

# Highly accurate quantum-chemical calculations

T. Helgaker

Centre for Theoretical and Computational Chemistry,

Department of Chemistry, University of Oslo, Norway

A. C. Hennum and T. Ruden, University of Oslo, Norway

S. Coriani, University of Trieste, Italy

W. Klopper, University of Karlsruhe, Germany

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

The Eighth Informal Conference on Atmospheric and Molecular Science

LO-Skolen Conference Center, Helsingør, Denmark

June 8–10, 2007

## High-precision quantum-chemical calculations

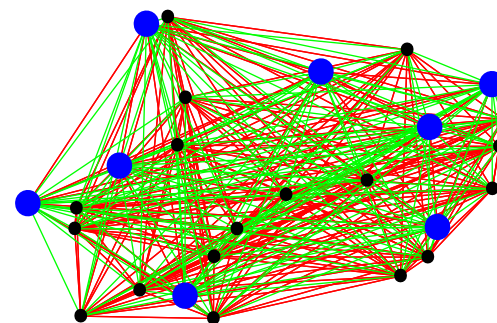
- The last decade has seen a dramatic improvement in our ability to treat molecular electronic systems accurately
  - development of techniques for systematic convergence towards the exact solution
  - extensive benchmarking on small and light molecular systems
- In many cases, we can now confidently confirm or reject experimental observations
- Many black-box methods have been developed
  - well-defined levels of theory
  - relatively easy to use by the nonspecialist
- The purpose of the present talk is to give an introduction to such methods
  - the emphasis is on high-accuracy methods
  - the explicit construction of a wave function
  - do not touch density-functional theory

## The many-body problem of quantum chemistry

- All information about the electronic system is contained in the **wave function**, obtained by solving the **Schrödinger equation**.

“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.”

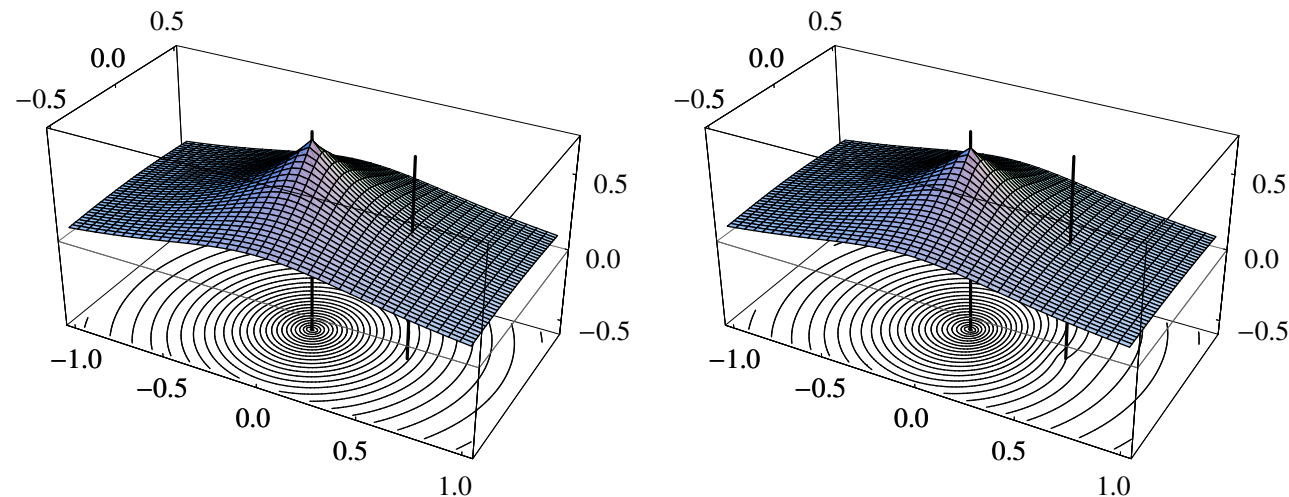
P. A. M. Dirac, 1929



- We are thus forced to make **approximations**—that is, set up simplified descriptions that incorporate the most important features of the system.
- This should preferably be done in an orderly fashion, so that the exact solution can be approached in a systematic manner.
- In this manner, we establish **hierarchies of approximations**—that is, a system of ever more accurate and expensive computational models.

## The Hartree–Fock approximation

- The Hartree–Fock model—the fundamental approximation of wave-function theory
  - each electron moves in the mean field of all other electrons
  - provides an uncorrelated description: average rather than instantaneous interactions
  - gives rise to the concept of molecular orbitals
  - typical errors: 0.5% in the energy; 1% in bond distances, 5%–10% in other properties
  - forms the basis for more accurate treatments
- The Hartree–Fock and exact wave functions in helium:



- concentric Hartree–Fock contours, reflecting an uncorrelated description
- in reality, the electrons see each other and the contours become distorted

## Electron correlation and virtual excitations

- **electron correlation:**
  - to improve upon the Hartree–Fock 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**:

$$\hat{X}_{ij}^{ab} = t_{ij}^{ab} a_b^\dagger a_a^\dagger a_i a_j$$

- 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 Hartree–Fock state, we obtain an improved, **correlated description** of the electrons:

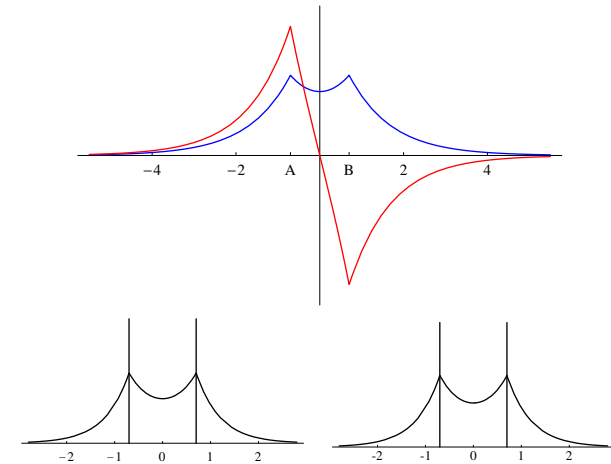
$$|\text{HF}\rangle \rightarrow (1 + \hat{X}_{ij}^{ab})|\text{HF}\rangle$$

