Part I. Introduction to Molecular Modeling

1. What is molecular modeling
Model is some kind of a representation of a system, usually simplified, that allows for description, investigation and prediction of properties of interest.

Molecular modeling, which is more or less synonymous with computational chemistry, is a discipline concerned with developing models of molecular systems, chemical reactions and phenomena of interest to chemistry.

Why chemists use models:
a) to help with analysis and interpretation of experimental data
b) to uncover new laws and formulate new theories
c) to help solve problems and hint solutions before doing experiments
d) to help design new experiments
e) to predict properties and quantities that are difficult or even impossible to observe experimentally

Simulations and computer “experiments” can be designed to mimic reality, however, are always based on assumptions, approximations and simplifications (i.e. models).

Important characteristics of models are:
a) Level of simplification: very simple to very complex
b) Generality: general or specific, i.e. relate only to specific systems or problems
c) Limitations: one must always be aware of the range of applicability and limits of accuracy of any model.
d) Cost and efficiency: CPU time, memory, disk space

Computable quantities:
a) molecular structures: closely tied to energy (best structure - one for which the energy is minimum)
b) energy: potential energy surfaces (PES) - extremely important! PES dictate essentially everything about the molecule or system
c) molecular properties that can be compared to/used to interpret experiments: thermodynamics, kinetics, spectra (IR, UV, NMR)
d) properties that are not experimental observables: bond order, aromaticity, molecular orbitals ….

2. Model chemistries
Molecular modeling (a.k.a. computational chemistry) is based on numerically (on the computer) solving the equations that express fundamental physical laws. This can be done at various levels: we talk about “levels of theory” or “model chemistries”

2.1. Quantum mechanics
Quantum mechanics, also electronic structure methods, are the most fundamental level of theory at which molecules can be treated. In seeks the solution to the Schrodinger equation:

\[
\hat{H} \psi = E \psi
\]
The price one pays is that the equation cannot be solved exactly and, as we will see, drastic approximations are often required. Even such approximate solutions can be difficult and computationally expensive and are generally limited to small molecules or systems.

The specific approximations used in solutions define the more specific “level of theory” or “model chemistry” used.

The basic classification of quantum mechanical methods is:
   a) Semi-empirical methods
   b) Ab initio methods
   c) Density functional theory (DFT) methods
   d) Compound methods

The specific level of theory is defined by the combination of the theoretical method (which can be roughly thought of as the approximation to the Hamiltonian $\hat{H}$ in eqn. 1), and the basis set, (which is an approximation for the wavefunction $\psi$). For example, HF/6-31G(d) level of theory is Hartree-Fock method with 6-31G(d) basis set. Semiempirical methods use default basis set and the Hamiltonian alone defined the level, for example CNDO. Density functional methods are specified by different density functionals (e.g. BLYP with the same basis set would give BLYP/6-31G(d) level).

Another important component to specify, though not strictly a part of the model chemistry, is whether the calculation is spin restricted or unrestricted. This defines how electrons with different spins are handled: for closed shell systems, where all orbitals are doubly occupied by electrons with opposite spins the restricted calculations are done. For open shell systems, such as radicals, excited states etc., unrestricted spin model is usually used and specified with an “U”: e.g. UHF/6-31G(d).

2.2. Molecular mechanics
Molecular mechanics methods are based on the laws of classical physics. They do not explicitly treat the electrons: the electronic effects, such as chemical bonds, are included implicitly in the energy function, known as the force field, which describes the interactions between the nuclei only. As such, the molecular mechanics methods can be viewed as a “coarse-grained” model: the electrons are lumped into empirical parameters.

Molecular mechanics methods are characterized by the force fields, many of which have been parameterized. The advantage of molecular mechanics methods is much higher computational efficiency, which allows modeling of very large systems with many atoms, simulations of molecular ensembles etc. The main disadvantages are that a) the force fields are typically limited to specific classes of molecules for which they were parameterized, and b) since electrons are not treated, properties that depend on details of the electronic structure cannot be modeled. This primarily involves chemical reactions in which bonds are formed and/or broken.

