

# Many-body methods in quantum chemistry

T. Helgaker, Department of Chemistry, University of Oslo, Norway

P. Jørgensen, J. Olsen, University of Aarhus, Denmark

W. Klopper, University of Karlsruhe, Germany

Computational advances in the nuclear many-body problem

March 11–13 2004

Department of Physics and Center of Mathematics for Applications

University of Oslo

Most calculations presented here were carried out with DALTON

<http://www.kjemi.uio.no/software/dalton/dalton.html>

## Experimental vs. theoretical chemistry

- Chemistry is traditionally an experimental science!
- A theoretical calculation provides numbers but no understanding!

“Every attempt to employ mathematical methods in the study of chemical questions must be considered **profoundly irrational**. If mathematical analysis should ever hold a prominent place in chemistry—an aberration which is happily impossible—it would occasion a rapid and widespread degradation of that science.”

August Comte, 1748–1857

- Quantum chemistry is based on a deep understanding of nature!

“The more progress sciences make, the more they tend to enter the domain of mathematics, which is a kind of center to which they all converge. We may even judge the **degree of perfection** to which a science has arrived by the facility with which it may be submitted to calculation.”

Adolphe Quetelet, 1796–1874

– what exactly constitutes this understanding?

## Molecules and quantum mechanics

- A molecule is a collection of  $N$  particles of different masses and charges, bound by predominantly **electrostatic forces**.

- Their behaviour is described by the wave function:

$$\Psi = \Psi(x_1, y_1, z_1, s_1, \dots, x_N, y_N, z_N, s_N, t) \leftarrow \text{a function of } 4N \text{ coordinates and of time}$$

– the wave function contains all information about the system

- The wave function is a solution to a differential equation: **the Schrödinger equation**,

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H} \Psi$$

– the spin and fieldfree nonrelativistic **Hamiltonian operator** is given by

$$\hat{H} = - \sum_i \frac{\hbar^2}{2m_i} \left( \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right) + \frac{e^2}{4\pi\epsilon_0} \sum_{i>j} \frac{Z_i Z_j}{r_{ij}}$$

– in the relativistic case, we solve the Dirac equation instead.

- We are often (but by no means always) interested in **stationary states**:

$$\Psi_n = \exp(-iE_n t/\hbar) \psi_n, \quad \hat{H} \psi_n = E_n \psi_n$$

– **quantization**: boundary conditions often restrict the solutions to discrete energies

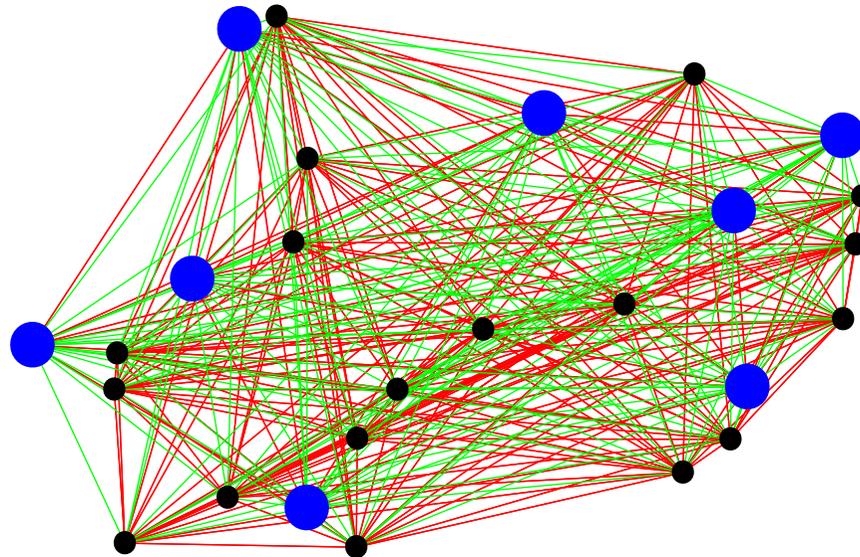
## The many-body problem

- Thus, about 1930, the mathematical foundation of chemistry was understood:

“The underlying physical laws necessary for the mathematical theory of a large part of physics and **the whole of chemistry** are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.”

P. A. M. Dirac, 1929

- The complications arise from the large number of particles:



## Computers and computational chemistry

- With the development of **modern computers** over the last 50 years, the molecular many-body problem has become tractable.
- Today, a large number of **chemical problems** have become amenable to calculation:
  - molecular structure and spectroscopic constants
  - reaction enthalpies and equilibrium constants
  - reactivity, reaction rates, and dynamics
  - interaction with applied electromagnetic fields and radiation
- In 1995 computation constituted about **15% of all chemical research**:
  - this proportion increases by about 1% every year
  - this growth is likely to flatten out in about 2035, when half of all chemistry has become computational
- Nowadays, quantum-chemical calculations are routinely carried out by **nonspecialists**:
  - chemists have for many years been among the heaviest users of supercomputers
  - mostly, this number crunching has been directed towards the approximate solution of the electronic Schrödinger equation for molecular systems

## The holy grail of quantum chemistry: chemical accuracy

- Quantum chemistry has been important in providing understanding and many **qualitative models** in chemistry
  - such models are good and useful
  - however, they do not constitute the bread and butter of quantum chemistry
- If quantum chemistry is to have a decisive impact on chemistry, we need to provide tools that can **compete with experiment**
  - ideally, our results should be as accurate as experiment: **chemical accuracy**
  - if we cannot consistently provide high accuracy, we will soon be out of business
- We must **control our accuracy** and approach the exact solution in a systematic manner
  - build **hierarchies of approximation**
  - perform careful **benchmarking**
- In **ab initio** theory, no empirical parameters except the fundamental constants are used
  - empirical parameters abound in simplified models (not treated here)
- **Error cancellation**—the Achilles' heel of quantum chemistry
  - error cancellation often leads to a **fortuitous agreement** with experiment
  - such cancellations are treacherous and must be avoided or carefully treated
  - **the right answer for the right reason**

## The Born–Oppenheimer approximation: separation of electrons and nuclei

- Molecules consist of two types of particles:
  - negatively charged, light **electrons** (Fermions) and positively charged, heavy **nuclei**
- In the **Born–Oppenheimer approximation**, the light electrons react instantaneously to the motion of the nuclei
  - for each nuclear geometry  $\mathbf{R}$ , the electrons are described by an **electronic wave function**, which is the eigenfunction of an **electronic Hamiltonian** (atomic units):

$$\begin{aligned}\hat{H}_{\text{el}}(\mathbf{R}) \psi_n(\mathbf{r}; \mathbf{R}) &= E_n(\mathbf{R}) \psi_n(\mathbf{r}; \mathbf{R}) \\ \hat{H}_{\text{el}}(\mathbf{R}) &= -\frac{1}{2} \sum_i \nabla_i^2 - \sum_{iK} \frac{Z_K}{|\mathbf{r}_i - \mathbf{R}_K|} + \sum_{i>j} \frac{1}{r_{ij}} + \sum_{K>L} \frac{Z_K Z_L}{R_{KL}}\end{aligned}$$

