary, it can also be defined as

$$C(t) = \langle A(\tau)A(t + \tau) \rangle \quad (109)$$

- * Velocity autocorrelation function: correlates the velocity of an atom with itself. By correlating the velocity of an atom at a given time with its velocity at a later time, it reveals what effect the interatomic forces have had on this atom's motion, i.e., if there were no forces, the atom's velocity would never change, and the correlation would stay at a fixed value for all time
- The Wiener-Khintchine Theorem connects the autocorrelation in time domain with the Fourier transform:

$$\langle A(\tau)A(t + \tau) \rangle_{\tau} = \frac{1}{2\pi} \int | \int A(t) \exp\{-i\omega t\} dt |^2 \exp\{i\omega t\} d\omega \quad (110)$$

- Amplitude of a mode is inversely proportional to the eigenvalue frequency ω and the square root of the mass, so to obtain the same intensity for each mode one should rescale the spectrum by ω^2 , which is formally equivalent to correlating time derivatives:

$$P(\omega) = m\omega^2 \int \langle \mathbf{r}(\tau)\mathbf{r}(t + \tau) \rangle_{\tau} \exp\{-i\omega t\} dt \quad (111)$$

$$= m \int \langle \dot{\mathbf{r}}(\tau)\dot{\mathbf{r}}(t + \tau) \rangle_{\tau} \exp\{-i\omega t\} dt \quad (112)$$

- * In the case of infrared, we autocorrelate the derivative of the dipole moment, resulting in

$$A(\omega) \propto \int \langle \dot{\boldsymbol{\mu}}(\tau)\dot{\boldsymbol{\mu}}(t + \tau) \rangle_{\tau} \exp\{-i\omega t\} dt \quad (113)$$

Lecture 11: Linear Scaling DFT and High Performance Computing

- The goal of DFT is to minimize the energy functional E_{KS} with respect to the electron density $n(r)$, i.e.,

$$\min_{n(r)} E_{KS}[n(r)] \quad , \quad (114)$$

where $n(r) = \sum_{i=1}^N |\Psi_i(r)|^2$ and the single atom orbitals Ψ_i are expanded with basis functions (can be non-orthogonal), i.e., $\Psi_i(r) = \sum_{\alpha} C_i^{\alpha} \phi_{\alpha}(r)$

- Results in a generalized eigenvalue equation of the form

$$\underline{\mathbf{H}}_{KS} \mathbf{C}_i = \underline{\mathbf{S}} \mathbf{C}_i \epsilon_i \quad , \quad (115)$$

where $\underline{\mathbf{H}}_{KS}$ is the Kohn-Sham matrix, $\underline{\mathbf{S}}$ the overlap matrix, \mathbf{C}_i the orbital coefficient vector i and ϵ_i the orbital eigenvalue i . The Kohn-Sham energy is then computed via density matrix $P = \sum \mathbf{C}_i \mathbf{C}_i^T$, i.e.,

$$E_{KS}[n(r)] = E_{KS}[P] \quad (116)$$

- Orthonormality is fulfilled, i.e.,

$$\mathbf{C}_i^T \underline{\mathbf{S}} \mathbf{C}_i = \delta_{ij} \quad (117)$$

- The matrix $\underline{\mathbf{P}}$ is calculated by diagonalizing the Kohn-Sham matrix $\underline{\mathbf{H}}_{KS}$
 - Only 10% – 50% of the eigenvectors are needed
 - The goal of the Lanczos algorithm is to find the k largest/smallest eigenvalues and eigenvectors
 - * Run-time is $O(N^2)$ for a dense and $O(N)$ for a sparse matrix
- Self-consistent field (SCF) equations are traditionally solved in $O(N^3)$
 - Linear scaling SCF: avoid finding 20% lowest eigenvectors, results in $O(N)$

The Density Matrix $\underline{\mathbf{P}}$ as a function of the Hamiltonian $\underline{\mathbf{H}}_{KS}$

- The density matrix $\underline{\mathbf{P}}$ can be expressed as a function of the Hamiltonian $\underline{\mathbf{H}}_{KS}$, resulting in a Fermi function:

$$\underline{\mathbf{P}} \underline{\mathbf{S}} = \frac{1}{1 + \exp\left\{\frac{\underline{\mathbf{S}}^{-1} \underline{\mathbf{H}}_{KS} - \mu \underline{\mathbf{I}}}{k_B T}\right\}} \quad (118)$$