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

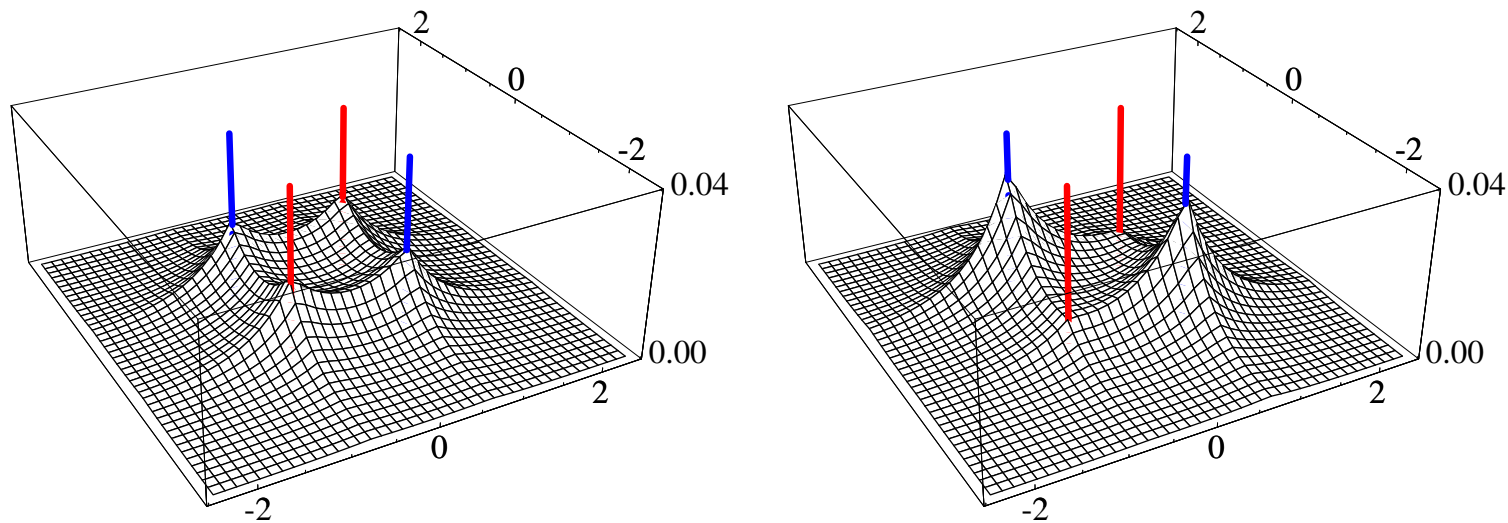
- Consider the effect of a double excitation in H<sub>2</sub>:

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

- 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 excitation operators

$$|\text{CC}\rangle = \underbrace{\left(1 + \hat{X}_i^a\right)}_{\text{singles}} \cdots \underbrace{\left(1 + \hat{X}_{ij}^{ab}\right)}_{\text{doubles}} \cdots \underbrace{\left(1 + \hat{X}_{ijk}^{abc}\right)}_{\text{triples}} \cdots \underbrace{\left(1 + \hat{X}_{ijkl}^{abcd}\right)}_{\text{quadruples}} \cdots |\text{HF}\rangle$$

- with each excitation, there is an associated **probability amplitude**  $t_{ijk\dots}^{abc\dots}$
- **single excitations** represent orbital adjustments rather than interactions
- **double excitations** are particularly important, arising from pair interactions
- **higher excitations** should become progressively less important
- This classification provides a **hierarchy of ‘truncated’ CC wave functions**:
  - CCS, CCSD, CCSDT, CCSDTQ, CCSDTQ5, ...
  - errors are typically reduced by a factor of three to four at each new level
- Lower-order excitations work in tandem to produce higher-order excited configurations

$$\left(1 + \hat{X}_{ij}^{ab}\right) \left(1 + \hat{X}_{kl}^{cd}\right) |\text{HF}\rangle = |\text{HF}\rangle + \hat{X}_{ij}^{ab} |\text{HF}\rangle + \hat{X}_{kl}^{cd} |\text{HF}\rangle + \hat{X}_{ij}^{ab} \hat{X}_{kl}^{cd} |\text{HF}\rangle$$

- the important thing is to parameterize the excitations rather than the resulting states

## Basis sets of Gaussian functions

- In our calculations, we expand the molecular orbitals in one-electron Gaussian-type functions (GTOs):

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

- 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 number of GTOs per atom increases rapidly:

SZ	DZ	TZ	QZ	5Z	6Z
5	14	30	55	91	140

- Large basis sets are needed to generate a flexible virtual space!



## The two-dimensional chart of nonrelativistic quantum chemistry

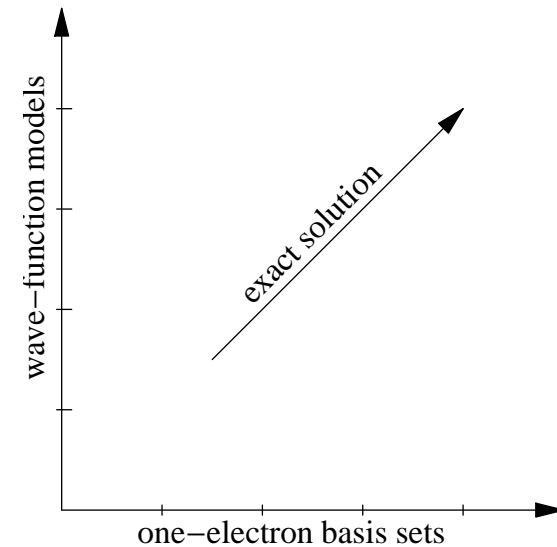
- The quality of nonrelativistic molecular electronic-structure calculations is determined by the description of
  1. the  $N$ -electron space (wave-function model),
  2. the one-electron space (basis set).
- In each space, there is a hierarchy of levels of increasing complexity:

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

HF, CCSD, CCSDT, CCSDTQ, ...

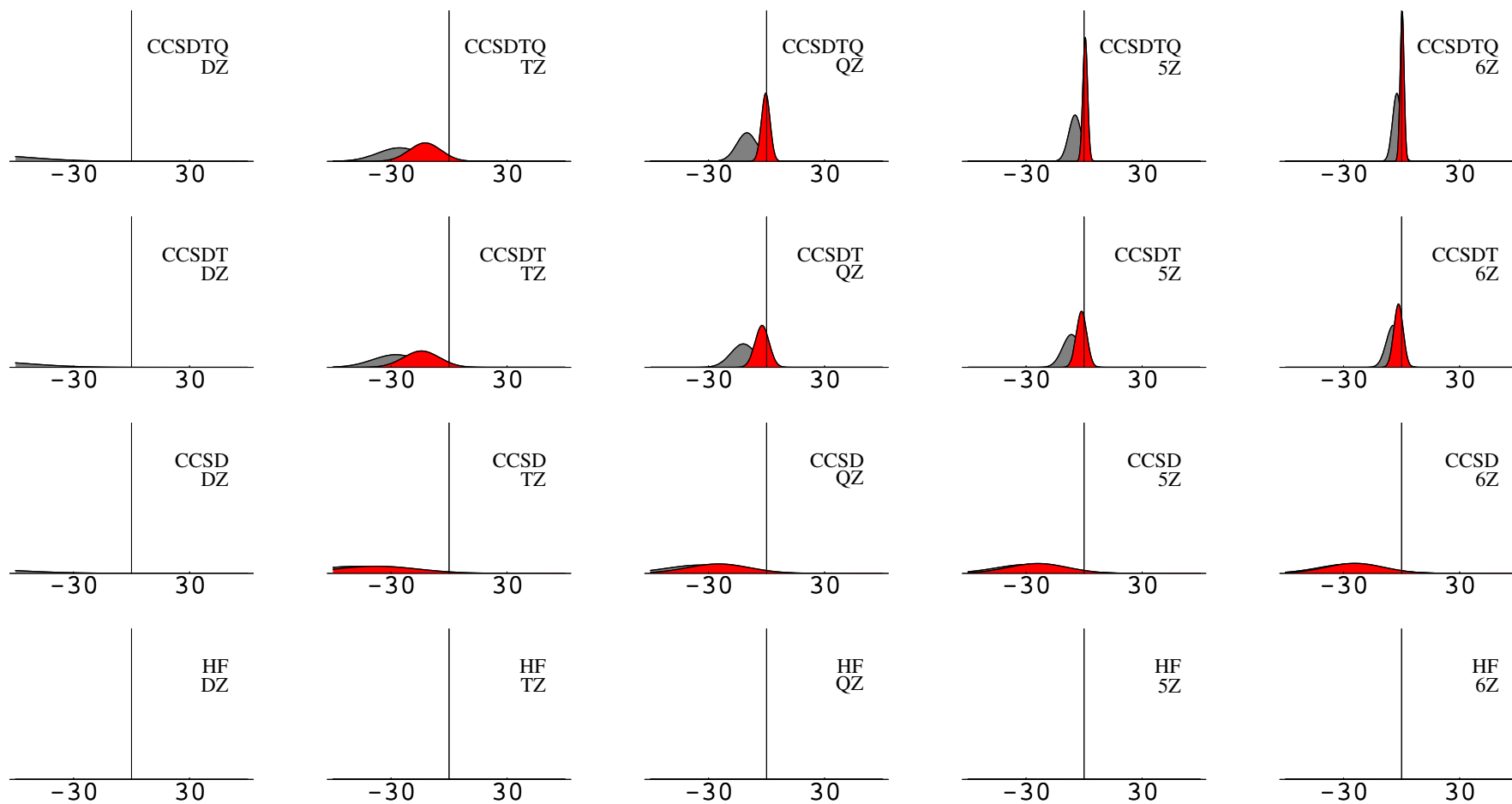
2. the one-electron hierarchy:  
correlation-consistent basis sets