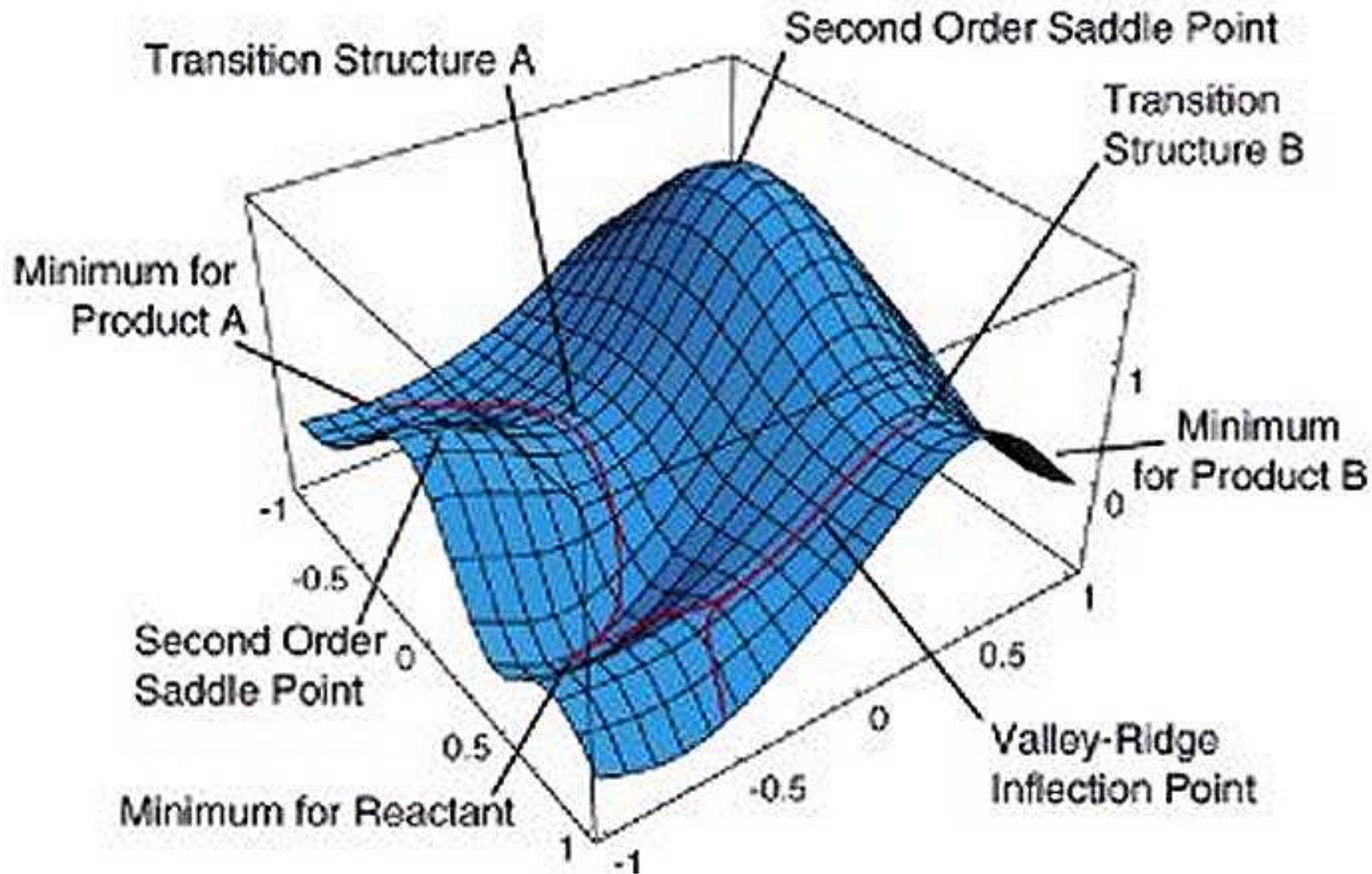
- its efficient solution for a given  $\mathbf{R}$  is the subject of **electronic-structure theory**
- The electronic energy  $E_n(\mathbf{R})$  represents a multidimensional **potential energy surface (PES)**, which in turn determines the motion of the nuclei:

$$\left[ -\frac{1}{2} \sum_K M_K^{-1} \nabla_K^2 + E_n(\mathbf{R}) \right] \Phi_{nv}(\mathbf{R}) = \mathcal{E}_{nv} \Phi_{nv}(\mathbf{R})$$

- To a very good approximation, the total **molecular wave function** is a product of the electronic and nuclear wave functions:

$$\Psi_{nv}(\mathbf{r}, \mathbf{R}) = \psi_n(\mathbf{r}; \mathbf{R}) \Phi_{nv}(\mathbf{R})$$

The molecular potential energy surface (PES)



## The electronic-structure problem: approximate wave functions

- We are now ready to consider the problem of constructing electronic wave functions.
- Some exact properties that may or may not be incorporated in approximate models:
  - for bound states, the exact wave function is square-integrable:

$$\langle \Psi | \Psi \rangle = \int \Psi^* (\mathbf{x}) \Psi (\mathbf{x}) d\mathbf{x} = 1 \quad \leftarrow \text{always satisfied}$$

- the exact wave function is antisymmetric in the electron coordinates:

$$P_{ij} \Psi = -\Psi \quad \leftarrow \text{always satisfied}$$

- the exact wave function is variational (i.e., the energy is stable):

$$\langle \delta \Psi | \Psi \rangle = 0 \quad \Rightarrow \quad \langle \delta \Psi | \hat{H} | \Psi \rangle = 0 \quad \leftarrow \text{not always satisfied}$$

- the exact wave function is size-extensive, implying that:

$$\hat{H} = \sum_i \hat{H}_i \quad \Rightarrow \quad E = \sum_i E_i \quad \leftarrow \text{not always satisfied}$$

- the exact nonrelativistic wave function is a spin eigenfunction:

$$\hat{S}^2 \Psi = S(S+1)\Psi; \quad \hat{S}_z \Psi = M_S \Psi \quad \leftarrow \text{not always satisfied}$$

- Many models and approximations have been proposed over the years
  - we shall focus on the coupled-cluster hierarchy of wave-function models

## The Hartree–Fock model

- For a system of **noninteracting** electrons, the exact wave function takes the form of **an antisymmetric product of molecular orbitals MOs** (a Slater determinant):

$$\Phi = \hat{A} \prod_{i=1}^N \phi_i(\mathbf{x}_i), \quad \langle \phi_i | \phi_j \rangle = \delta_{ij}; \quad |\text{HF}\rangle = \left( \prod_{i=1}^N a_i^\dagger \right) |\text{vac}\rangle, \quad [a_i^\dagger, a_j^\dagger]_+ = 0$$

- for **interacting** electrons, the exact wave function cannot be written in product form:
- still, we may take this form as a useful **ansatz** for an approximate description
- to make the most of this *ansatz*, we invoke the **variation principle** and minimize the expectation value of the energy with respect to the form of the MOs:

$$E = \min_{\phi_i} \frac{\langle \Phi | \hat{H} | \Phi \rangle}{\langle \Phi | \Phi \rangle} \geq E_{\text{exact}} \quad \leftarrow \text{nonrelativistic Hamiltonian bounded from below}$$

- this approach constitutes **the Hartree–Fock (HF) model**
- In the Hartree–Fock **independent-particle** description, each electron is
  - described by taking into account only the mean effect of all other electrons
  - affected by the average rather than instantaneous positions of the remaining electrons
- The HF model is the **cornerstone of *ab initio* theory**:
  - it constitutes a useful, qualitative model on its own; applicable to large systems
  - it forms the starting point for more accurate models

## Electron correlation and virtual excitations

- **electron correlation:**

- to improve upon the HF model, we must take into account the instantaneous interactions among the electrons
- in real space, the electrons are constantly being scattered by **collisions**
- in the orbital picture, these collisions manifest themselves as **excitations** from occupied to virtual (unoccupied) spin orbitals

- **double excitations:**

- the most important events are collisions between two electrons
- in the orbital picture, such an event corresponds to an excitation from two occupied to two virtual spin orbitals, known as **pair excitations** or **double excitations**
- consider the following **double excitation operator**:

$$\hat{X}_{ij}^{ab} = t_{ij}^{ab} a_b^\dagger a_a^\dagger a_i a_j; \quad [a_p, a_q]_+ = 0, \quad [a_p^\dagger, a_q^\dagger]_+ = 0, \quad [a_p, a_q^\dagger]_+ = \delta_{pq}$$

- the amplitude  $t_{ij}^{ab}$  represents the probability that the electrons in  $\phi_i$  and  $\phi_j$  will interact and be excited to  $\phi_a$  and  $\phi_b$
- by applying  $1 + \hat{X}_{ij}^{ab}$  to the HF state, we obtain an improved, **correlated description** of the electrons:

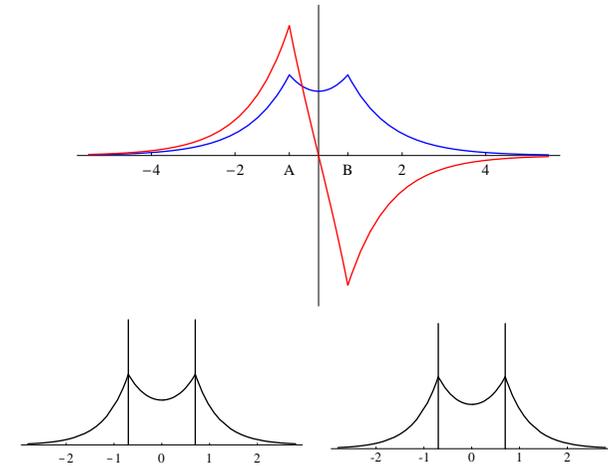
$$|\text{HF}\rangle \rightarrow (1 + \hat{X}_{ij}^{ab})|\text{HF}\rangle$$

