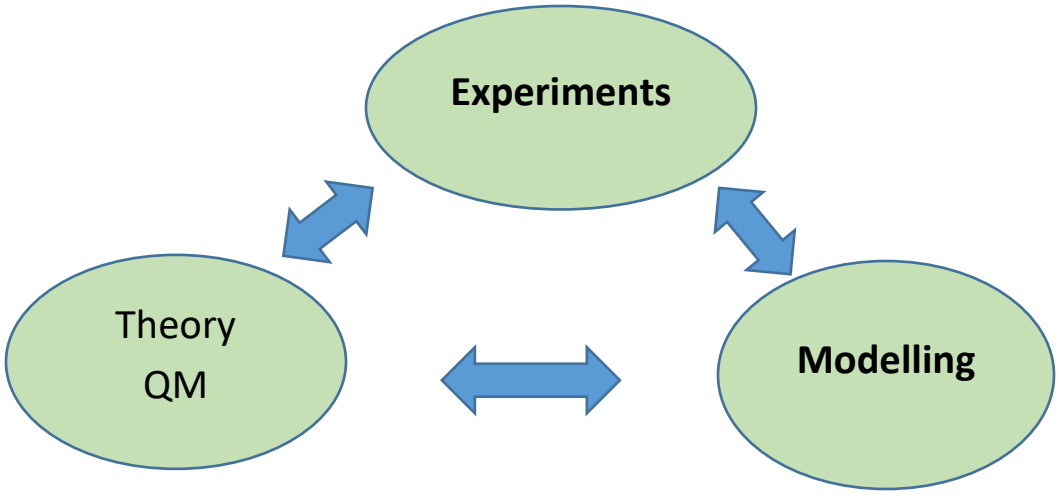
Computational Chemistry II 2023

Text book: Cramer: Essentials of Quantum Chemistry, Wiley (2 ed.)

Chapter 1. Intro

Introduction

This course will focus on **computational chemistry and computational materials screening**. The computational approach can be viewed as the third pillar of scientific knowledge. Naturally, experimental observations are the basis of Natural Science but we need the theoretical models to rationalize the experimental results. However, often the theoretical models are very complex and they need numerical methods to solve them. Even more, the fundamental theories will provide a basis to model complex real phenomena. In chemistry, almost all phenomena, in principle, can be derived from the equations of **quantum mechanics** (Schrödinger or Dirac equation). In practice, this is very difficult. First, the equations are difficult to solve and secondly the molecules are moving which may need to be taken into account.



Modelling

The strength of modelling is that it can provide **reliable** values of experimentally observable quantities. With modern computers, the modelling is often a faster method to get information than the experiments. In addition, in simulations every atom is known and very **detailed control** of the systems is achieved. This is very interesting when, for example, chemical reactions are studied. We can investigate the role of an individual atom to the reactivity. Because simulations can be faster than the experiments, they can be used to **screen potentially interesting materials** for experimental studies. This is complicated since the structure-property relations are often very complex and to unravel them a lot of good quality data is needed.

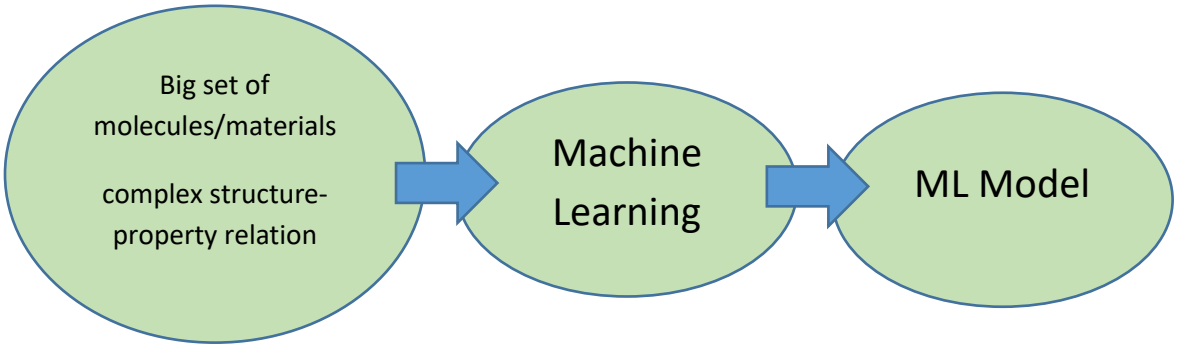
Quantum Chemistry

The Quantum Chemistry (QC) will focus mostly on the **electrons wave function** and **total energy** of the system. Most of the properties of a single molecule can be computed using QC but a group of molecules at finite temperature needs some kind of dynamical simulations, which can take the **entropy** into account. The dynamical simulations will use either Molecular Dynamics or Monte-Carlo simulations. These has been discussed in course Computational Chemistry I (CHEM-E4115) The basics of electrons wave function and total energy have been discussed in course CHEM-E4115 and CHEM-E4100, Quantum Mechanics and Spectroscopy, which are prerequisites for this course.

