

Lecture 40 — Computational Quantum Chemistry

Reading: Engel 4th ed., Chapter 15

Learning Objectives

- Describe the hierarchy of quantum chemistry methods (HF → CI → CC → DFT)
 - Explain basis sets and their role in computational chemistry
 - Distinguish semi-empirical, ab initio, and DFT methods
 - Interpret output from computational chemistry calculations
 - Understand the cost-accuracy tradeoffs of different methods
-

1. The Computational Chemistry Hierarchy

The Challenge

For any molecule beyond H_2^+ , we cannot solve the Schrödinger equation exactly. Computational chemistry develops systematic approximations of increasing accuracy.

Born-Oppenheimer Approximation

Separate nuclear and electronic motion: solve the electronic Schrödinger equation at fixed nuclear positions, creating a **potential energy surface** (PES).

$$\hat{H}_{\text{elec}} \Psi_{\text{elec}} = E_{\text{elec}}(R) \Psi_{\text{elec}}$$

2. Basis Sets

What Are They?

A **basis set** is a set of mathematical functions used to represent molecular orbitals. MOs are expanded as linear combinations of basis functions:

$$\phi_i = \sum_{\mu} c_{\mu i} \chi_{\mu}$$

Common Basis Sets

Basis Set	Description	Quality
STO-3G	Minimal: 3 Gaussians per Slater orbital	Low (qualitative)
3-21G	Split-valence	Basic
6-31G(d)	Polarized split-valence	Good for organics
6-311+G(d,p)	Triple-zeta + diffuse + polarization	High
cc-pVDZ/TZ/QZ	Correlation-consistent	Systematic convergence

Key Concepts

- **Slater-type orbitals (STOs):** Physically correct ($e^{-\zeta r}$) but integrals are hard
- **Gaussian-type orbitals (GTOs):** $e^{-\alpha r^2}$ — products of Gaussians are Gaussians, making integrals fast
- **Polarization functions:** Add higher- l functions (d on C, p on H) to describe bond distortion
- **Diffuse functions (+):** Spatially extended functions needed for anions, excited states, long-range interactions

[!NOTE] **Concept Check 40.1** In a basis set, what is the purpose of adding "polarization functions" (like d -functions on a Carbon atom)? How do they improve the description of a chemical bond?

3. Methods Hierarchy

Hartree-Fock (HF)

- Single Slater determinant, self-consistent field
- Captures ~99% of total energy, but misses **electron correlation**
- Scales as N^4 (number of basis functions)
- Useful for geometry optimization and qualitative trends

Post-Hartree-Fock Methods

Method	Accuracy	Scaling	Description
MP2	Good	N^5	Second-order perturbation theory for correlation
CCSD	High	N^6	Coupled cluster with singles and doubles
CCSD(T)	Very high ("gold standard")	N^7	+ perturbative triples
Full CI	Exact (in basis)	Exponential	All possible configurations — impractical except for tiny systems

Configuration Interaction (CI)

Expand the wavefunction as a sum of Slater determinants formed by promoting electrons from occupied to virtual orbitals:

$$\Psi = c_0\Phi_0 + \sum_a c_a\Phi_a + \sum_{ab} c_{ab}\Phi_{ab} + \dots$$

Density Functional Theory (DFT)

Based on the **Hohenberg-Kohn theorems**: the ground-state energy is a unique functional of the electron density $\rho(\mathbf{r})$:

$$E[\rho] = T[\rho] + V_{\text{ne}}[\rho] + J[\rho] + E_{\text{xc}}[\rho]$$

- The exchange-correlation functional $E_{\text{xc}}[\rho]$ is unknown and must be approximated
- Common functionals: B3LYP, PBE, ω B97X-D, M06-2X
- Scales as N^3 – N^4 — similar to HF but includes some correlation
- Excellent cost-accuracy ratio for most chemistry

Why DFT Dominates Modern Chemistry

- Accuracy of $\sim\pm 4$ kcal/mol for thermochemistry (B3LYP)
- Can handle hundreds of atoms
- Good geometries, frequencies, and relative energies
- Weakness: struggles with dispersion, charge-transfer, and strongly correlated systems

[!NOTE] **Concept Check 40.2** Density Functional Theory (DFT) is one of the most widely used methods in computational chemistry. What is the fundamental difference

between DFT and wavefunction-based methods like Hartree-Fock or Coupled Cluster?

4. What Calculations Can Tell Us

Property	Computational method
Geometry (r_e , angles)	Geometry optimization on PES
Vibrational frequencies	Second derivative of energy (Hessian)
Bond energies	Energy differences
Transition states	Saddle points on PES
Reaction mechanisms	Intrinsic reaction coordinate (IRC)
Electronic spectra	TD-DFT (time-dependent DFT)
NMR shifts	GIAO methods

5. Semi-Empirical Methods

- Parameterized versions of HF with reduced integral evaluation
- Methods: AM1, PM3, PM6, PM7
- Very fast — can handle thousands of atoms
- Less accurate — best for screening and qualitative insights

Machine Learning Potentials (Emerging)

Neural network potentials trained on DFT data can achieve DFT-level accuracy at force-field cost, enabling quantum-level molecular dynamics on large systems.

Key Equations Summary

Equation	Expression
LCAO expansion	$\phi_i = \sum_{\mu} c_{\mu i} \chi_{\mu}$
HF energy	$E_{\text{HF}} = \sum_i h_i + \frac{1}{2} \sum_{ij} (J_{ij} - K_{ij})$
DFT energy	$E[\rho] = T_s[\rho] + V_{\text{ne}}[\rho] + J[\rho] + E_{\text{xc}}[\rho]$
Correlation energy	$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}$

Recent Literature Spotlight

"Machine Learning Accurate Exchange and Correlation Functionals of the Electronic Density" S. Dick, M. Fernandez-Serra, *Nature Communications*, **2020**, 11, 3509. [DOI](#)

This paper demonstrates that neural networks can learn exchange-correlation functionals directly from accurate quantum chemical data, achieving DFT-level accuracy without fitting to analytic functional forms. The approach builds on the Kohn-Sham framework taught in this lecture — using the electron density $\rho(\mathbf{r})$ as the fundamental variable — and shows how machine learning can systematically improve upon the local density and generalized gradient approximations.

Practice Problems

- Basis set convergence.** Explain why results improve as the basis set increases from STO-3G \rightarrow 6-31G(d) \rightarrow 6-311+G(2d,2p) \rightarrow cc-pVQZ. What is the **basis set limit**?
 - Method comparison.** For the reaction $\text{H}_2 + \text{F}_2 \rightarrow 2\text{HF}$, rank the following methods by expected accuracy: HF/6-31G(d), B3LYP/6-31G(d), CCSD(T)/cc-pVTZ. Explain why HF fails most.
 - DFT functional choice.** You want to study (a) conformational energies of a drug molecule, (b) a transition metal complex, (c) weak van der Waals interactions in a protein-ligand complex. Which DFT functional would you choose for each?
-

Next lecture: NMR Spectroscopy