## Example: electron correlation in H<sub>2</sub>

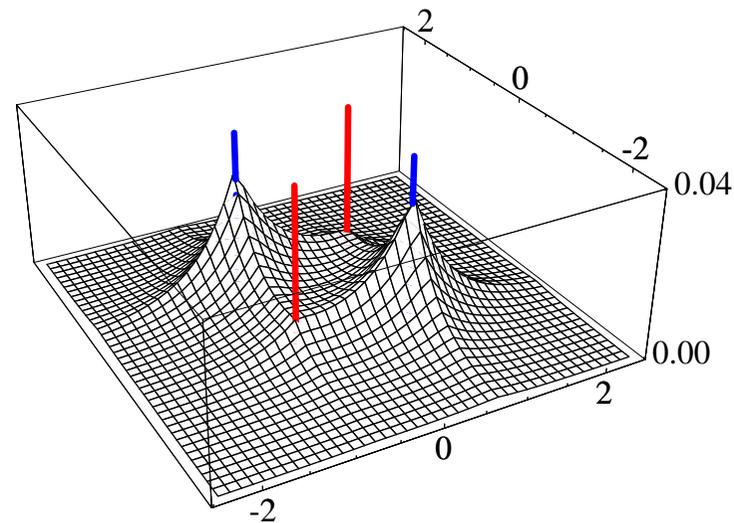
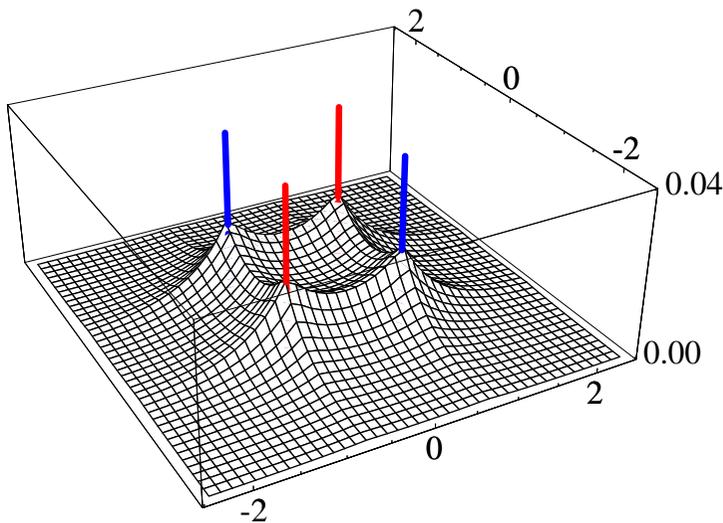
- Consider the effect of a double excitation in H<sub>2</sub>:

$$|1\sigma_g^2\rangle \rightarrow (1 + \hat{X}_{gg}^{uu})|1\sigma_g^2\rangle = |1\sigma_g^2\rangle - 0.11|1\sigma_u^2\rangle$$

- The one-electron density  $\rho(z)$  is hardly affected:



- The two-electron density  $\rho(z_1, z_2)$  changes dramatically:



## Coupled-cluster theory

- In **coupled-cluster (CC) theory**, we generate the correlated state from the HF reference state by applying all possible operators  $1 + \hat{X}_\mu$ :

$$|\text{CC}\rangle = \left[ \prod_{ai} (1 + \hat{X}_i^a) \right] \left[ \prod_{abij} (1 + \hat{X}_{ij}^{ab}) \right] \cdots |\text{HF}\rangle; \quad [\hat{X}_\mu, \hat{X}_\nu] = 0$$

- It is reasonable to assume that **lower-order excitations** are more important than higher-order ones.
- **Double excitations** are particularly important, arising from pair interactions.
- This classification provides a **hierarchy of ‘truncated’ CC wave functions**:
  - **CCSD**: CC with all single and double excitations ( $n^6$ )
  - **CCSDT**: CC with all single, double, and triple excitations ( $n^8$ )
  - **CCSDTQ**: CC with up to quadruple excitations ( $n^{10}$ )
  - **CCSDTQ5**: CC with up to quintuple excitations ( $n^{12}$ )
- The advantage of the CC product form is that it enables us to approximate the **full configuration-interaction (FCI) wave function** in a compact manner:

$$|\text{FCI}\rangle = \sum_{\mu} C_{\mu} |\mu\rangle \quad \leftarrow \text{intractable sum over all possible determinants}$$

- Errors are typically reduced by a factor of three to four with each new excitation level.

## Connected and disconnected clusters

- Expanding the CCSD product state, we obtain:

$$\begin{aligned} |\text{CCSD}\rangle &= \left[ \prod_{ai} (1 + \hat{X}_i^a) \right] \left[ \prod_{abij} (1 + \hat{X}_{ij}^{ab}) \right] |\text{HF}\rangle \\ &= |\text{HF}\rangle + \sum_{ai} \hat{X}_i^a |\text{HF}\rangle + \sum_{abij} (\hat{X}_{ij}^{ab} + \hat{X}_i^a \hat{X}_j^b) |\text{HF}\rangle + \dots \end{aligned}$$

- The doubly-excited determinants have two distinct contributions:
  - from pure double excitations: **connected doubles**
  - from products of single excitations: **disconnected doubles**
- In large systems, the disconnected excitations become more important.
- In **configuration-interation (CI) theory**, we retain only the connected excitations:

$$|\text{CISD}\rangle = \left( 1 + \sum_{ai} \hat{X}_i^a + \sum_{abij} \hat{X}_{ij}^{ab} \right) |\text{HF}\rangle$$

- without the disconnected excitations, the wave function is no longer **size-extensive**
- the CISD model works only for few electrons (about 10)
- for large systems, it is no better than the Hartree–Fock model!
- The CI model has been abandoned in favor of the CC model.

## The CC exponential ansatz

- The CC wave function is usually written in exponential form:

$$|\text{CC}\rangle = \exp(\hat{T})|\text{HF}\rangle; \quad \hat{T} = \sum_{ai} \hat{X}_i^a + \sum_{abij} \hat{X}_{ij}^{ab} + \dots$$

- Equivalence with the product form is easily established since, for example:

$$\exp(\hat{X}_i^a) = 1 + \hat{X}_i^a + \frac{1}{2} \hat{X}_i^a \hat{X}_i^a + \dots = 1 + \hat{X}_i^a \quad \Leftarrow \quad \hat{X}_i^a \hat{X}_i^a = 0$$

- For technical reasons, the CC energy is not determined variationally:

$$E \neq \min \frac{\langle \text{CC} | \hat{H} | \text{CC} \rangle}{\langle \text{CC} | \text{CC} \rangle}$$

- Multiplying the CC Schrödinger equation in the form

$$\exp(-\hat{T}) \hat{H} \exp(\hat{T}) |\text{HF}\rangle = E |\text{HF}\rangle \quad \leftarrow \text{similarity-transformed Hamiltonian}$$

from the left by  $\langle \text{HF} |$  and the excited determinants  $\langle \mu |$ , we obtain

