

High-Accuracy Quantum Chemistry

Trygve Helgaker

Hylleraas Centre for Quantum Molecular Sciences,
Department of Chemistry, University of Oslo, Norway

4th Winter School of Computational Chemistry Program (online),
Sharif University of Technology, Tehran, Iran
February 13, 2024

High-precision quantum-chemical calculations

- ▶ Recent 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
- ▶ We can now work with confidence
 - ▶ predict, confirm or reject experimental observations
- ▶ Many black-box methods have been developed
 - ▶ well-defined levels of theory
 - ▶ relatively easy to use by the nonspecialist
- ▶ Still, the exact solution can be approached in infinitely many ways
 - ▶ at many (incomplete) levels of theory, agreement with experiment may be obtained fortuitously
 - ▶ error cancellation is treacherous: the right answer for the wrong reason
- ▶ We here review nonrelativistic high-precision quantum chemistry
 - ▶ 'Molecular Electronic Structure Theory' by Helgaker, Jørgensen and Olsen (Wiley, 2000)

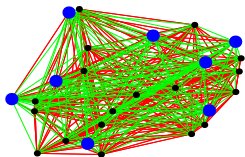
- ▶ **Many-electron problem**
 - ▶ Hartree–Fock theory
 - ▶ electron correlation and virtual excitations
 - ▶ coupled-cluster theory
 - ▶ convergence to the “exact” solution
- ▶ **Orbital expansions and basis sets**
 - ▶ electron cusp and the Coulomb hole
 - ▶ basis-set convergence
 - ▶ basis-set extrapolation
 - ▶ explicit correlation
- ▶ **Accurate calculations and calibration**
 - ▶ atomization energies
 - ▶ reaction enthalpies
 - ▶ bond distances
 - ▶ vibrational frequencies

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.”*

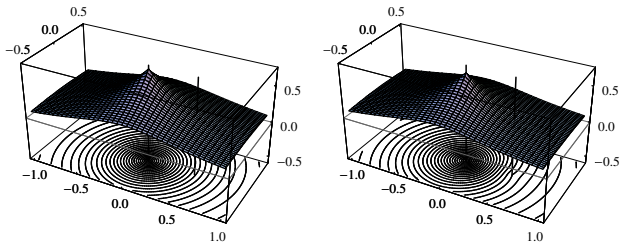
P. A. M. Dirac, 1929



- ▶ 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 the **helium atom**:



- ▶ 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:

- ▶ going beyond Hartree–Fock theory, we describe the **instantaneous interactions** among 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:

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

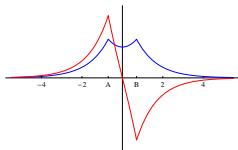
- ▶ t_{ij}^{ab} is the probability that the electrons in ϕ_i and ϕ_j will interact and be excited to ϕ_a and ϕ_b
- ▶ applying $1 + \hat{X}_{ij}^{ab}$ to the Hartree–Fock state, we obtain a correlated description of the electrons:

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

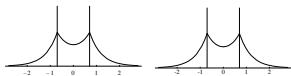
Example: electron correlation in H₂

- ▶ Consider the effect of a double excitation in H₂:

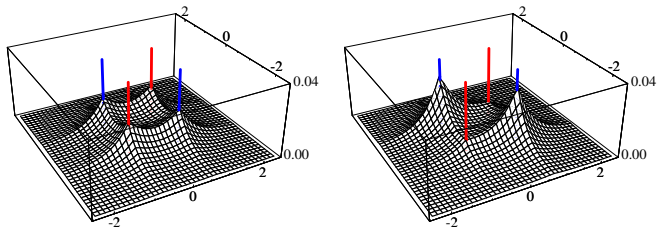
$$|1\sigma_g^2\rangle \rightarrow (1 + \hat{X}_{gg}^{uu})|1\sigma_g^2\rangle = |1\sigma_g^2\rangle - 0.11|1\sigma_u^2\rangle$$



- ▶ 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

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

- ▶ 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

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

- ▶ 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 **Gaussian-type functions (GTOs)**:

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

- ▶ Hierarchy of **basis sets** of atomic functions:

- ▶ **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 **size of the basis sets** increases rapidly:

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

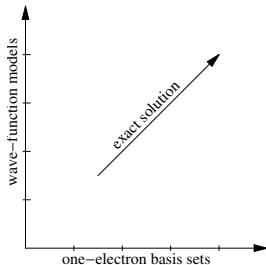
- ▶ Large basis sets are needed to generate a flexible virtual space!