- In the limit of small $k_B T$, Eq. (118) can be conveniently written as

$$\underline{\mathbf{P}} \underline{\mathbf{S}} = \frac{1}{2} (1 - \text{sign}(\underline{\mathbf{S}}^{-1} \underline{\mathbf{H}}_{KS} - \mu \underline{\mathbf{I}})) \quad (119)$$

- The matrix sign function:

$$\text{sign}(\underline{\mathbf{A}}) = \underline{\mathbf{A}}(\underline{\mathbf{A}}^2)^{-1/2} \quad (120)$$

- For diagonalizable $\underline{\mathbf{A}}$, eigenvectors of $\underline{\mathbf{A}}$ are eigenvectors of $\text{sign}(\underline{\mathbf{A}})$ with eigenvalues of -1 and 1 respectively

- To compute $\text{sign}(\underline{\mathbf{A}})$, one can use the Newton-Schultz iteration, which only requires matrix multiplications:

$$\underline{\mathbf{X}}_{n+1} = \frac{1}{2}\underline{\mathbf{X}}_n(3\underline{\mathbf{I}} - \underline{\mathbf{X}}_n^2) \quad , \quad (121)$$

with $X_\infty = \text{sign}(A)$

Lecture 12: Selected Topics

Periodic Systems

- In a metal there is a partially filled band (valence band) and there are always empty states available for conductivity (conduction band); both the conduction and valence band are overlapping
 - Semiconductors only conduct if the electrons are excited from the valence band to the conduction band through the forbidden band gap (E_g); the conduction and valence band are separated by the energy E_g

Electrons in a Periodic Potential

- Crystal is described in real space in terms of lattice vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and the positions of atoms inside the primitive unit cell:

$$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \quad (122)$$

- The lattice vectors connect all equivalent points in space (Bravais lattice)

- The reciprocal lattice is the inverse space of the real lattice. We can construct vectors connecting all equivalent points in reciprocal space and call it \mathbf{G} (in analogy to the Bravais lattice vectors):

$$\mathbf{G} = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3 \quad , \quad (123)$$

with \mathbf{b}_i being the reciprocal lattice vectors

-

$$\mathbf{R} \cdot \mathbf{G} = 2\pi l, \quad l = n_1m_1 + n_2m_2 + n_3m_3 \quad (124)$$

- Any function can be written with the periodicity of the Bravais lattice as

$$f(\mathbf{r}) = \sum_{\mathbf{G}} \exp\{i\mathbf{G} \cdot \mathbf{r}\} f(\mathbf{G}) \quad (125)$$

Bloch's Theorem

- A Bloch wave is a type of wave function for a particle in a periodically-repeating environment, most commonly an electron in the crystal
- When the potential in the single-particle Hamiltonian has the translational periodicity of the Bravais lattice, i.e.,

$$V^{sp}(\mathbf{r} + \mathbf{R}) = V^{sp}(\mathbf{r}) \quad , \quad (126)$$

then the single-particle wave functions have the same symmetry (up to a phase factor):

$$\Psi_{\mathbf{k}}(\mathbf{r} + \mathbf{R}) = \exp\{i\mathbf{k} \cdot \mathbf{R}\}\Psi_{\mathbf{k}}(\mathbf{r}) \quad (127)$$

The Tight-Binding Method

- The tight-binding model (TBM) is an approach to the calculation of the electronic band structure using an approximate set of wave functions based upon superposition of wave functions for isolated atoms located at each atomic site
 - The electron band structure of a solid describes those ranges of energy that an electron within a solid may (energy bands) or may not (band gaps) have
 - In a metal, the atoms share some of their electrons and their energies fill the energy band from the bottom. As soon as the energy band is filled, the electrons become very difficult to move and the material no longer conducts electric current, thus becoming an insulator

1. In a crystal, the single atom Hamiltonian is given by

$$H = H_{at} + \Delta U \quad , \quad (128)$$

where H_{at} is the single particle Hamiltonian and ΔU encodes all the differences between the true potential in the crystal and the potential of an isolated atom