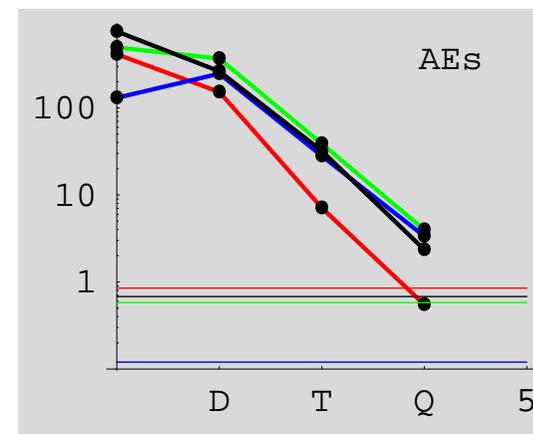
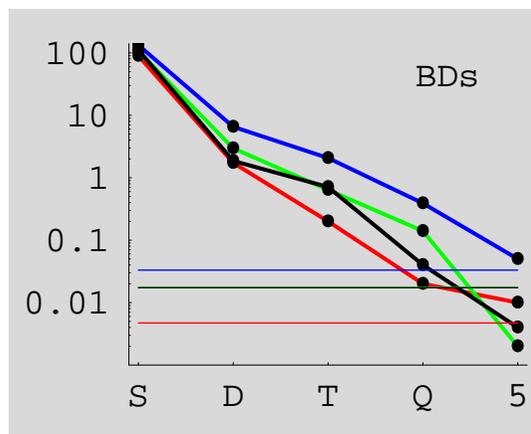
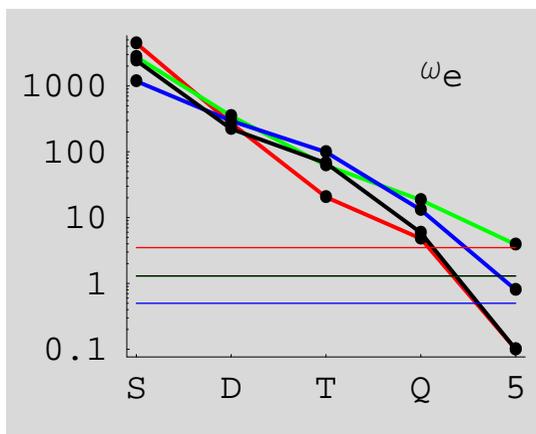
$$\langle \text{HF} | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) |\text{HF}\rangle = E \quad \leftarrow \text{energy (not an upper bound)}$$

$$\langle \mu | \exp(-\hat{T}) \hat{H} \exp(\hat{T}) |\text{HF}\rangle = 0 \quad \leftarrow \text{amplitudes}$$

- From these equations, the CC energy and amplitudes are determined.

## Excitation-level convergence

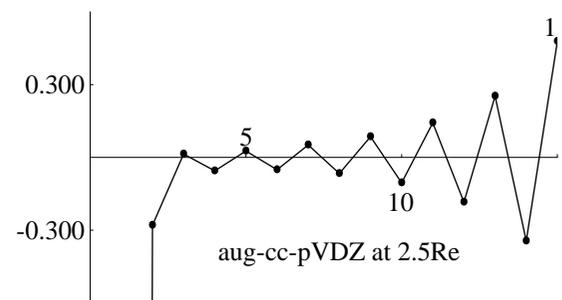
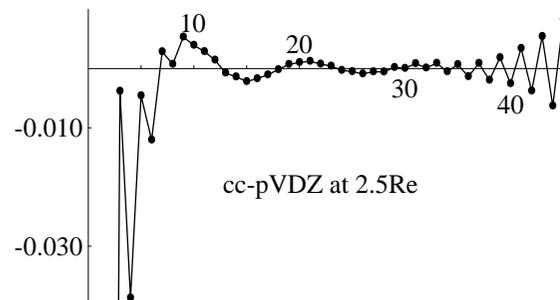
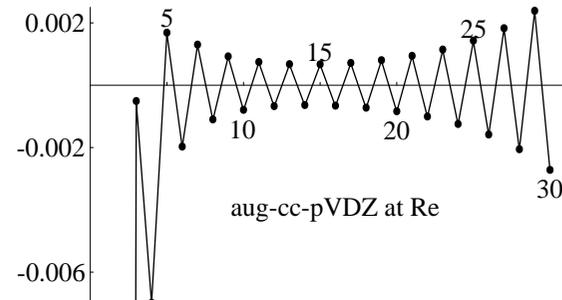
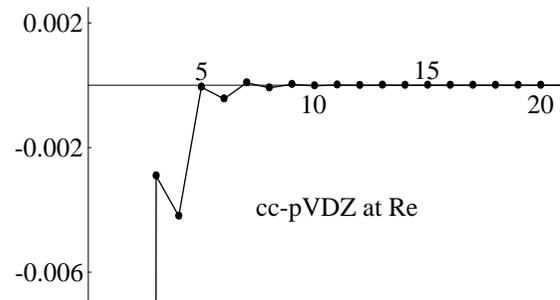
- Log plots of contributions to harmonic frequencies ( $\text{cm}^{-1}$ ), bond lengths (pm), and atomization energies (kJ/mol):



- color code: HF (red), N<sub>2</sub> (green), F<sub>2</sub> (blue), and CO (black)
- straight lines indicate first-order relativistic corrections
- Excitation-level convergence is approximately exponential
- Relativity becomes important beyond connected quadruples
- Chemical accuracy is achieved at the CCSDT or CCSDTQ level of theory

## Many-body perturbation theory: an alternative to coupled-cluster theory

- Alternatively, we may determine the coupled-cluster amplitudes by perturbation theory
- Unfortunately, the resulting **many-body perturbation theory (MBPT)** series is frequently divergent, even in very simple cases
- Here are some examples for the HF molecule (10 electrons):



- However, to lowest order, the MBPT expansion is frequently very useful (MP2)

## Example of a quantum-chemical calculation: the atomization energy of CO

- The energies of interest in chemistry are often small compared with the total energies.
- Example: the atomization energy (AE) of CO (**exp. 1071.8 kJ/mol**)

	$E_C$	+	$E_O$	−	$E_{CO}$	=	$D_{CO}$	err.
HF	−98964.9	−	196437.1	+	296132.2	=	730.1	−341.7
SD	−388.4	−	639.6	+	1350.2	=	322.1	−19.6
(T)	−7.7	−	11.9	+	54.2	=	34.6	15.0
vib.	0.0	+	0.0	−	12.9	=	−12.9	2.1
rel.	−40.1	−	139.0	+	177.1	=	−2.0	−0.1
tot.	−99401.1	−	197227.6	+	297700.8	=	1071.9	

- The AE constitutes less than 1% of the total energy.

- Bonds are broken—correlation is important:

	HF	SD	(T)
% of $E$	99.5	0.5	0.01
% of AE	67	30	3

- Nevertheless, the coupled-cluster convergence is rapid:  
the error is reduced by an order of magnitude with each new level.
- The zero-point vibrational energy is of the same order of magnitude as the triples energy but in the opposite direction.

## One-electron basis functions

- Our correlated  $N$ -electron wave functions are linear combinations of products of one-electron functions (MOs)
- Each MO is expanded in simple Gaussian functions:

$$G_{ijk}(r_A, \alpha) = x_A^i y_A^j z_A^k \exp(-\alpha r_A^2)$$

- these basis functions are fixed to the atomic nuclei and called atomic orbitals (AOs)
- their functional form is taken to resemble that of the hydrogenic eigenfunctions
- however, a Gaussian rather than exponential radial form is chosen for convenience
- the exponents  $\alpha$  are fixed in atomic calculations and tabulated (basis sets)
- Basis sets of increasing size:
  - minimal or single-zeta (SZ) basis sets:
    - \* one set of GTOs for each occupied atomic shell ( $2s1p$ )
    - \* gives a rudimentary description of electron structure
  - double-zeta (DZ) basis sets:
    - \* two sets of GTOs for each occupied atomic shell ( $3s2p1d$ )
    - \* sufficient for a qualitative description of the electron system
  - triple-zeta (TZ), quadruple-zeta (QZ) and larger basis sets:
    - \* needed for a quantitative description of the electronic system

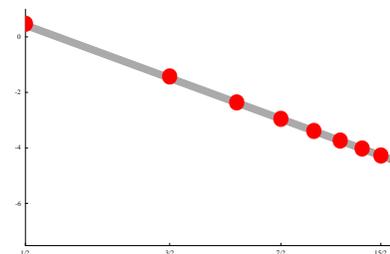
# The principal expansion and correlation-consistent basis sets

- The energy contribution from each AO in large CI calculations on helium:

$$\varepsilon_{nlm} \approx n^{-6} \quad \leftarrow \text{Carroll et al. (1979)}$$

- The principal expansion:** include all AOs belonging to the same shell simultaneously, in order of increasing principal quantum number  $n$ :

$$\varepsilon_n \approx n^2 n^{-6} = n^{-4}$$



- Practical realization: the correlation-consistent basis sets **cc-pVXZ (Dunning, 1989)**
- Energy-optimized AOs are added one shell at a time:

SZ	cc-pVDZ	cc-pVTZ	cc-pVQZ	number of AOs
$2s1p$	$3s2p1d$	$4s3p2d1f$	$5s4p3d2f1g$	$\frac{1}{3}(X+1)(X+\frac{3}{2})(X+2) \propto X^3$
	$+3s3p3d$	$+4s4p4d4f$	$+5s5p5d5f5g$	$(X+1)^2 \propto X^2$

- The error in the energy is equal to the contributions from all omitted shells:

$$\Delta E_X \approx \sum_{n=X+1}^{\infty} n^{-4} \approx X^{-3} \approx N^{-1} \approx T^{-1/4}$$

- Each new digit in the energy therefore costs 10000 times more CPU time!