The two-dimensional chart of nonrelativistic quantum chemistry

- ▶ The quality of nonrelativistic 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:

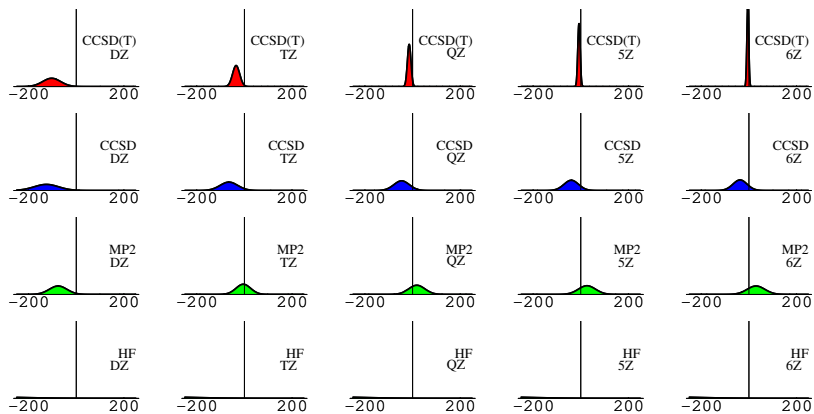
- 1 the N -electron hierarchy:
coupled-cluster excitation levels
HF, CCSD, CCSDT, CCSDTQ, ...

- 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.

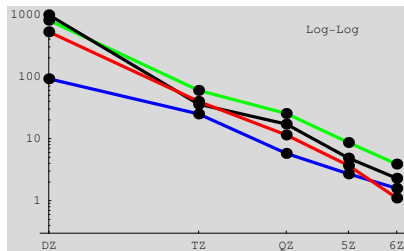
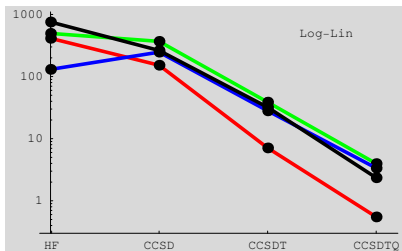
Atomization energies (kJ/mol)



Coupled-cluster (CC) vs. basis-set (AO) convergence

- ▶ Atomization energies of **HF**, **N₂**, **F₂**, and **CO**

- ▶ contributions of each CC excitation level (left) and AO basis-set shell (right)



- ▶ **Excitation-level convergence** is approximately linear (**log-linear plot**)
 - ▶ each new excitation level reduces the error by about an order of magnitude
- ▶ **Basis-set convergence** is much slower (**log-log plot**)
 - ▶ each shell contributes an energy proportional to X^{-4} where X is the cardinal number
- ▶ Convergence is **smooth** in both directions!
 - ▶ we shall consider the CC convergence first, the AO convergence next

CC convergence: atomization energies (kJ/mol)

- ▶ The energy required to split a molecule into its constituent atoms
 - ▶ complete basis-set limit results

	RHF	SD	T	Q	rel.	vib.	total	experiment	error
CH ₂	531.1	218.3	9.5	0.4	-0.7	-43.2	715.4	714.8±1.8	0.6
H ₂ 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 ₂	482.9	426.0	42.4	3.9	-0.6	-14.1	940.6	941.6±0.2	-1.1
F ₂	-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 N₂
- ▶ All purely **electronic contributions** are positive (except Hartree–Fock for F₂)
 - ▶ very large correlation contributions
- ▶ **Relativistic contributions** are small ($\approx 0.5\%$) but needed for agreement with experiment
 - ▶ nearly cancel quadruples contributions (error cancellation possible)
- ▶ **Vibrational contributions** are substantial
 - ▶ similar to triples contributions but oppositely directed (error cancellation)

CC convergence: bond distances (pm)

- ▶ Equilibrium bond distances – no vibrational corrections
 - ▶ complete basis-set limit results

	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 ₂	106.54	2.40	0.67	0.14	0.03	0.00	109.78	109.77	0.01
F ₂	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 F₂
- ▶ Hartree–Fock theory underestimates bond distances by up to 8.6 pm (for F₂)
- ▶ All **correlation contributions** are positive
 - ▶ approximate linear convergence, slowest for F₂
 - ▶ triples contribute up to 2.0 pm, quadruples up to 0.4 pm, and quintuples 0.03 pm
- ▶ **Relativistic corrections** are small except for F₂ (0.05 pm)
 - ▶ of the same magnitude and direction as the quintuples
 - ▶ no sources of error cancellation (in a complete AO basis)