2.3. Hybrid QM/MM methods
Hybrid methods try to overcome the limitations of quantum and classical methods by combining them. In such model chemistry, a part (small) of the system that may undergo reactions is treated at quantum mechanical level, which the rest (large) is treated classically using a force field. Different schemes can be used to model the interactions between the quantum and classical parts.
3. Units
Proper units are an essential part of specifying and reporting of any physical quantity. To get any meaningful results from computational modeling (as from experiments) one needs to know how to use and convert units. Computational chemistry, in particular quantum, often uses so called atomic units, but not exclusively. Angstroms ($10^{-10}$ m), calorie based energy units (in particular cal/mol) are common as are wavenumbers (cm$^{-1}$), cgs units (Debye) and others. Some commonly encountered units and conversions are given in the tables below.

### Fundamental atomic units

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Name</th>
<th>Symbol</th>
<th>Value in SI units</th>
</tr>
</thead>
<tbody>
<tr>
<td>mass</td>
<td>electron rest mass</td>
<td>$m_e$</td>
<td>$9.1093826(16)\times10^{-31}$ kg</td>
</tr>
<tr>
<td>charge</td>
<td>elementary charge</td>
<td>$e$</td>
<td>$1.60217653(14)\times10^{-19}$ C</td>
</tr>
<tr>
<td>angular momentum</td>
<td>Reduced Planck's constant</td>
<td>$\hbar = \frac{h}{2\pi}$</td>
<td>$1.05457168(18)\times10^{-34}$ J·s</td>
</tr>
<tr>
<td>electric constant</td>
<td>Coulomb force constant</td>
<td>$\frac{1}{(4\pi\varepsilon_0)}$</td>
<td>$8.9875517873681\times10^9$ kg·m³·s⁻²·C⁻²</td>
</tr>
</tbody>
</table>

### Derived atomic units

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Name</th>
<th>Symbol</th>
<th>Expression</th>
<th>Value in SI units</th>
<th>Value in more common units</th>
</tr>
</thead>
<tbody>
<tr>
<td>length</td>
<td>Bohr radius</td>
<td>$a_0$</td>
<td>$\frac{4\pi\varepsilon_0 h^2}{m_e c^2}$</td>
<td>$5.291772108(18)\times10^{-11}$ m</td>
<td>0.052918 nm = 0.52918 Å</td>
</tr>
<tr>
<td>energy</td>
<td>Hartree energy</td>
<td>$E_h$</td>
<td>$\frac{m_e^4}{(4\pi\varepsilon_0 \hbar)^2}$</td>
<td>$4.35974417(75)\times10^{-18}$ J</td>
<td>27.211 eV</td>
</tr>
<tr>
<td>time</td>
<td></td>
<td></td>
<td>$\frac{\hbar}{E_h}$</td>
<td>$2.418884326505(16)\times10^{-17}$ s</td>
<td></td>
</tr>
<tr>
<td>velocity</td>
<td></td>
<td>$a_0 E_h / \hbar = \frac{\alpha c}{2}$</td>
<td>$2.1876912633(73)\times10^6$ m·s⁻¹</td>
<td></td>
<td></td>
</tr>
<tr>
<td>force</td>
<td></td>
<td>$E_d/a_0$</td>
<td></td>
<td>$8.2387225(14)\times10^{-8}$ N</td>
<td>82.387 nN = 51.421 eV·Å⁻¹</td>
</tr>
<tr>
<td>temperature</td>
<td></td>
<td>$E_d/k_B$</td>
<td></td>
<td>$3.1577464(55)\times10^3$ K</td>
<td>27.211 eV</td>
</tr>
<tr>
<td>pressure</td>
<td></td>
<td>$E_d/a_0^3$</td>
<td></td>
<td>$2.9421912(19)\times10^{13}$ Pa</td>
<td></td>
</tr>
<tr>
<td>electric field</td>
<td></td>
<td>$E_d/(e a_0)$</td>
<td></td>
<td>$5.1421\times10^{11}$ V·m⁻¹</td>
<td>514.21 V·nm⁻¹ = 51.421 V·Å⁻¹</td>
</tr>
</tbody>
</table>