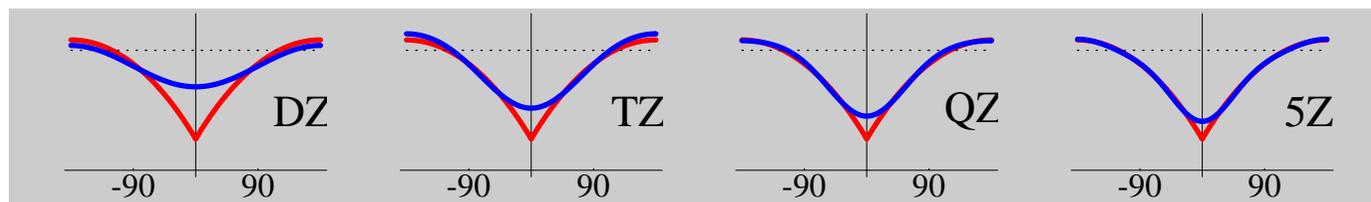
1 minute     $\rightarrow$     1 week     $\rightarrow$     200 years

## Basis-set convergence

- Convergence of the contributions to the atomization energy of CO (kJ/mol):

	$N_{\text{bas}}$	HF ( $n^4$ )		SD ( $n^6$ )		(T)( $n^7$ )		CCSD(T)	error
cc-pCVDZ	36	710.2	+	277.4	+	24.5	=	1012.1	-74.8
cc-pCVTZ	86	727.1	+	297.3	+	32.6	=	1057.0	-29.9
cc-pCVQZ	168	730.3	+	311.0	+	33.8	=	1075.1	-11.8
cc-pCV5Z	290	730.1	+	316.4	+	34.2	=	1080.7	-6.2
cc-pcV6Z	460	730.1	+	318.8	+	34.4	=	1083.3	-3.6
limit	$\infty$	730.1	+	322.1	+	34.6	=	1086.9	0.0

- The **doubles** converge very slowly—chemical accuracy requires 460 AOs (6Z)!
- The **Hartree–Fock** and **triples** contributions are less of a problem.
- The slow convergence arises from a poor description of short-range (dynamical) correlation in the orbital approximation (since  $r_{ij}$  is not present in the wave function):



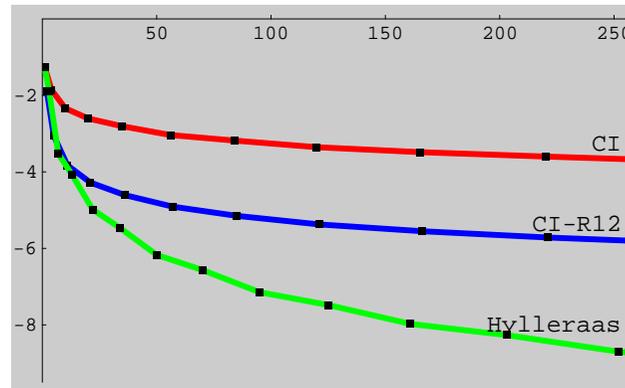
- How can we solve this problem?

## Solutions to slow basis-set convergence

### 1. Use explicitly correlated methods!

- Include **interelectronic distances**  $r_{ij}$  in the wave function (Hylleraas 1928):

$$\Psi_{R12} = \sum_K C_K \Phi_K + C_{R12} \Phi_0$$



- We use CCSD-R12 (Klopper and Kutzelnigg, 1987) for benchmarking

### 2. Use basis-set extrapolation!

- Exploit the smooth convergence  $E_\infty = E_X + AX^{-3}$  to extrapolate to basis-set limit

$mE_h$	DZ	TZ	QZ	5Z	6Z	R12
plain	194.8	62.2	23.1	10.6	6.6	1.4
extr.		21.4	1.4	0.4	0.5	

- The formula is linear and contains no parameters; applicable to many properties

### 3. Use density-functional theory!

## *Ab initio* hierarchies

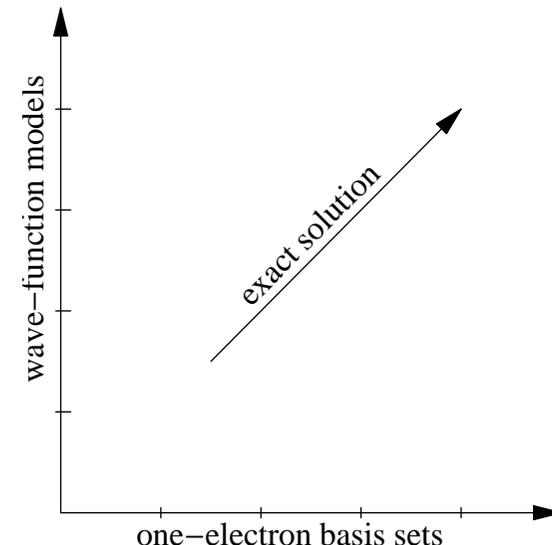
- The quality of *ab initio* calculations is determined by the description of
  1. the  $N$ -electron space (wave-function model);
  2. the one-electron space (basis set).
- In each space, there is a hierarchy of levels of increasing complexity:

1.  $N$ -electron hierarchy:  
coupled-cluster excitation levels

HF, CCSD, CCSDT, CCSDTQ, ... FCI

2. one-electron hierarchy:  
correlation-consistent basis sets

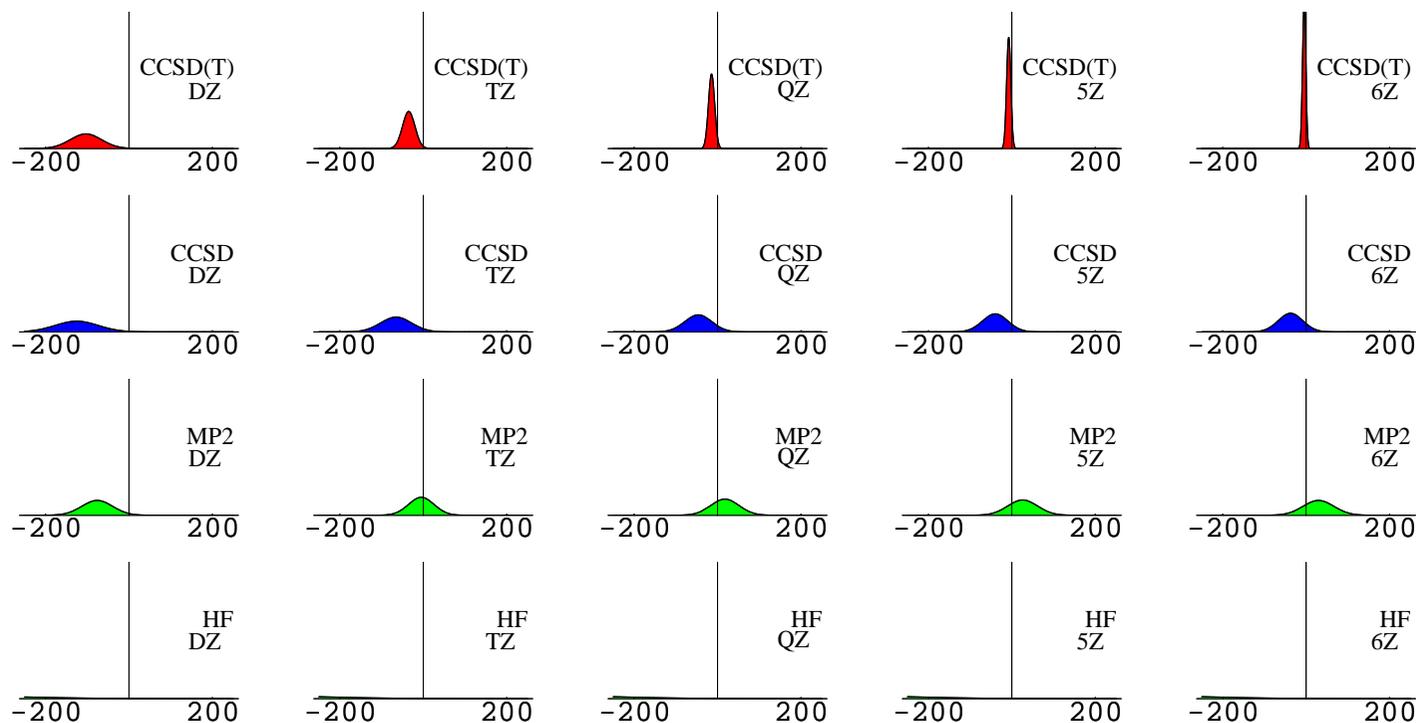
DZ, TZ, QZ, 5Z, 6Z, ...