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



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

# Atomization energies (kJ/mol)



## Basis-set convergence

- We have a systematic approach towards the exact solution of the Schrödinger equation
- However, basis-set convergence is painfully slow:

	$N_{\text{bas}}$	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.
- What is the reason for this slow convergence?
  - can it be accelerated?
- To understand this problem, it is sufficient to study the helium atom.
  - this was the first many-body problem studied quantum-mechanically

## The local kinetic energy

- Consider the local energy of the helium atom

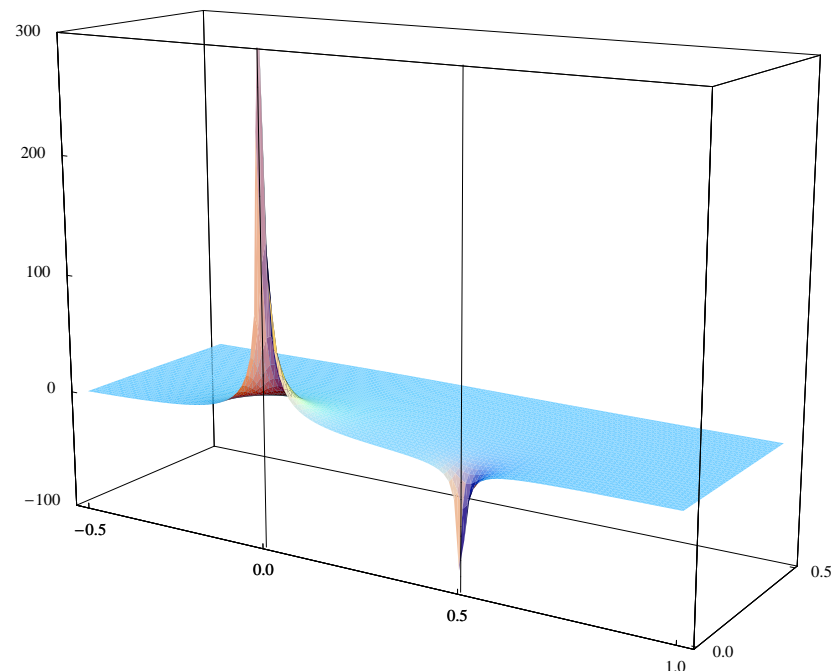
$$E_{\text{loc}} = (H\Psi)/\Psi \quad \leftarrow \text{constant for exact wave function}$$

- The electronic Hamiltonian has singularities at points of coalescence

$$H = -\frac{1}{2}\nabla_1^2 - \frac{1}{2}\nabla_2^2 - \frac{2}{r_1} - \frac{2}{r_2} + \frac{1}{r_{12}}$$

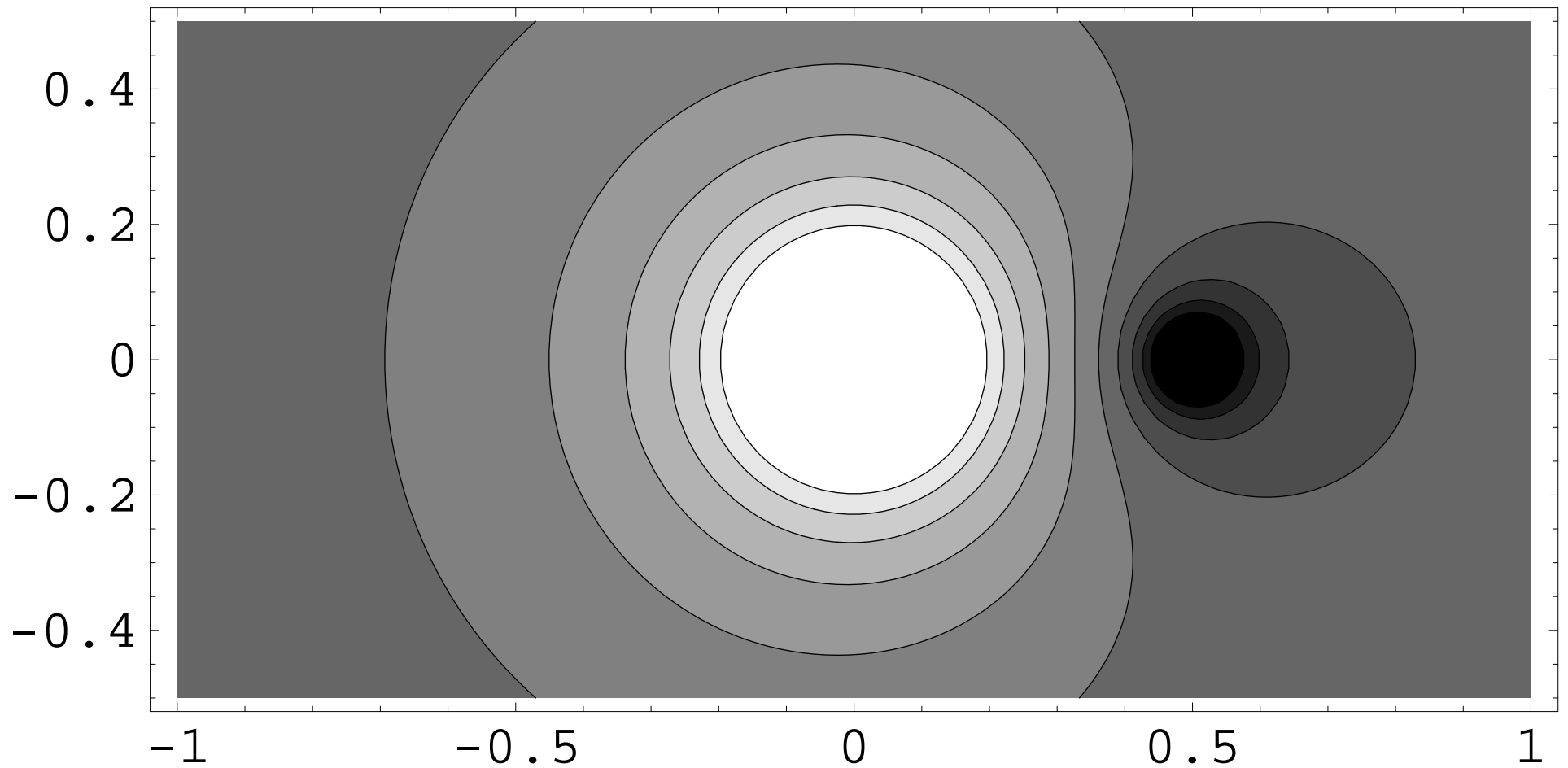
– the infinite potential terms must be canceled by infinite kinetic terms at coalescence