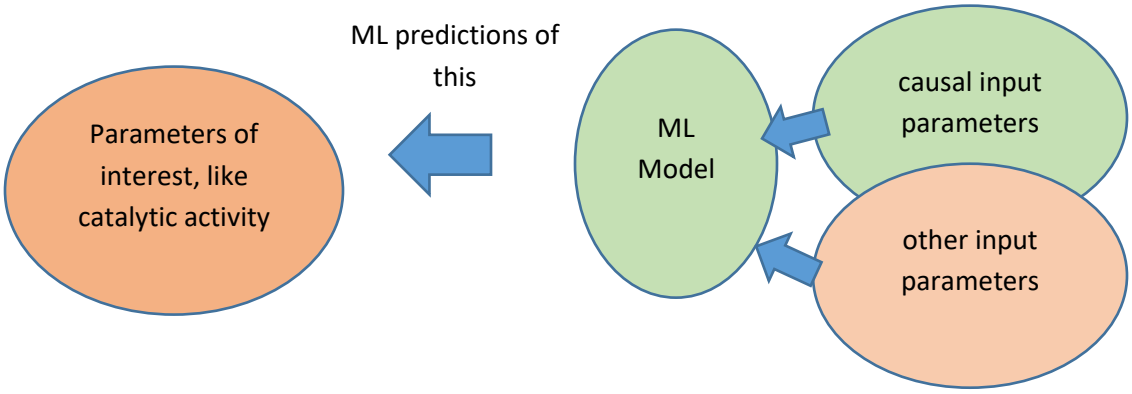
Machine Learning in Chemistry

Chemistry is rather ideal field for Machine Learning (ML) since almost all the problems can be described on atomic level and almost all properties of molecules can be computed. In fact, the ML can be used in any field of chemistry or chemical engineering if there is enough reliable data but this course focus on molecular systems. The principles are very similar in any ML application and in Aalto there are several ML courses.

The computational quantum chemical methods can provide a lot data of several molecules or materials. Typically, we want to optimize some useful property like catalytic activity, dielectric constant, etc. This interesting property will depend in a complex way of the materials composition and possibly structure. If we want to test a large amount of systems we need Machine Learning methods to find the correlations. In ideal case the ML model will be almost as accurate as the original Quantum Chemical calculations (or any original data). The ML model do not need time consuming QC step so it is much faster than the original data collection.

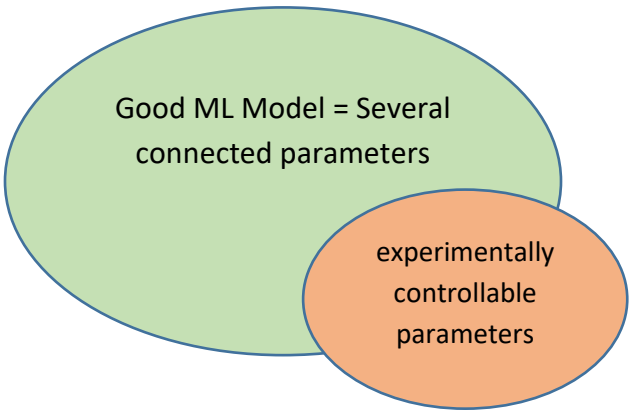


There are several challenges to this approach. The original data set need to be large, diverse enough and consistent to make a good basis for the ML. The ML is very good at finding correlations and it can handle several variables but not all the correlations have physically justified connection to the interesting variables (causality)



The ML methods **CANNOT DISTINGUISH** the causal input parameters form the others. It can only find the correlations. This will affect the predictive power of the ML methods in materials science. With good descriptors this is probably not a big problem.

Typically, the input of ML contain many variables and often variable that are not easy to control experimentally.



How to use the ML models to help the experiments? This usually has to be done in collaboration with experimental groups otherwise the ML predictions may be very difficult to fulfill in practice.

Total energy

Before going to the wave functions and Schrödinger equation we can look what can be learned from the **atomic total energy**. We assume that the electrons are at their **ground state** so the total energy depend only on the atomic coordinates.

$$E_{tot}^{el}(R_1,..R_N)$$

Here we have used the **Born-Oppenheimer approximation** in which we have assumed that the atoms are point-like from the point of view of the electrons. Here we do not take into account the atomic wave functions. This is a very good approximation at room temperature and for all atoms except hydrogen. The atomistic quantum effects can be later taken into account using harmonic

approximation near the minima of the total energy (This will be discussed later) or using the Path-Integral Methods.

Free energy

Unfortunately, the chemically important "energy" is the free energy and it contain the entropy term. In general, the entropy is difficult to compute but if the molecules do not interact strongly it can be estimated with gas phase thermodynamical methods. The Enthalpy (H) is very close to total energy but it also have small finite temperature correction

$$G_{tot}(T, P, R_1, .. R_N) = H_{tot}(T) - TS_{tot}(T, P)$$