- The quality is systematically improved upon by going up in the hierarchies.

## Example of an *ab initio* hierarchy: atomization energies

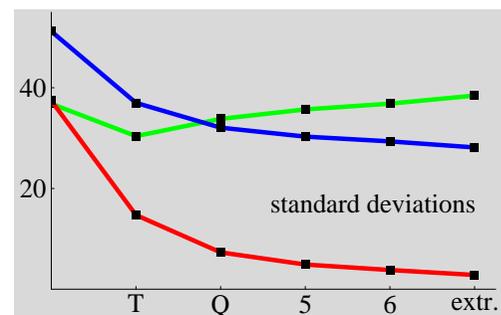
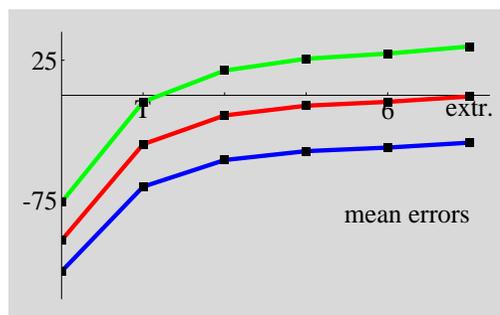
- Normal distribution of errors (kJ/mol) for RHF, MP2, CCSD, and CCSD(T) in the cc-pCVXZ basis sets:



- The *ab initio* hierarchy of wave-function theory enables it to approach the exact solution in an orderly, controlled manner.
- In this aspect, it distinguishes itself from density-functional theory.

## Atomization energies (AEs)

- Statistics based 20 closed-shell organic molecules (kJ/mol)



- AEs increase with excitation level in the coupled-cluster hierarchy:

$$\text{HF} < \text{CCSD} < \text{CCSD(T)} < \text{MP2}$$

- Mean abs. cc-pcV6Z err. (kJ/mol): 423 (HF), 37 (MP2), 30 (CCSD), 4 (CCSD(T))
- AEs increase with cardinal number.
- CCSD(T) performs excellently, but DZ and TZ are inadequate:

kJ/mol	DZ	TZ	QZ	5Z	6Z
raw	103.1	34.0	13.5	6.6	4.1
extrapolated		14.5	1.7	0.9	0.8

## Comparison of CCSD(T) and experimental AEs

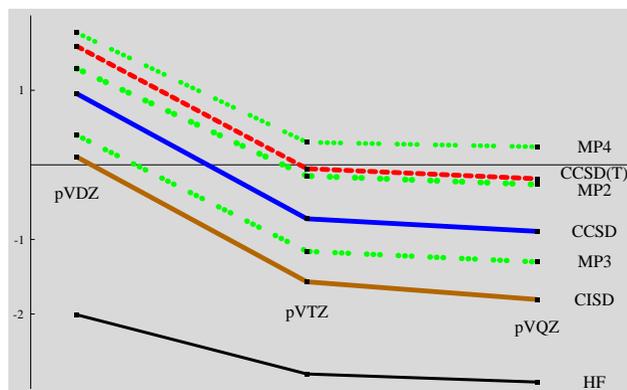
	cc-pCVQZ		cc-pCV6Z		exp.	vib.	rel.
F <sub>2</sub>	153.4	-9.9	161.1	-2.2	163.4(06)	5.5	-3.3
H <sub>2</sub>	456.6	-1.5	458.1	0.1	458.0(00)	26.0	0.0
HF	586.1	-7.0	593.3	0.1	593.2(09)	24.5	-2.5
O <sub>3</sub>	583.6	-32.6	605.5	-10.7	616.2(17)	17.4	-3.9
HOF	649.5	-25.4	662.9	-12.0	674.9(42)	35.9	-3.5
CH <sub>2</sub>	751.3	-5.7	757.9	0.9	757.1(22)	43.2	-0.7
HNO	842.7	-18.8	860.4	-1.1	861.5(03)	35.8	-2.1
N <sub>2</sub>	936.3	-19.9	954.9	-1.3	956.3(02)	14.1	-0.6
H <sub>2</sub> O	963.5	-11.8	975.5	0.2	975.3(01)	55.4	-2.1
CO	1075.5	-11.2	1086.9	0.2	1086.7(05)	12.9	-2.0
NH <sub>3</sub>	1232.7	-15.1	1247.4	-0.5	1247.9(04)	89.0	-1.1
HCN	1294.1	-18.6	1311.0	-1.7	1312.8(26)	40.6	-1.4
CH <sub>2</sub> O	1552.4	-14.2	1568.0	1.4	1566.6(07)	69.1	-2.7
CO <sub>2</sub>	1612.3	-20.1	1633.2	0.7	1632.5(05)	30.3	-4.2
C <sub>2</sub> H <sub>2</sub>	1681.0	-16.8	1697.1	-0.8	1697.8(10)	68.8	-1.9
CH <sub>4</sub>	1749.9	-9.4	1759.4	0.1	1759.3(06)	115.9	-1.2
C <sub>2</sub> H <sub>4</sub>	2343.6	-16.2	2360.8	1.0	2359.8(10)	132.2	-2.1

Reaction enthalpies (kJ/mol)

	DFT	CCSD(T)	exp.	vib.	rel.
$\text{CO} + \text{H}_2 \rightarrow \text{CH}_2\text{O}$	-34	-23	-21(1)	30.2	0.7
$\text{H}_2\text{O} + \text{F}_2 \rightarrow \text{HOF} + \text{HF}$	-119	-118	-129(4)	-0.6	0.5
$\text{N}_2 + 3\text{H}_2 \rightarrow 2\text{NH}_2$	-166	-165	-164(1)	86.1	1.6
$\text{C}_2\text{H}_2 + \text{H}_2 \rightarrow \text{C}_2\text{H}_4$	-208	-206	-203(2)	37.4	0.2
$\text{CO}_2 + 4\text{H}_2 \rightarrow \text{CH}_4 + 2\text{H}_2\text{O}$	-211	-244	-244(1)	92.7	1.0
$\text{CH}_2\text{O} + 2\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$	-234	-250	-251(1)	50.3	0.5
$\text{CO} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$	-268	-273	-272(1)	80.5	1.2
$\text{HCN} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{NH}_2$	-320	-321	-320(3)	85.4	0.9
$\text{HNO} + 2\text{H}_2 \rightarrow \text{H}_2\text{O} + \text{NH}_2$	-429	-446	-444(1)	56.7	1.0
$\text{C}_2\text{H}_2 + 3\text{H}_2 \rightarrow 2\text{CH}_4$	-450	-447	-446(2)	85.2	0.4
$\text{CH}_2 + \text{H}_2 \rightarrow \text{CH}_4$	-543	-543	-544(2)	46.8	0.4
$\text{F}_2 + \text{H}_2 \rightarrow 2\text{HF}$	-540	-564	-563(1)	17.6	1.6
$2\text{CH}_2 \rightarrow \text{C}_2\text{H}_4$	-845	-845	-844(3)	45.8	0.6
$\text{O}_3 + 3\text{H}_2 \rightarrow 3\text{H}_2\text{O}$	-909	-946	-933(2)	72.5	2.3
mean & std. dev.	8 & 13	-1&5			
mean abs. & max. abs.	11 & 33	3 & 13			