- Local kinetic energy in the helium atom
  - positive around the nucleus
  - negative around the second electron
- Negative kinetic energy counterintuitive
  - classical forbidden region
  - internal “tunneling”
  - w. f. decays towards the singularity



## The Coulomb hole: the forbidden region

- Each electron is surrounded by a classically forbidden region: the Coulomb hole
  - without a good description of this region, our results will be inaccurate

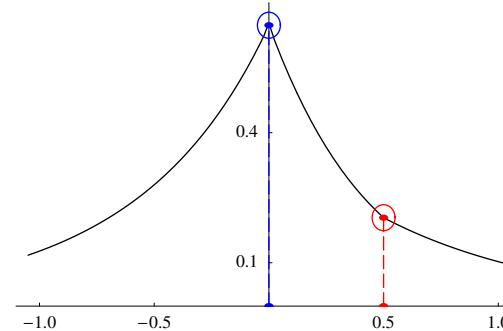


## The electron cusp and the Coulomb hole

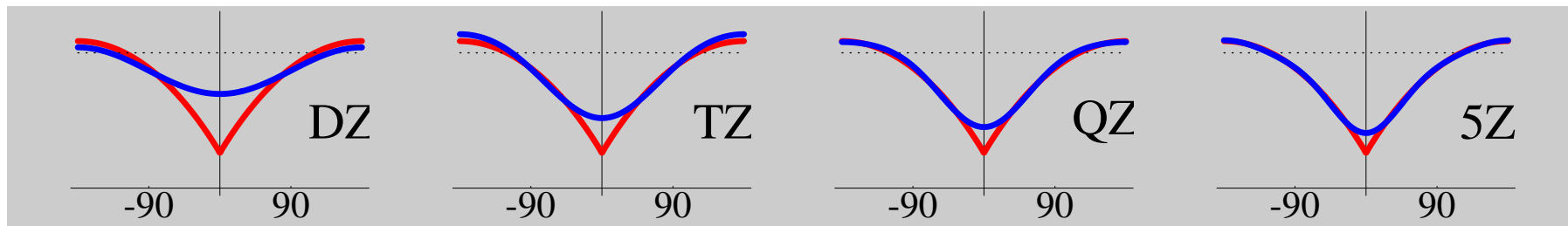
- The behaviour of the wave function at coalescence was established by Slater (1928)
  - nuclear and electronic cusp conditions

$$\left(\frac{\partial\Psi}{\partial r_i}\right)_{r_i=0,\text{ave}} = -Z\Psi(r_i=0)$$

$$\left(\frac{\partial\Psi}{\partial r_{12}}\right)_{r_{12}=0,\text{ave}} = \frac{1}{2}\Psi(r_{12}=0)$$



- The electronic cusp condition implies the existence of a **Coulomb hole**
  - its description is awkward without the use of  $r_{ij}$  in the wave function:



- standard wave functions use orbitals and do not contain  $r_{ij}$  explicitly

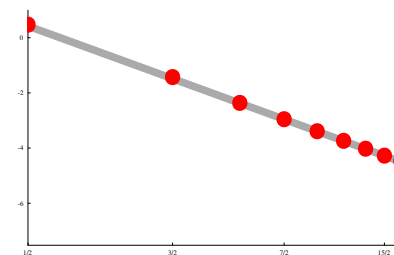
# The principal expansion and correlation-consistent basis sets

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

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

- 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
	+3s3p3d	+4s4p4d4f	+5s5p5d5f5g	$\propto X^2$
2s1p	3s2p1d	4s3p2d1f	5s4p3d2f1g	$\propto X^3$

- 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!

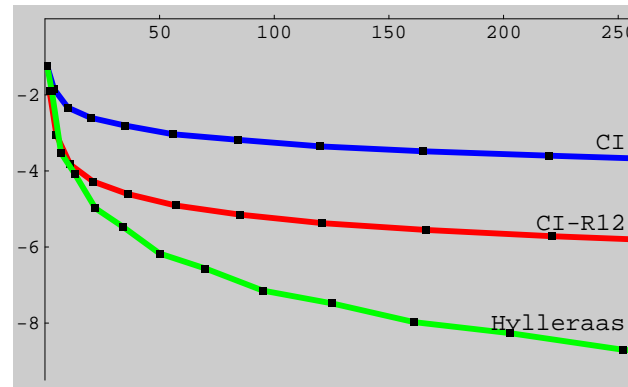
$$1 \text{ minute} \quad \rightarrow \quad 1 \text{ week} \quad \rightarrow \quad 200 \text{ years}$$

## Solutions to slow basis-set convergence

### 1. Use explicitly correlated methods!

- Include interelectronic distances  $r_{ij}$  in the wave function:

$$\Psi_{R12} = \sum_K C_K \Phi_K + C_{R12} r_{12} \Phi_0$$



### 2. Use basis-set extrapolation!

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

$$E_\infty = \frac{X^3 E_X - Y^3 E_Y}{X^3 - Y^3}$$

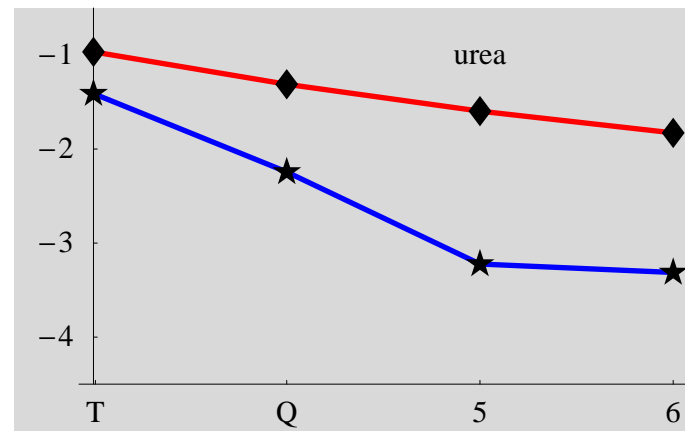
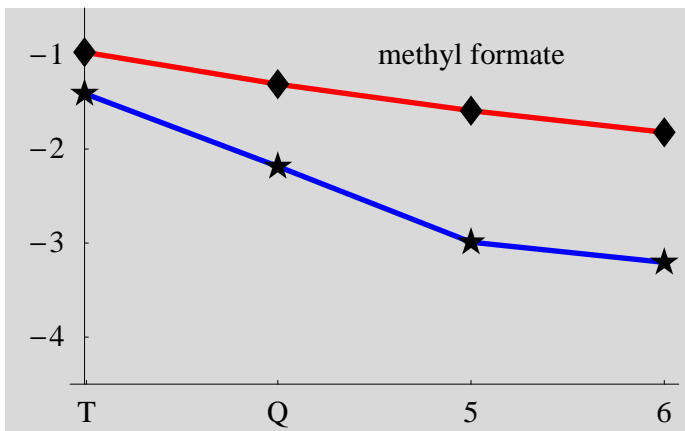
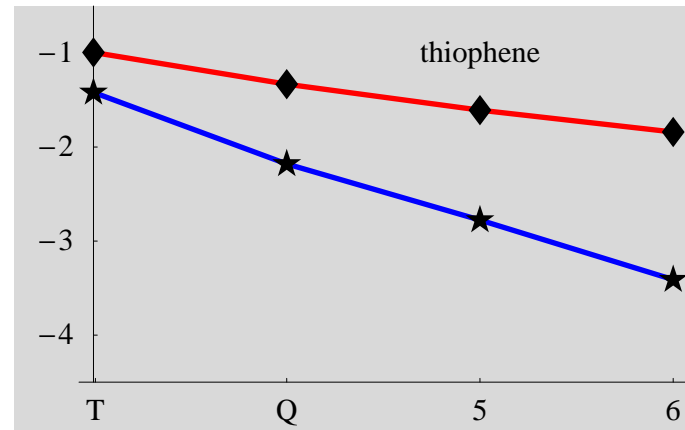
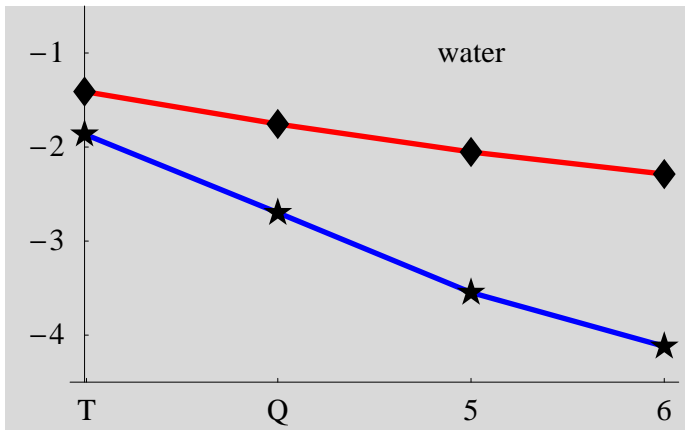
$mE_h$	DZ	TZ	QZ	5Z	6Z	R12
plain	194.8	62.2	23.1	10.6	6.6	1.4
extr.		21.4	1.4	0.4	0.5	