- We assume $\Delta U \rightarrow 0$ at the centre of each atom in a crystal

2. The single atom states in a crystal are then $\psi_{n\mathbf{k}}(\mathbf{r})$, where

$$H\psi_{n\mathbf{k}}(\mathbf{r}) = E_{n\mathbf{k}}\psi_{n\mathbf{k}}(\mathbf{r}) \quad , \quad (129)$$

where n is the band index, and \mathbf{k} is a wavevector in the first Brillouin zone

- It is required that the single atom states $\psi_{n\mathbf{k}}$ obey Bloch's theorem

3. The atomic wave functions $\phi_l(\mathbf{r} - \mathbf{t}_i)$ are eigenstates of the single particle Hamiltonian H_{at} :

$$H_{at}\phi_l(\mathbf{r} - \mathbf{t}_i) = \epsilon_i\phi_l(\mathbf{r} - \mathbf{t}_i) \quad , \quad (130)$$

where ϵ_i is the energy of the i th energy level in an isolated atom and l is the index of the atomic orbital (1s, 2s, etc.)

- $\phi_l(\mathbf{r} - \mathbf{t}_i)$ is also called a *atomic orbital* and is the atomic wave function at the lattice point \mathbf{t}_i
 - Atomic orbitals describe the wave-like behaviour of an electron (or a pair of electrons) in an atom
 - For CaCO_3 we would need 15 atomic wave functions ϕ_l to represent the 15 valence orbitals consisting of the $4s$ orbital for calcium, the $2s$ and $3 \times 2p$ -orbitals for carbon and the $2s$ and $3 \times 2p$ -orbitals for oxygen
4. It is required that the single particle states must obey Bloch's theorem; a single atomic wave function ϕ_l does not satisfy Bloch's theorem, so we make a linear combination of atomic orbitals (LCAO) that does,

$$\chi_{\mathbf{k}li}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} \exp\{i\mathbf{k} \cdot \mathbf{R}\} \phi_l(\mathbf{r} - \mathbf{t}_i - \mathbf{R}) \quad , \quad (131)$$

where \mathbf{R} is the real space translation vector for a unit cell defined in Eq. (122) and thus resulting in a summation over all atomic sites (all lattices and bases)

- $\chi_{\mathbf{k}li}$ in Eq. (131) is called a *Bloch wave* or *Bloch state* and represents a molecular orbital consisting of a linear combination of atomic orbitals
 - A molecular orbital describes the wave-like behaviour of an electron in a molecule and is usually constructed by combining atomic (or hybrid) orbitals from each atom of the molecule
 - The σ -bond and π -bond are both molecular orbitals
 - The principles of molecular orbital theory state that the total number of molecular orbitals is equal to the total number of atomic orbitals used to make them
5. In general, a material will have more than one atom in the unit cell. The Bloch wave in Eq. (131) accounts only for the state of a single atom in the unit cell. For multiple atoms i in the unit cell with positions \mathbf{t}_i (see Eq. (131)), a new Bloch state can be constructed from a linear combination of the Bloch states defined in Eq. (131), resulting in the *tight binding wave function* $\psi_{n\mathbf{k}}$ defined as

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{l,i} c_{n\mathbf{k}li} \chi_{\mathbf{k}li}(\mathbf{r}) \quad . \quad (132)$$

6. The energy of the tight binding wave function $\psi_{n\mathbf{k}}$ in Eq. (132) can be evaluated by substituting in into the time independent Schrödinger equation and using the single atom Hamiltonian shown in Eq. (128),

$$H\psi_{n\mathbf{k}} = \epsilon_{n\mathbf{k}}\psi_{n\mathbf{k}} \quad . \quad (133)$$

Eq. (133) can be rewritten as a component equation,

$$\sum_{l,i} [\langle \chi_{\mathbf{k}mj} | H | \chi_{\mathbf{k}li} \rangle - \epsilon_{n\mathbf{k}} \langle \chi_{\mathbf{k}mj} | \chi_{\mathbf{k}li} \rangle] c_{n\mathbf{k}li} = 0 \quad , \quad (134)$$

with

$$\langle \chi_{\mathbf{k}m_j} | \chi_{\mathbf{k}l_i} \rangle = \sum_{\mathbf{R}} \exp\{i\mathbf{k} \cdot \mathbf{R}\} \langle \phi_m(\mathbf{r} - \mathbf{t}_j) | \phi_l(\mathbf{r} - \mathbf{t}_i - \mathbf{R}) \rangle \quad (135)$$