CC convergence: harmonic constants ω_e (cm^{-1})

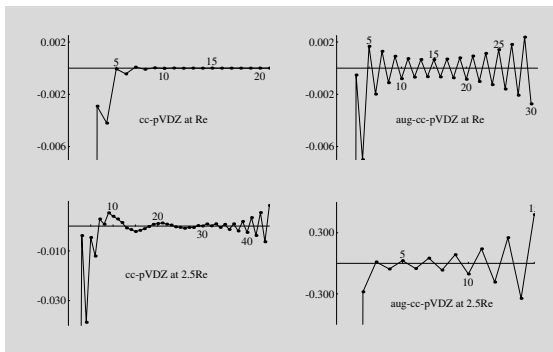
- ▶ Curvature of energy curve at equilibrium
 - ▶ complete basis-set limit results

	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 ₂	2730.3	-275.8	-72.4	-18.8	-3.9	-1.4	2358.0	2358.6	-0.6
F ₂	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 cm^{-1} except for F₂
- ▶ **Hartree–Fock theory** overestimates harmonic frequencies by up to 38% (in F₂).
- ▶ All **correlation contributions** are large and negative
 - ▶ triples contribute up to 95 cm^{-1} , quadruples 20 cm^{-1} , and quintuples 4 cm^{-1}
 - ▶ sextuples are sometimes needed for convergence to within 1 cm^{-1}
- ▶ **Relativistic corrections** are of the order of 1 cm^{-1}
 - ▶ of the same magnitude and direction as the quadruples or quintuples
- ▶ No sources of error cancellation in a complete AO basis

Many-body perturbation theory: approximate coupled-cluster theory

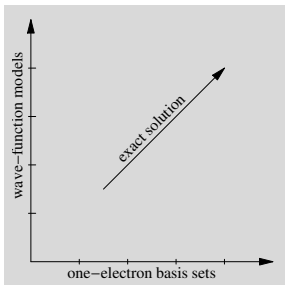
- ▶ Coupled-cluster amplitudes may be estimated by **perturbation theory**
- ▶ Caveat: the resulting perturbation series is **frequently divergent**, even in simple cases
 - ▶ here are some examples for the HF molecule (10 electrons):



- ▶ However, **to lowest order**, perturbational corrections are very useful and popular
 - ▶ **MP2** (approximate CCSD) and **CCSD(T)** (approximate CCSDT)
 - ▶ correlation effects are typically overestimated, leading to fortuitously good results

Basis-set convergence

- ▶ In all examples up to now, we have worked in a complete AO basis
- ▶ However, the overall quality is determined by the description of



- 1 the wave-function model
- 2 the AO basis set

- ▶ We shall now consider **basis-set convergence**
 - ▶ new sources of errors
 - ▶ new opportunities for error cancellation
- ▶ We begin by investigating convergence in the helium atom

AO convergence: the helium atom

- ▶ The helium atom contains only two electrons
 - ▶ ideal system to study basis-set convergence
 - ▶ pair interactions dominate also molecules
- ▶ Historical interest: the first many-body system treated with quantum mechanics
 - ▶ experimental ionization potential of helium: 24.59 eV
- ▶ Unsöld 1927: 20.41 eV
 - ▶ first-order perturbation theory—not much better than Bohr theory
- ▶ Hylleraas 1928: 24.47 eV
 - ▶ expansion in antisymmetric orbital products
 - ▶ excruciatingly slow AO convergence
- ▶ Hylleraas 1929: 24.58 eV
 - ▶ introduced the interelectronic coordinate r_{12} to arbitrary powers
 - ▶ the discrepancy of 0.01 eV due to relativistic corrections
 - ▶ full agreement between experiment and quantum mechanics
- ▶ Hylleraas discovered both the slow CI convergence and the efficacy of introducing r_{12}
 - ▶ the question of CI expansions vs. explicit correlation is still with us today

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

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

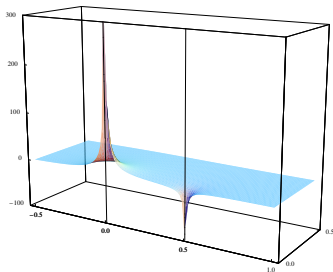
- ▶ at coalescence, the infinite potential terms must be canceled by infinite kinetic terms