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



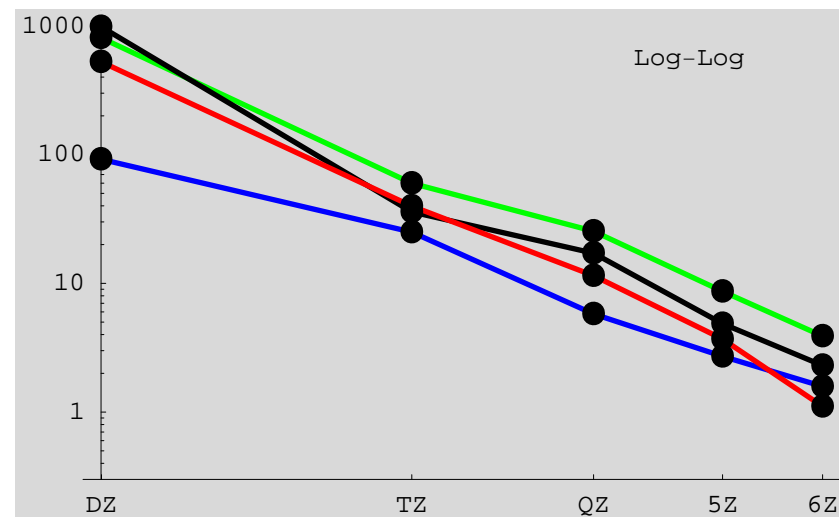
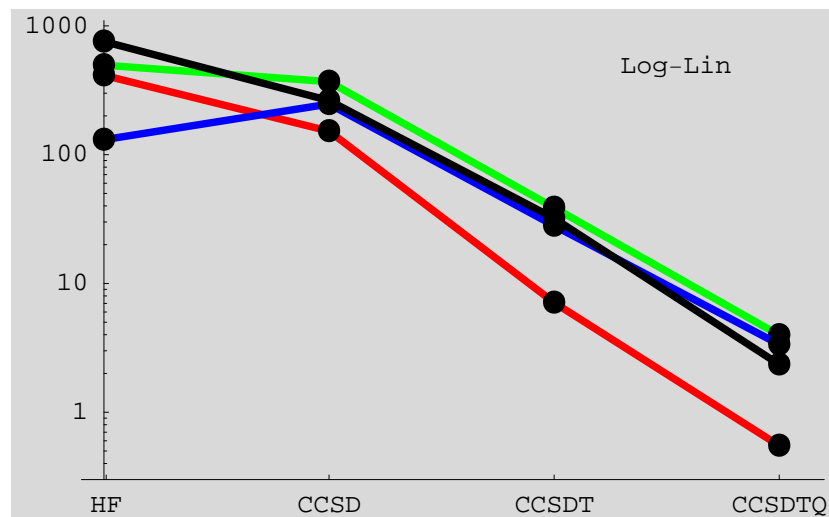
## Extrapolation vs. explicit correlation

- For some purposes, extrapolation is competitive with explicit correlation
  - logarithmic errors in **plain** and **extrapolated** energies relative to R12



## Review and preview

- We have studied the two-dimensional chart of quantum chemistry:
  - excitation level ( $N$ -electron space) and basis sets (one-electron space)
  - example: atomization energies of **HF**, **N<sub>2</sub>**, **F<sub>2</sub>**, and **CO**
  - contributions of each excitation level (left) and each basis-set shell (right)



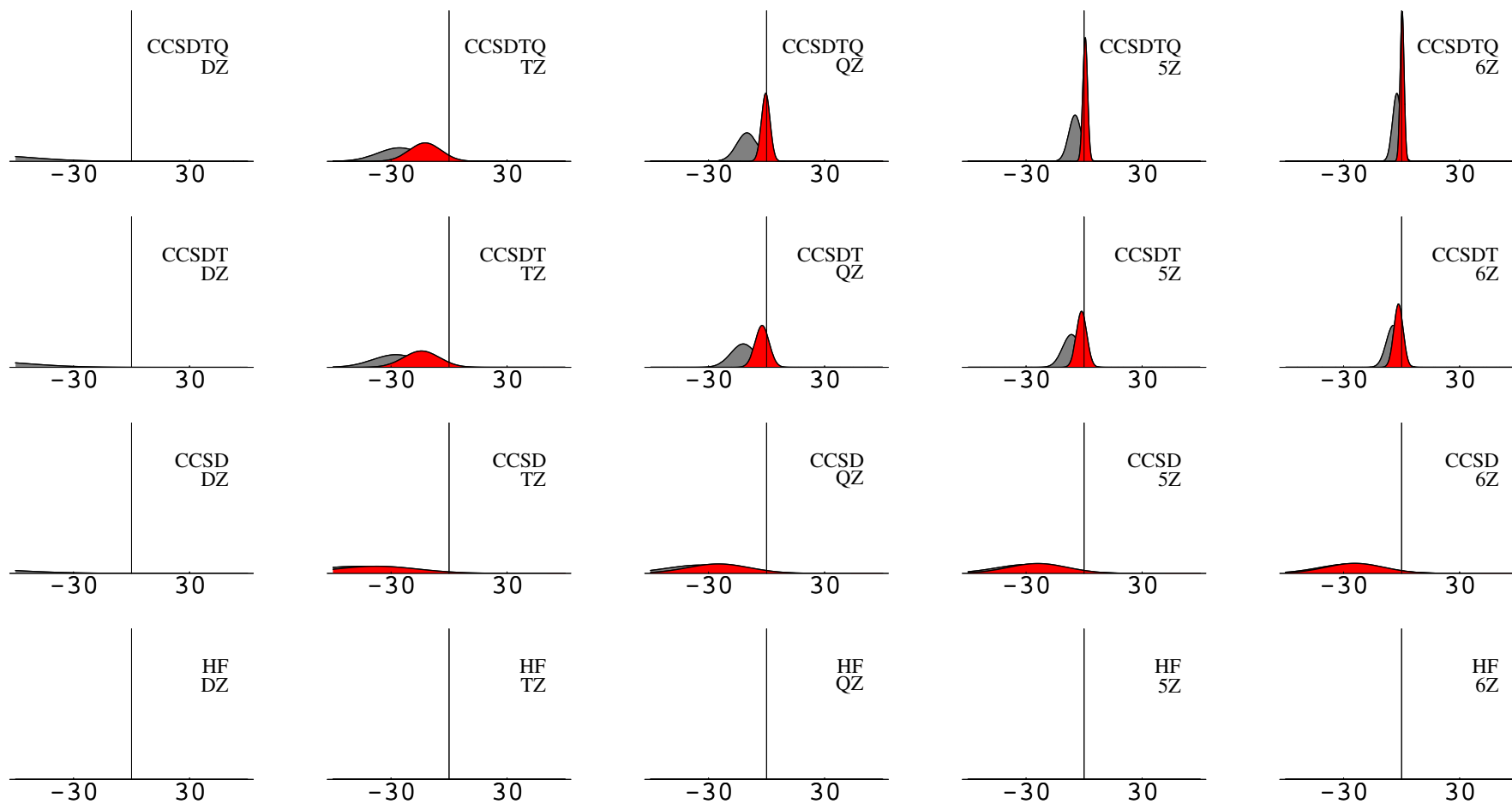
- convergence is smooth but slow in both directions
- We shall now consider in detail three properties:
  1. atomization energies
  2. bond distances
  3. harmonic and anharmonic vibrational constants