### Other units often used in (computational) chemistry

<table>
<thead>
<tr>
<th>Name</th>
<th>Quantity</th>
<th>Symbol</th>
<th>Value in SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>calorie</td>
<td>energy</td>
<td>cal</td>
<td>$4.184$ J</td>
</tr>
<tr>
<td>wavenumber (inverse centimeter)</td>
<td>frequency (inverse of wavelength)</td>
<td>cm⁻¹</td>
<td>$10^2$ m⁻¹</td>
</tr>
<tr>
<td>atomic mass unit</td>
<td>mass</td>
<td>amu</td>
<td>$1.660 538 782(83)\times10^{-27}$ kg</td>
</tr>
<tr>
<td>Ångström</td>
<td>length</td>
<td>Å</td>
<td>$10^{-10}$ m</td>
</tr>
</tbody>
</table>
Part II. Modeling of Essential Molecular Properties

1. Single Point Energy Calculations
Single point energy calculation amounts to solving the Schrödinger equation (eqn. 1) at a given, fixed molecular geometry – which is why it is called “single point”. It refers to a “single point” on the molecule’s Potential Energy Surface (PES – see the next section).

Single point energy calculations are used to obtain not only the total energy, which on its own is not all that useful (we will discuss its meaning more later), but more importantly various properties related to the wavefunction, such as:

- **Molecular orbitals (and orbital energies).** As you well know, electron orbitals are very important for understanding chemistry and reactivity, especially the so-called Frontier orbitals, such as HOMO (Highest Occupied Molecular Orbital) and LUMO (Lowest Unoccupied Molecular Orbital). Single point energy calculations are often used to compute and visualize molecular orbitals.

- **Electron density, atomic partial charges, dipole moment and higher multipole moments.** These properties are related to the square absolute value of the wavefunction, and their calculation is called Population analysis. Some of these properties, namely the atomic partial charges, are somewhat arbitrary, which relates to the fact that the wavefunction cannot be unambiguously partitioned between individual atoms: we will again revisit this issue later.

- **Wavefunction stability analysis.** This is a more technical problem, which has to do with ensuring that the correct solution of Schrödinger equation has been found. Once again, this will be expanded upon later in the course, after we review all the necessary theoretical background.

While single point energy calculations can be done for any molecular geometry, generally we are most interested in geometries that correspond to molecules as they exist in reality. These are found through geometry optimization calculations, which are described next. Because optimizations are lengthy, multi-step computations, often single point energy calculations at more accurate (and computationally intensive) level are done for structures that have been optimized at lower (and faster) level, to get better energies, orbitals, dipole moments etc.

2. Geometry Optimizations

2.1. The Potential Energy Surface
The potential energy surface (PES) is a central concept in computational chemistry. It specifies how the energy of a molecular system varies with small changes in its structure. A potential energy surface is a mathematical relationship linking molecular structure and the resultant energy.

For a diatomic molecule, it is a two-dimensional plot with the inter-nuclear separation on the X-axis (the only way that the structure of such a molecule can vary), and the potential energy at that bond distance on the Y-axis, producing a curve.
For larger systems, the surface has as many dimensions as there are degrees of freedom within the molecule.

Generally, a non-linear $N$ atomic molecule, has $3N - 6$ degrees of freedom, or internal coordinates. This is because all $N$ atoms can move in 3 dimensions ($x$, $y$ and $z$) giving $3N$ degrees of freedom. However, 6 of those: three translations - in $x$, $y$, $z$ directions, and three rotations - along $x$, $y$ and $z$ axes of the molecule as a whole - do not produce any change in energy. There are also referred to as external coordinates.