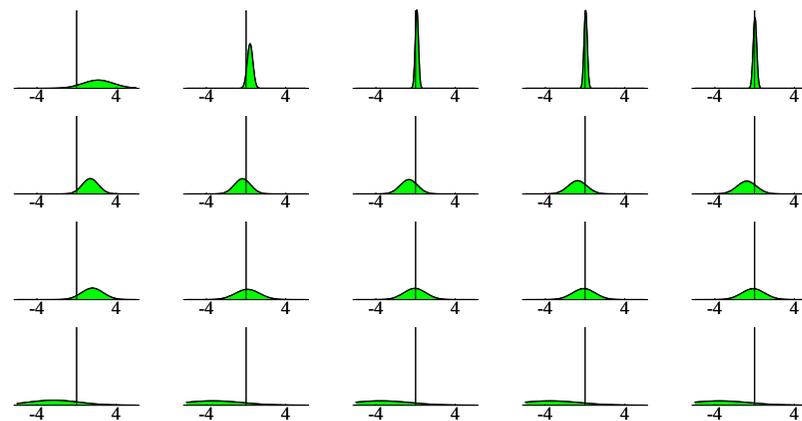
## Bond distances I

- Mean and mean absolute errors for 28 bond distances (pm):



$ \bar{\Delta} $	DZ	TZ	QZ
CCSD(T)	1.68	0.20	0.16
CCSD	1.19	0.64	0.80
MP2	1.35	0.56	0.51
HF	1.94	2.63	2.74

- Bonds shorten as we increase the basis set:
  - DZ  $\rightarrow$  TZ  $\approx$  1 pm
  - TZ  $\rightarrow$  QZ  $\approx$  0.1 pm
- Bonds lengthen as we improve the  $N$ -electron model:
  - singles  $<$  doubles  $<$  triples  $<$   $\dots$



## Bond distances II

		HF	MP2	CCSD	CCSD(T)	emp.	exp.
H <sub>2</sub>	$R_{\text{HH}}$	73.4	73.6	74.2	74.2	74.1	74.1
HF	$R_{\text{FH}}$	89.7	91.7	91.3	91.6	91.7	91.7
H <sub>2</sub> O	$R_{\text{OH}}$	94.0	95.7	95.4	95.7	95.8	95.7
NH <sub>3</sub>	$R_{\text{NH}}$	99.8	100.8	100.9	101.1	101.1	101.1
N <sub>2</sub> H <sub>2</sub>	$R_{\text{NH}}$	101.1	102.6	102.5	102.8	102.9	102.9
C <sub>2</sub> H <sub>2</sub>	$R_{\text{CH}}$	105.4	106.0	106.0	106.2	106.2	106.2
C <sub>2</sub> H <sub>4</sub>	$R_{\text{CH}}$	107.4	107.8	107.9	108.1	108.1	108.1
N <sub>2</sub>	$R_{\text{NN}}$	106.6	110.8	109.1	109.8	109.8	109.8
CH <sub>2</sub> O	$R_{\text{CH}}$	109.3	109.8	109.9	110.1	110.1	110.1
CH <sub>2</sub>	$R_{\text{CH}}$	109.5	110.1	110.5	110.7	110.6	110.7
CO	$R_{\text{CO}}$	110.2	113.2	112.2	112.9	112.8	112.8
HCN	$R_{\text{CN}}$	112.3	116.0	114.6	115.4	115.3	115.3
CO <sub>2</sub>	$R_{\text{CO}}$	113.4	116.4	115.3	116.0	116.0	116.0
HNC	$R_{\text{CN}}$	114.4	117.0	116.2	116.9	116.9	116.9
C <sub>2</sub> H <sub>2</sub>	$R_{\text{CC}}$	117.9	120.5	119.7	120.4	120.4	120.3
CH <sub>2</sub> O	$R_{\text{CO}}$	117.6	120.6	119.7	120.4	120.5	120.3
N <sub>2</sub> H <sub>2</sub>	$R_{\text{NN}}$	120.8	124.9	123.6	124.7	124.6	124.7
C <sub>2</sub> H <sub>4</sub>	$R_{\text{CC}}$	131.3	132.6	132.5	133.1	133.1	133.1

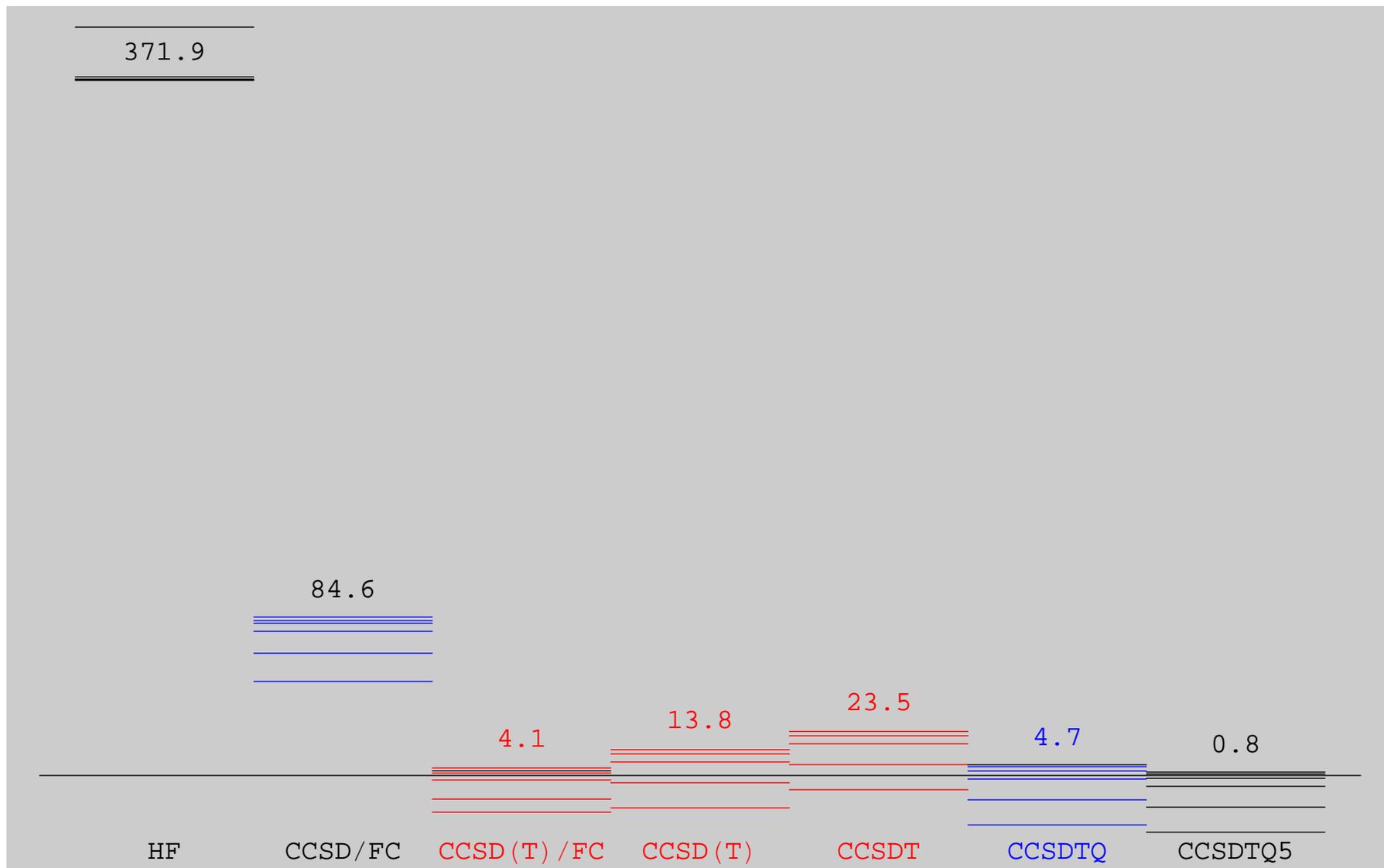
## Harmonic frequencies of some small molecules

- Harmonic frequencies vibrational frequencies ( $\text{cm}^{-1}$ ):

	HF	N <sub>2</sub>	F <sub>2</sub>	CO
CCSD	4191.0	2443.2	1026.5	2238.5
triples	-48.8	-71.9	-92.6	-72.2
quadruples	-4.1	-18.8	-15.3	-5.8
quintuples	-0.1	-3.9	-0.8	0.0
core correlation	4.0	9.8	1.6	9.9
relativistic correction	-3.5	-1.4	-0.5	-1.3
adiabatic correction	0.4	0.0	0.0	0.0
theory	4138.9	2358.0	918.9	2169.1
experiment	4138.3	2358.6	916.6	2169.8
error	0.1	-0.6	2.3	0.7