## Atomization energies (kJ/mol)

	RHF	SD	T	Q	rel.	vib.	total	experiment	error
CH <sub>2</sub>	531.1	218.3	9.5	0.4	-0.7	-43.2	715.4	714.8±1.8	0.6
H <sub>2</sub> O	652.3	305.3	17.3	0.8	-2.1	-55.4	918.2	917.8±0.2	0.4
HF	405.7	178.2	9.1	0.6	-2.5	-24.5	566.7	566.2±0.7	0.5
N <sub>2</sub>	482.9	426.0	42.4	3.9	-0.6	-14.1	940.6	941.6±0.2	-1.1
F <sub>2</sub>	-155.3	283.3	31.6	3.3	-3.3	-5.5	154.1	154.6±0.6	-0.5
CO	730.1	322.2	32.1	2.3	-2.0	-12.9	1071.8	1071.8±0.5	-0.0

- agreement with experiment in all cases except for N<sub>2</sub>
- all purely electronic contributions are positive (except Hartree–Fock for F<sub>2</sub>)
  - very large correlation contributions
- relativistic corrections are small ( $\approx 0.5\%$ ) but needed for agreement with experiment
  - nearly cancel quadruples contributions (error cancellation possible)
- vibrational corrections are substantial
  - similar to triples contributions but oppositely directed (error cancellation)

# Atomization energies (kJ/mol)

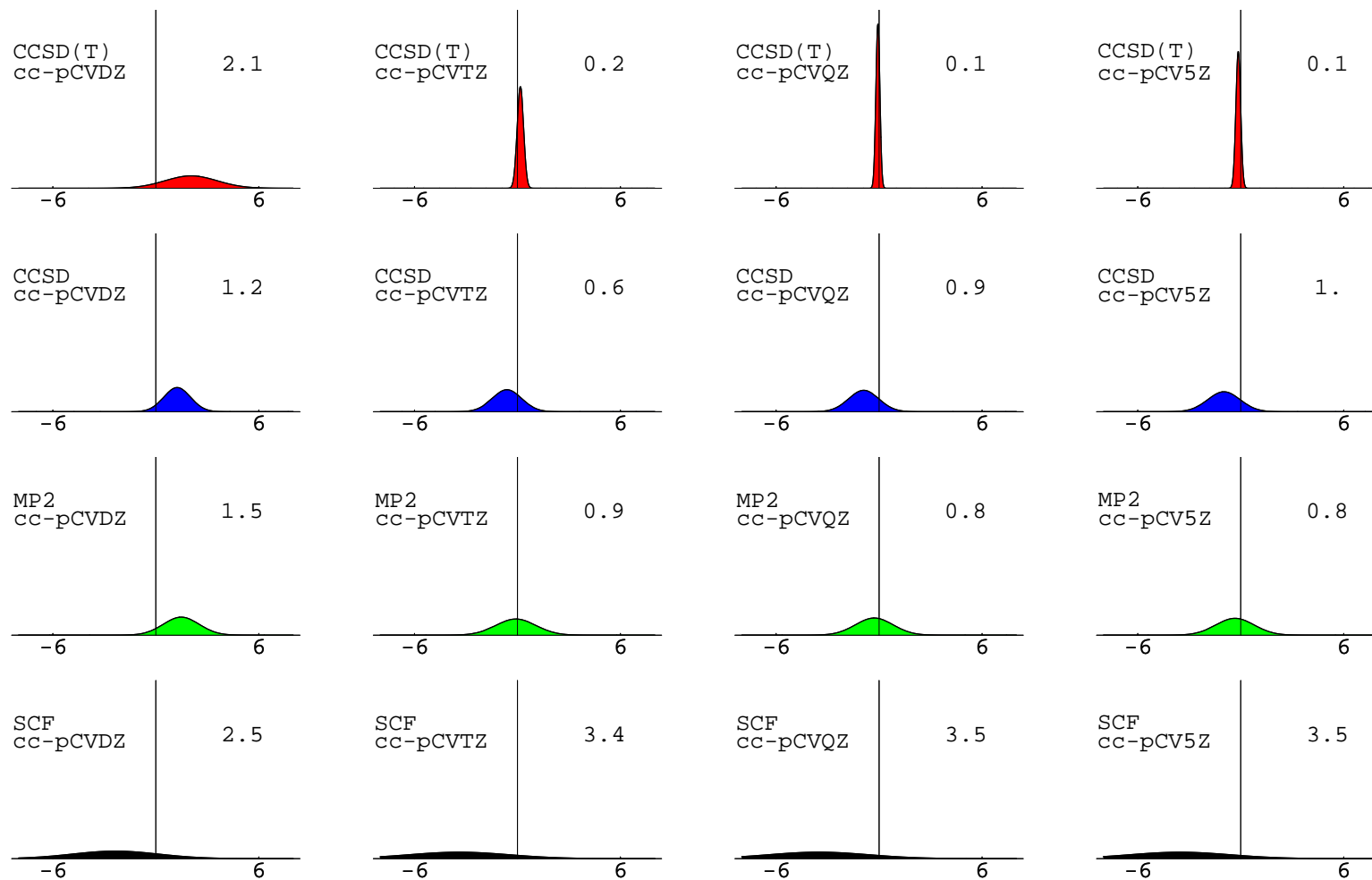


## Bond distances (pm)

	RHF	SD	T	Q	5	rel.	theory	exp.	err.
HF	89.70	1.67	0.29	0.02	0.00	0.01	91.69	91.69	0.00
N <sub>2</sub>	106.54	2.40	0.67	0.14	0.03	0.00	109.78	109.77	0.01
F <sub>2</sub>	132.64	6.04	2.02	0.44	0.03	0.05	141.22	141.27	-0.05
CO	110.18	1.87	0.75	0.04	0.00	0.00	112.84	112.84	0.00