For a linear molecule there are only two rotations and the number of internal coordinates (degrees of freedom) is $3N - 5$.

A potential energy surface (PES) is often represented by illustrations like the one below. This sort of drawing considers only two of the degrees of freedom within the molecule, and plots the energy above the plane defined by them, creating a literal surface. Each point corresponds to the specific values of the two structural variables-and thus represents a particular molecular structure-with the height of the surface at that point corresponding to the energy of that structure.

### 2.2 Stationary points of PES

There are two minima on this potential surface. A minimum is the bottom of a valley on the potential surface. From such a point, motion in any direction-a physical metaphor corresponding to changing the structure slightly-leads to a higher energy.

A minimum can be a local minimum, meaning that it is the lowest point in some limited region of the potential surface, or it can be the global minimum, the lowest energy point anywhere on the potential surface. Minima occur at equilibrium structures for the system, with different minima corresponding to different conformations or structural isomers in the case of single molecules, or reactant and product molecules in the case of multi-component systems.

Peaks and ridges correspond to maxima on the potential energy surface. A peak is a maximum in all directions (i.e., both along and across the ridge). A low point along a ridge - a mountain pass in our topographical metaphor - is a local minimum in one direction (along the ridge), and a
maximum in the other. A point which is a maximum in one direction and a minimum in the other (or in all others in the case of a larger dimensional potential surface) is called a saddle point (based on its shape).

For example, the saddle point in the diagram is a minimum along its ridge and a maximum along the path connecting minima on either side of the ridge. A saddle point corresponds to a transition structure connecting the two equilibrium structures.

The slice along the “lowest-energy” path connecting the global minimum, the local minimum and the transition state, that is, along the “reaction coordinate” gives a one-dimensional energy surface. Here the transition state (saddle point) looks like a maximum. Note that the horizontal axis (i.e. the reaction coordinate) is a combination of the two coordinates (bond length and angle) from the original plot.

Together, minima, maxima and saddle points are called stationary points. They have in common that at any stationary point (minimum, maximum) the surface is flat, i.e. parallel to the horizontal
line corresponding to the one geometric parameter (or to the plane corresponding to two geometric parameters etc). A marble placed on a stationary point will remain balanced, i.e. stationary, while at any other point on a potential surface the marble will roll toward a region of lower potential energy.

Mathematically, the stationary points are characterized by zero first derivatives with respect to the geometrical parameters:

$$\frac{\partial V}{\partial q_1} = \frac{\partial V}{\partial q_2} = \cdots = \frac{\partial V}{\partial q_n} = 0$$  \hspace{1cm} (2.1)

where the potential energy is denoted by $V$ and:

- $n = 3N - 6$ for internal coordinates
- $n = 3N$ for Cartesian coordinates
- $n > 3N - 6$ for redundant internal coordinates, i.e. internal coordinates, such as bond lengths, angles and torsions, but more of them than we actually need.

The derivatives can be arranged into a vector, which is called **gradient of energy**

$$\text{grad } V = \nabla V = \begin{pmatrix} \frac{\partial V}{\partial q_1} \\ \frac{\partial V}{\partial q_2} \\ \cdots \\ \frac{\partial V}{\partial q_n} \end{pmatrix}$$  \hspace{1cm} (2.2)

as we know from basic physics, gradient of energy is closely related to force:

$$\mathbf{F} = -\text{grad } V = -\nabla V$$  \hspace{1cm} (2.3)

Therefore at a stationary point, the forces on the atoms in the molecule in all directions are all zero, which is another reason why they are stationary points.

**2.3. Locating minima**

Minima on potential energy structures are important because they correspond to the optimum molecular geometries, also called equilibrium structures. Geometry optimization is therefore more or less synonymous with **energy minimization**. From a mathematical point of view, it is a problem of finding a minimum of a function of many variables.