$$\langle \chi_{\mathbf{k}m_j} | H | \chi_{\mathbf{k}l_i} \rangle = \sum_{\mathbf{R}} \exp\{i\mathbf{k} \cdot \mathbf{R}\} \langle \phi_m(\mathbf{r} - \mathbf{t}_j) | H | \phi_l(\mathbf{r} - \mathbf{t}_i - \mathbf{R}) \rangle \quad (136)$$

$$\langle \phi_m(\mathbf{r} - \mathbf{t}_j) | \phi_l(\mathbf{r} - \mathbf{t}_i - \mathbf{R}) \rangle = \delta_{lm} \delta_{ij} \delta(\mathbf{R}) \quad (137)$$

$$\langle \phi_m(\mathbf{r} - \mathbf{t}_j) | H | \phi_l(\mathbf{r} - \mathbf{t}_i - \mathbf{R}) \rangle = \delta_{lm} \delta_{ij} \delta(\mathbf{R}) \epsilon_l \quad (138)$$

Example: One-dimensional crystal with one atom in the basis (linear chain with s -orbitals)

- Consider a one-dimensional crystal of atoms with one valence orbital ϕ_l (s -orbital). The tight binding wave function for this case is (only one valence orbital),

$$\chi_{kl}(x) = \sum_{n=-\infty}^{\infty} \exp\{ikna\} \phi_l(x - na) \quad , \quad (139)$$

where a is the width of the unit cell

- The overlap integrals (Eq. (137)) evaluate to

$$\langle \phi_l(x) | \phi_l(x - na) \rangle = \delta_{n0} \quad (140)$$

$$\langle \phi_l(x) | H | \phi_l(x - na) \rangle = \epsilon_l \delta_{n0} + t_l \delta_{n\pm 1} \quad , \quad (141)$$

where H is the single atom Hamiltonian from Eq. (128) and Eq. (141) depicts the interaction between nearest neighbours

- The tight binding wave function for this problem (single atom in a unit cell) is simply given by

$$\psi_k = c_k \chi_{kl}(x) \quad (142)$$

- Inserting the wave function from Eq. (142) into the time dependent Schrödinger equation from Eq. (133) and rewriting it using Eq (134) we get

$$\begin{aligned} & [\langle \chi_{kl}(x) | H | \chi_{kl}(x) \rangle - \epsilon_k \langle \chi_{kl}(x) | \chi_{kl}(x) \rangle] c_k = 0 \\ \Leftrightarrow & \sum_n \exp\{ikna\} [\epsilon_l \delta_{n0} + t_l \delta_{n\pm 1}] = \epsilon_k \sum_n \exp\{ikna\} \delta_{n0} \end{aligned} \quad (143)$$

- Since in the tight binding approximation only the on-site ($n = 0$) and nearest neighbour ($n \pm 1$) matrix elements are retained for the calculation, for the given problem Eq. (143) simplifies to

$$\begin{aligned} & \exp\{-ka\} \underbrace{[\epsilon_l \delta_{-10}]}_{=0} + \underbrace{t_l \delta_{-1\pm 1}}_{=t_l} + \exp\{0\} \underbrace{[\epsilon_l \delta_{00}]}_{=\epsilon_l} + \underbrace{t_l \delta_{0\pm 1}}_{=0} \\ & + \exp\{ka\} \underbrace{[\epsilon_l \delta_{10}]}_{=0} + \underbrace{t_l \delta_{1\pm 1}}_{=t_l} = \epsilon_k \exp\{0\} \delta_{00} \\ \Leftrightarrow & \underline{\epsilon_l + t_l (\exp\{-ka\} + \exp\{ka\}) = \epsilon_k} \end{aligned} \quad (144)$$

- By using the relation $\cos(\alpha) = 1/2(\exp\{i\alpha\} + \exp\{-i\alpha\})$ in Eq. (144), the energy ϵ_k for a one-dimensional crystal with one atom in the basis is given by

$$\underline{\underline{\epsilon_k = \epsilon_l + 2t_l \cos(ka)}} \quad (145)$$

References

- [1] J. VANDEVONDELE, 327-5102-001 molecular and materials modelling. <http://vvz.ethz.ch/Vorlesungsverzeichnis/lerneinheitPre.do?lerneinheitId=97794&semkez=2015S&lang=de>.