- agreement with experiment to within 0.01 pm except for F<sub>2</sub>
- Hartree–Fock theory underestimates bond distances by up to 8.6 pm (for F<sub>2</sub>)
- all correlation contributions are positive
  - approximate linear convergence, slowest for F<sub>2</sub>
  - triples contribute up to 2.0 pm, quadruples up to 0.4 pm, and quintuples 0.03 pm
- relativistic corrections are small except for F<sub>2</sub> (0.05 pm)
  - of the same magnitude and direction as the quintuples
  - no sources of error cancellation (in a complete AO basis)

# Bond distances $R_e$ of BH, CO, N<sub>2</sub>, HF, and F<sub>2</sub> (cm<sup>-1</sup>)

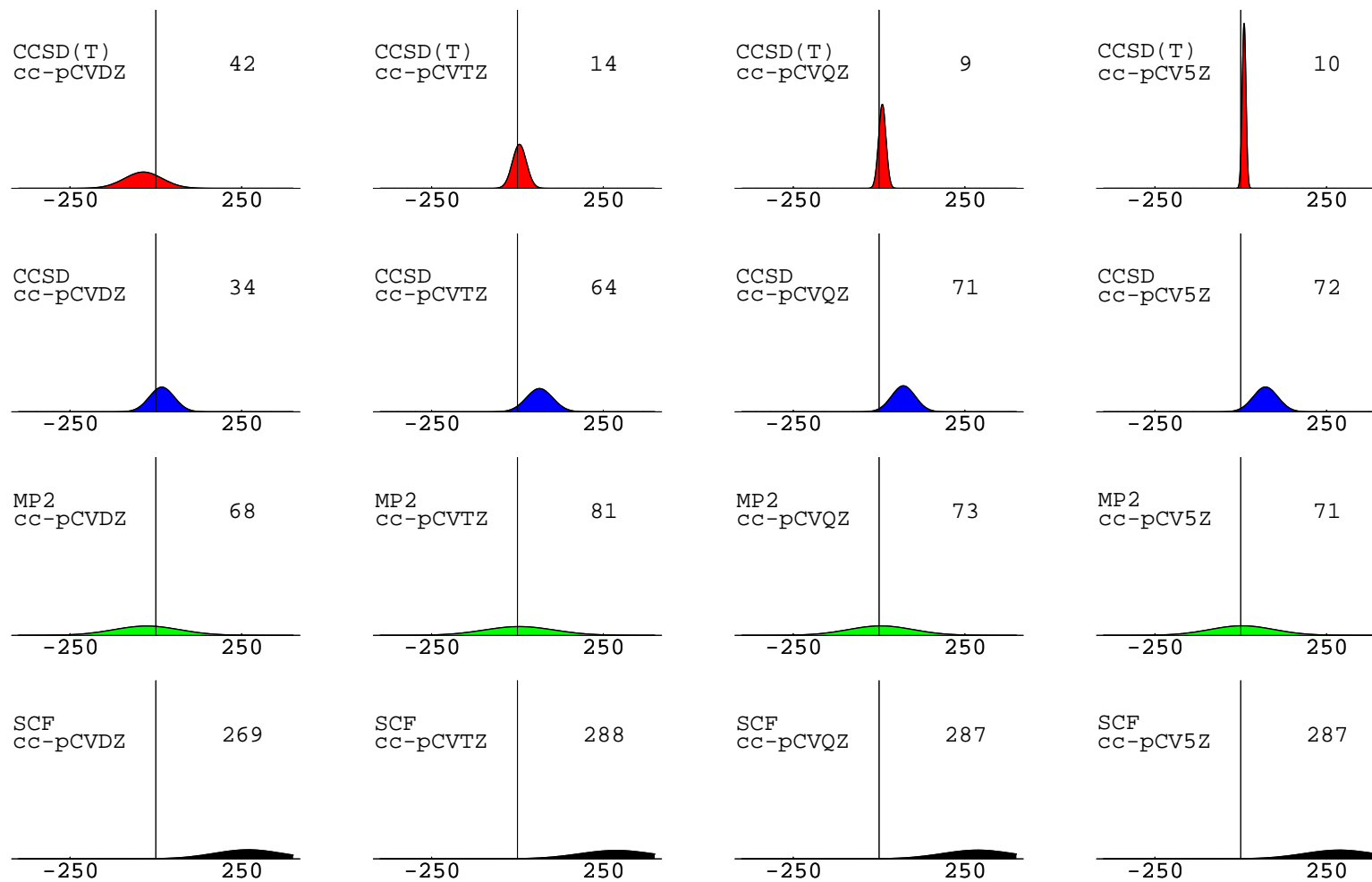


## Harmonic constants $\omega_e$ ( $\text{cm}^{-1}$ )

	RHF	SD	T	Q	5	rel.	theory	exp.	err.
HF	4473.8	-277.4	-50.2	-4.1	-0.1	-3.5	4138.5	4138.3	0.2
N <sub>2</sub>	2730.3	-275.8	-72.4	-18.8	-3.9	-1.4	2358.0	2358.6	-0.6
F <sub>2</sub>	1266.9	-236.1	-95.3	-15.3	-0.8	-0.5	918.9	916.6	2.3
CO	2426.7	-177.4	-71.7	-7.2	0.0	-1.3	2169.1	2169.8	0.7