The basic idea of a minimization algorithm is to vary the geometry of the molecule (i.e. atomic positions) in some systematic fashion so that the energy keeps decreasing; once the energy stops decreasing and only increases with further geometry variations, the minimum is found.
Important point to remember is that all minimization algorithms only find local minima: you are never guaranteed that the minimum found is a global one!

There are several classes of algorithms that have been developed for this purpose. They can be divided into three basic classes:

- only function values are used (no derivatives)
- first derivatives (gradients) are used
- second derivatives (Hessians) are used.

The first class represents the most robust, but least efficient way to find the minimum. The example is the Simplex algorithm (also known as “amoeba”). It is robust, because it can always find a minimum of pretty much any function, no matter how crazy. Least efficient means that for well behaved functions, the other two classes could do it ways faster! Because molecular PES are usually smooth and well behaved functions of atomic coordinates, these algorithms are usually not the first choice for optimizations.

If the gradients can be computed (note that the function has to be differentiable - i.e. continuous), they indicate in which direction the energy decreases. Remember the gradient is the negative of force, and the force pushes the atoms to their minimum energy position (= equilibrium) position. The simples gradient based algorithm is the so-called “steepest descent”, where steps are made in the direction of the gradient until the minimum is found (in that direction), then the gradient is recomputed, steps are made in the new direction etc. This is simple but in practice very inefficient.

With gradients available the direction toward the potential minimum is known, but not how far one has to step to get there. Imagine the surface could be approximated by a quadratic function (parabola). Assuming one dimension for illustration we could write a Taylor expansion around some point $q_0$:

$$V(q) \approx V(q_0) + \left( \frac{dV}{dq} \right)_{q_0} (q - q_0) + \frac{1}{2} \left( \frac{d^2V}{dq^2} \right)_{q_0} (q - q_0)^2 \quad (2.4)$$

taking a derivative with respect to $q$:

$$\left( \frac{dV}{dq} \right)_q \approx \left( \frac{dV}{dq} \right)_{q_0} + \left( \frac{d^2V}{dq^2} \right)_{q_0} (q - q_0) \quad (2.5)$$

Assuming the point $q_0$ is our starting point and we are looking for the next point “q” on our PES where the energy is minimum, we know that at “q” (the minimum) the gradient has to be zero. Setting the gradient at point “q”, which is the left-hand side of the last equation, equal to zero:

$$0 = \left( \frac{dV}{dq} \right)_{q_0} + \left( \frac{d^2V}{dq^2} \right)_{q_0} (q - q_0) \quad (2.6)$$
and solving for $q$:

$$q = q_0 - \left( \frac{dV}{dq} \right)_{q_0}^{-1} \left( \frac{d^2V}{dq^2} \right)_{q_0}^{-1} \left( \frac{dV}{dq} \right)_{q_0}^{-1}$$  \hspace{1cm} (2.7)$$

In more dimensions (remember we have $n$ coordinates, which can be $3N - 6$ or more for $N$ atoms depending on the coordinates), $q$ has $n$ components, gradient is a vector with $n$ components (eqn. 2.2), which we denote $\mathbf{g}$, and the second derivatives are so called *force constants*, which form a $n$-by-$n$ matrix called *Hessian*, which we will denote $H$:

$$H_{ij} = \frac{\partial^2 V}{\partial q_i \partial q_j} \hspace{1cm} H = \begin{pmatrix}
\frac{\partial^2 V}{\partial q_1 \partial q_1} & \frac{\partial^2 V}{\partial q_1 \partial q_2} & \cdots & \frac{\partial^2 V}{\partial q_1 \partial q_n} \\
\frac{\partial^2 V}{\partial q_2 \partial q_1} & \frac{\partial^2 V}{\partial q_2 \partial q_2} & \cdots & \frac{\partial^2 V}{\partial q_2 \partial q_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial^2 V}{\partial q_n \partial q_1} & \frac{\partial^2 V}{\partial q_n \partial q_2} & \cdots & \frac{\partial^2 V}{\partial q_n \partial q_n}
\end{pmatrix} \hspace{1cm} (2.8)$$

The equation then becomes:

$$q = q_0 - \left( H^{(0)} \right)^{-1} \mathbf{g}^{(0)} \hspace{1cm} (2.9)$$

where the superscript “0” symbolizes the starting point and the “-1” is matrix inverse.