- The accurate calculation of frequencies requires very high excitation levels:
  - it does not seem possible yet to obtain harmonic frequencies to within  $1 \text{ cm}^{-1}$
  - error in F<sub>2</sub> may be explained by a too short bond distance (141.19 vs. 141.27 pm)

The accurate calculation of  $\omega_e$  in  $N_2$

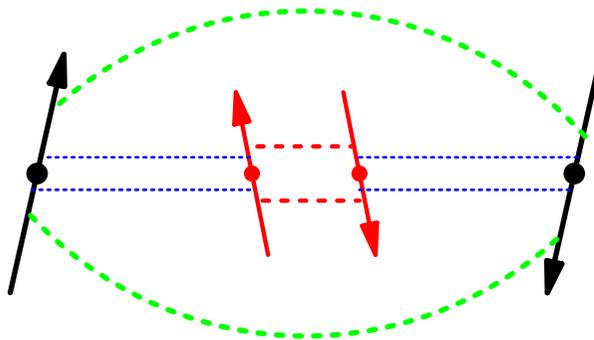


## Density-functional theory (DFT)

- In all methods discussed so far, we have calculated the electronic wave function—a very complicated function in  $4N$  coordinates  $\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ :
  - to determine the wave function, we frequently use several million parameters, making the calculation very expensive
- In principle, however, the energy is a **functional of the electron density  $E[\rho]$** :
  - it is thus possible to bypass the wave function and to concentrate instead on the electron density  $\rho(\mathbf{r})$ —a simple function of only three spatial coordinates  $\mathbf{r}$
  - this is the basis of **density-functional theory (DFT)**
- A problem with DFT is that the explicit form of  $E[\rho]$  is unknown:
  - a number of quite accurate **approximate density functionals** have been developed
  - unfortunately, these functionals have been arrived at in a rather ad hoc manner
  - there are no hierarchies of approximation
  - DFT calculations cannot be improved systematically towards the exact solution:  
**what you see is what you got**
- Nevertheless, over the last 10 years, DFT has been immensely successful:
  - it is relatively expensive (about the same cost as Hartree–Fock theory)
  - it can more easily be applied to large systems (several hundred atoms)

## NMR indirect nuclear spin–spin coupling constants

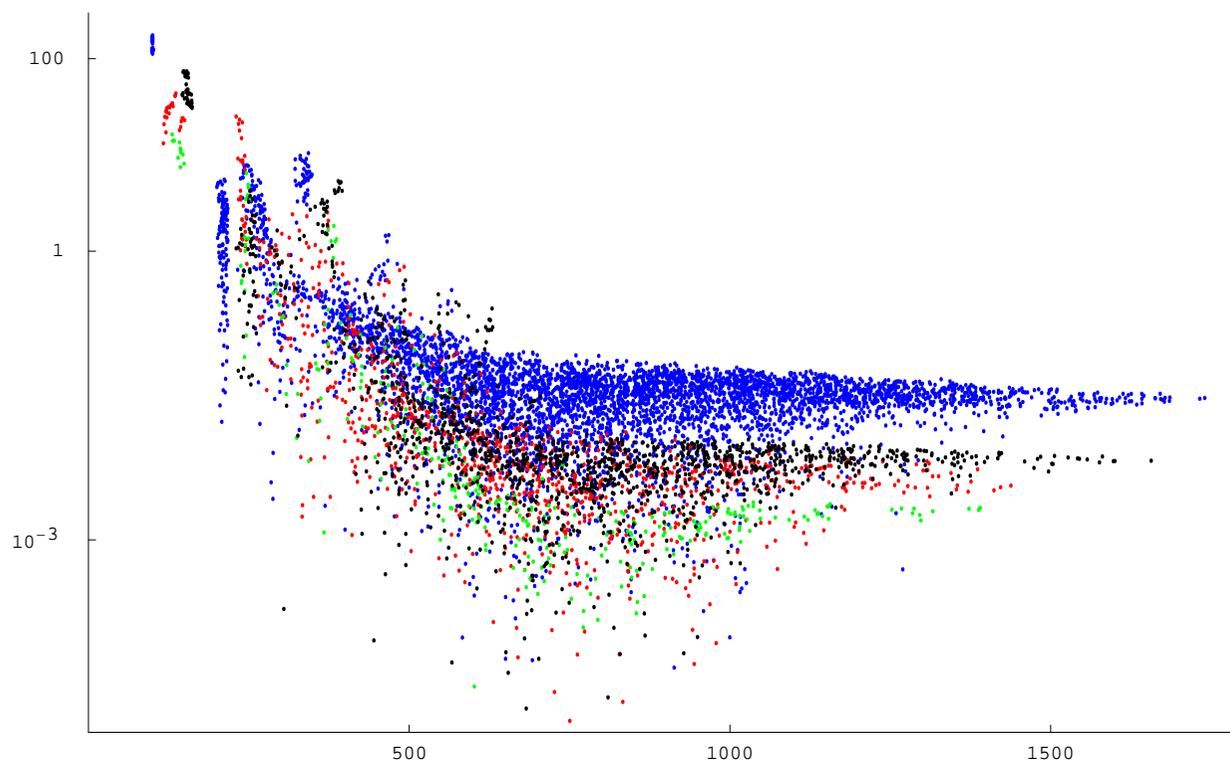
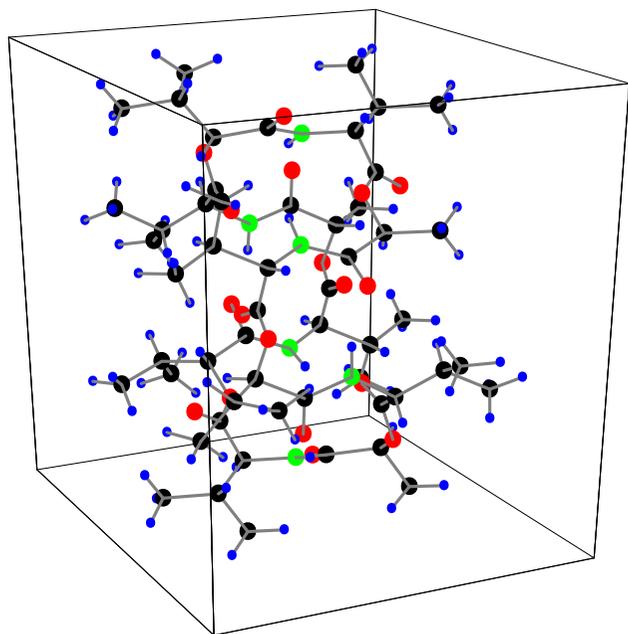
- As an example of the application of DFT, we shall consider the calculation of NMR indirect nuclear spin–spin coupling constants.
- The indirect nuclear spin–spin coupling tensors represent the coupling of the nuclear magnetic moments, as modified by the **intervening electrons**:



- The observed couplings are mediated by **hyperfine interactions** and are exceedingly weak.
- These constants determine the splitting of lines in NMR spectra.
- They are the second derivatives of the energy with respect to nuclear magnetic moments.
- The calculations are difficult for the following reasons:
  - its accurate calculation requires a flexible description of the wave function, in particular close to the nuclei
  - a large number of mechanisms contribute to the coupling (spin and orbital motion)

## Indirect nuclear spin–spin coupling constants in valinomycin

- We have calculated the indirect nuclear spin-spin coupling constants in valinomycin.
- To the right, we have plotted the absolute values of the 7587 spin–spin coupling constants (Hz) to all carbon atoms on a logarithmic scale, as the function of the internuclear separation (pm).



## Summary

- Quantum chemistry has evolved into an important tool in many areas of chemistry.
- We have reviewed the **Born–Oppenheimer approximation**:
  - the separation of nuclear and electronic degrees of freedom
- We have considered **wave-function methods**:
  - the independent-particle Hartree–Fock model
  - the coupled-cluster model of electron correlation
  - one-electron basis sets of atom-fixed Gaussian functions
  - one- and  $N$ -electron hierarchies
- We have briefly considered **density-functional theory**:
  - surprisingly accurate in many cases (often comparable to CCSD)
  - inexpensive and applicable to large systems
  - cannot be improved upon in a systematic manner
  - often contains empirical elements