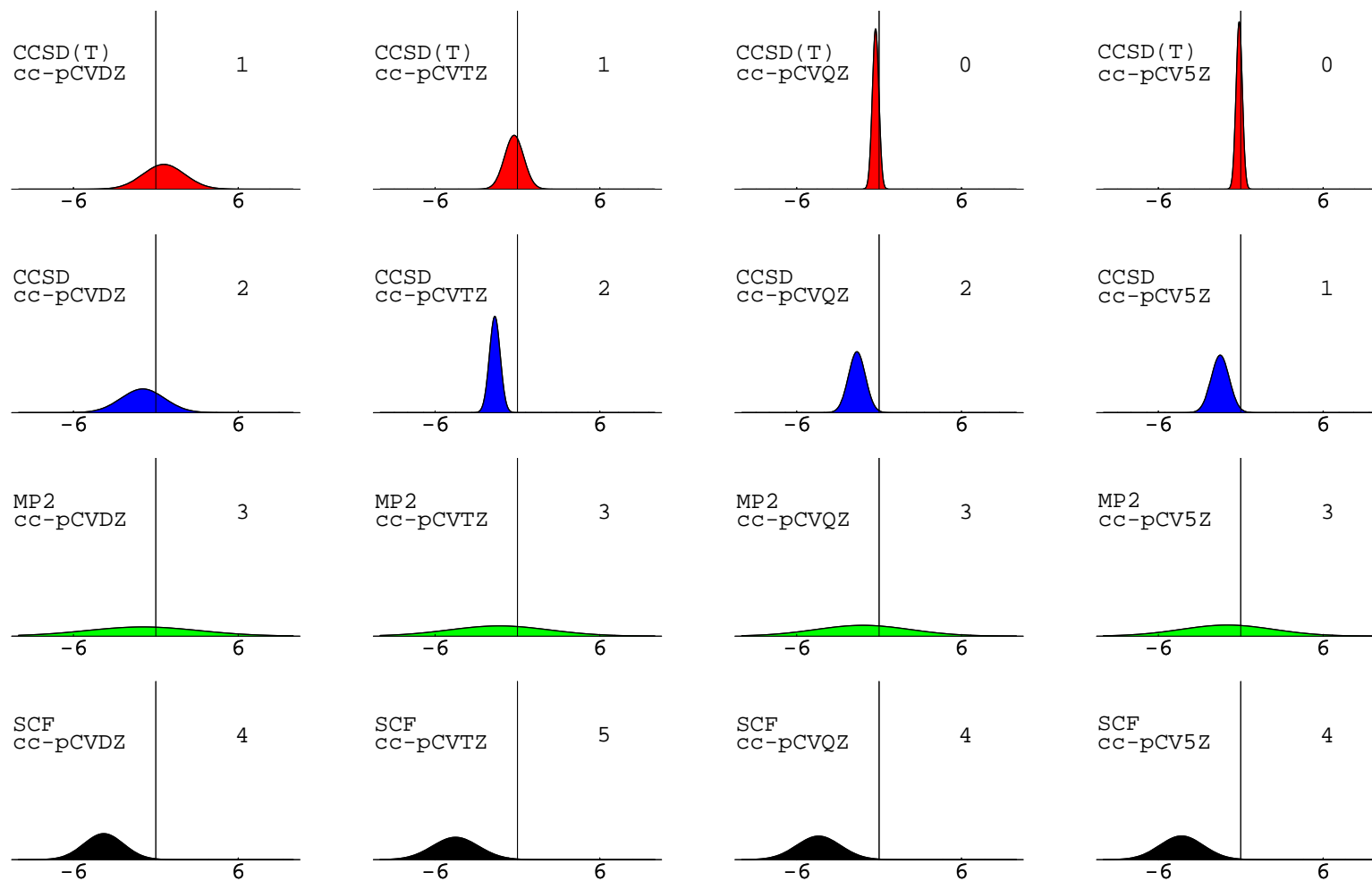
- agreement with experiment to within  $1 \text{ cm}^{-1}$  except for F<sub>2</sub>
- Hartree–Fock theory overestimates harmonic frequencies by up to 38% (in F<sub>2</sub>).
- all correlation contributions are large and negative
  - triples contribute up to  $95 \text{ cm}^{-1}$ , quadruples  $20 \text{ cm}^{-1}$ , and quintuples  $4 \text{ cm}^{-1}$
  - sextuples are sometimes needed for convergence to within  $1 \text{ cm}^{-1}$
- relativistic corrections are of the order of  $1 \text{ cm}^{-1}$ 
  - of the same magnitude and direction as the quadruples or quintuples
- no sources of error cancellation in a complete AO basis

# Harmonic constants $\omega_e$ of BH, CO, N<sub>2</sub>, HF, and F<sub>2</sub> (cm<sup>-1</sup>)

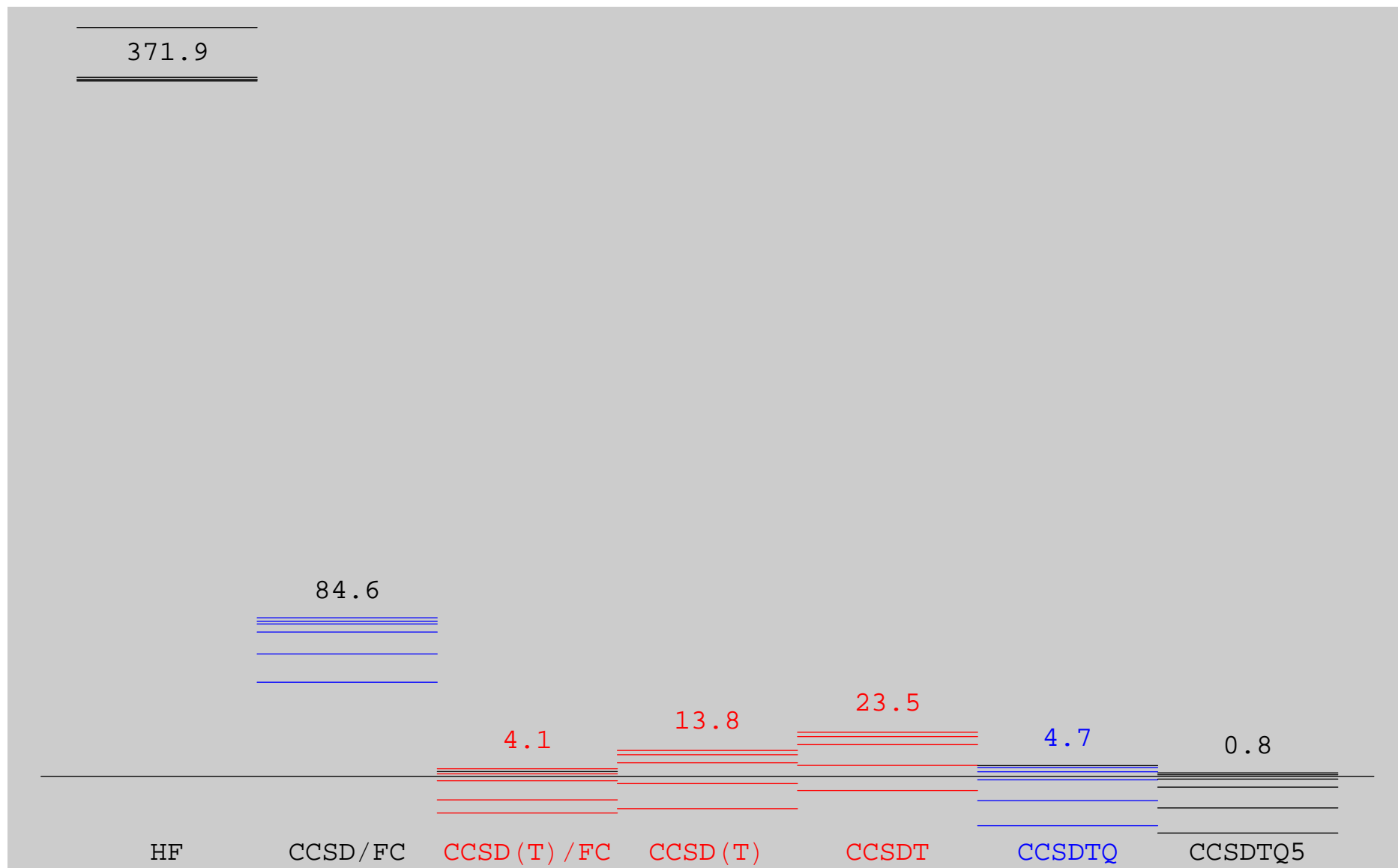




# Anharmonic constants $\omega_e x_e$ of BH, CO, N<sub>2</sub>, HF, and F<sub>2</sub> (cm<sup>-1</sup>)



# Calculations of $\omega_e$ in $N_2$



## Conclusions

- Quantum chemistry provides a set of well-defined levels of approximation
  - excitation-level expansion
  - basis-set expansion
- The “exact” result can be approached in a systematic manner
- In practice, we have to be content with low levels of theory
  - low excitation levels, approximated by perturbation theory
  - small basis sets
- Sometimes complicated interplay between different approximations
  - useful and reliable (balanced) levels of theory have been implemented in codes
  - can be usefully applied without deep knowledge of quantum chemistry
- Still, some knowledge of these methods, their errors and the interplay of these errors, will make the application of these methods more reliable
- Discipline is always needed!