For the quadratic (parabolic) surface, this formula will take us from any starting point straight to the minimum. If the surface is not exactly parabolic, this formula can be iterated, i.e. compute gradient and (inverted) Hessian at point $q_0$, move to the new point $q$ using formula (2.9), recalculate gradient and Hessian etc.

This algorithm is known as *Newton-Raphson method*.

The closer to the minimum one gets the better the quadratic approximation becomes: eventually this algorithm will converge to the minimum.

The disadvantages of Newton-Raphson method are the following:

- Hessian matrix is large and computationally expensive to construct at every iteration step
- Hessian must be inverted (inverting a matrix is also computationally expensive)
- Second derivatives are even more sensitive to discontinuities than gradients: for surfaces that are not “well behaved” the method may become unstable and diverge or crash.
Do avoid that, a class of methods known as *Quazi-Newton* (also variable metric methods) instead of using the true Hessian estimate it (actually its inverse) approximately, e.g. from the gradients at the successive steps (there are many schemes for estimating and updating Hessians, see e.g. *Numerical Recipes: The Art of Scientific Computing* by William H. Press, Brian P. Flannery, Saul A. Teukolsky and William T. Vetterling, www.nr.com). The exact Hessian can be calculated once in a while, if needed, typically when the optimization gets close to the minimum.

While this may sound like a step back - using an approximation instead of the true $H$, *quazi-Newton* methods are actually more reliable and efficient than the old *Newton-Raphson*. You can imagine more sophisticated algorithms for finding minima, using e.g. higher order (cubic) approximations to the surface etc. Some details of the algorithm used by Gaussian can be found on the Gaussian manual pages.

### 2.4. Analytical vs. numerical gradients and Hessians

Gaussian uses both gradients and approximate Hessians or true Hessians when available to find the minima. It is important to know how the computations of gradients and Hessians are implemented for each method.

For example, under MP & Double Hybrid DFT methods in Gaussian manual you will find this:

**AVAILABILITY**

**MP2, B2PLYP[D], mPW2PLYP[D]**: Energies, analytic gradients, and analytic frequencies.

**MP3, MP4(DQ) and MP4(SDQ)**: Energies, analytic gradients, and numerical frequencies.

**MP4(SDTQ) and MP5**: Analytic energies, numerical gradients, and numerical frequencies.

Here “frequencies” as we will see later, refer essentially to Hessians, because vibrational frequencies are calculated directly from them. The important distinction here is *analytical vs. numerical*. While there is no such thing as rigorously “analytical” in Gaussian calculations (everything is done numerically),

- **analytical gradients** (frequencies) mean that essentially *exact* derivatives are calculated by solving the corresponding derivative Shrodinger equation.

- **numerical gradients** are calculated by finite differencing from energies at two points:

$$
\left( \frac{dV}{dq} \right)_{q_0} \approx \frac{V(q_0 + \Delta q) - V(q_0)}{\Delta q}
$$

(2.10)

this is not only less accurate, but also more computationally costly (especially the larger the molecules get) because for each coordinates two energies have to be evaluated at two different points (i.e. the Shrodinger equation solved twice).

- **numerical frequencies** (or more precisely the force constants) are calculated either by numerical differentiation of the analytical gradients, or in the worst case scenario, by numerical differentiation of numerical gradients (which is equivalent to calculating finite difference second derivatives from the energies).
\[ \left( \frac{d^2V}{dq^2} \right)_{q_0} \approx \frac{V(q_0 + \Delta q) + V(q_0 - \Delta q) - 2V(q_0)}{\Delta q^2} \]  

(2.10)

Obviously, this is very slow and very inaccurate.

Fortunately, most methods, even the very high level ones nowadays have implemented at least analytical gradients.