- ▶ **Local kinetic energy** in the helium atom

- ▶ positive around the nucleus
- ▶ negative around the second electron

- ▶ **Negative kinetic energy** near coalescence

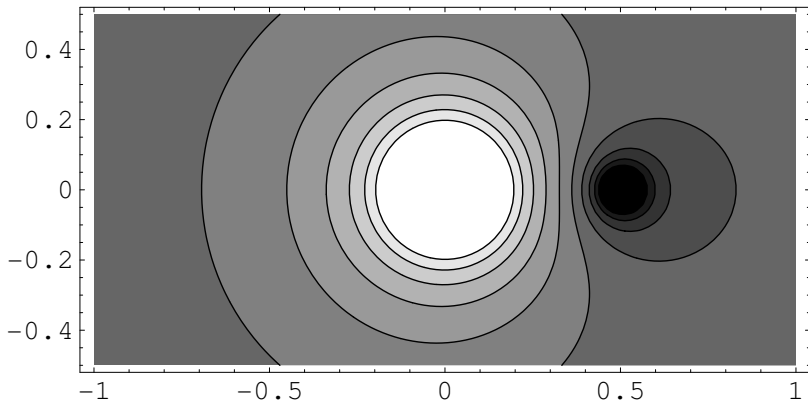
- ▶ classical forbidden region
- ▶ internal “tunneling”
- ▶ wavefunction decays towards the singularity
- ▶ the Coulomb hole



- ▶ The difficulty Hylleraas ran into was the description of the Coulomb hole!

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

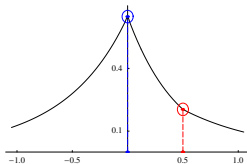


The electron cusp and the Coulomb hole

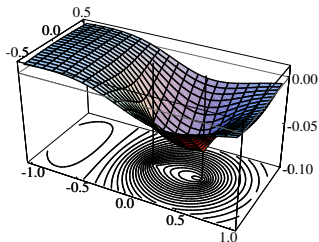
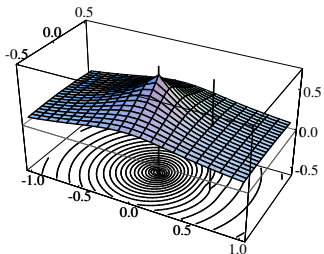
- ▶ The behaviour of the wave function at coalescence was established by Slater (1928)
 - ▶ nuclear and electronic cusp conditions

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

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



- ▶ The electronic cusp condition implies the existence of a **Coulomb hole**
 - ▶ the helium atom, with one electron fixed at a distance of $0.5a_0$ from the nucleus

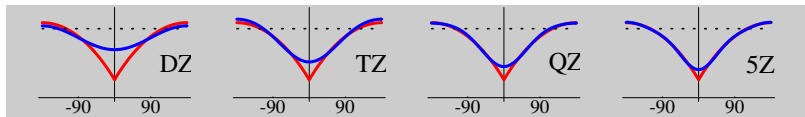


Basis-set convergence

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

	N_{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	∞	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):



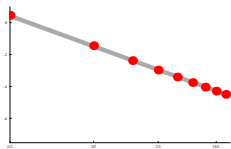
Principal expansion and correlation-consistent basis sets

- ▶ The energy contribution from each AO in large CI calculations on the helium atom depends only on the principal quantum number n :

$$\epsilon_{nlm} \approx n^{-6}$$

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

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



- ▶ 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:

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

Some observations

- ▶ The dependence of the error in the correlation energy on the number of AOs N :

$$\left. \begin{array}{l} \Delta \varepsilon_N \propto X^{-3} \\ N \propto X^3 \end{array} \right\} \Rightarrow \Delta \varepsilon_N \propto N^{-1}$$

- ▶ The dependence of the error in the correlation energy on the CPU time:

$$\left. \begin{array}{l} \Delta \varepsilon_N \propto N^{-1} \\ T \propto N^4 \end{array} \right\} \Rightarrow \Delta \varepsilon_n \propto T^{-1/4}$$

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

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

- ▶ The convergence is exceedingly **slow**!
- ▶ Clearly, a brute-force extension of the basis set until convergence (to some target accuracy) may not always be possible.
- ▶ Fortunately, the convergence is very **smooth**.
- ▶ We can treat slow basis convergence by **extrapolation** or **explicit correlation**

Extrapolations

- ▶ From two separate calculations with basis sets E_X and E_Y

$$E_\infty = E_X + AX^{-3}$$

$$E_\infty = E_Y + AY^{-3}$$

we eliminate A to obtain the following **two-point extrapolation formula**:

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

- ▶ Mean absolute error in the electronic energy of CH_2 , H_2O , HF , N_2 , CO , Ne , and F_2 :

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	

- ▶ For the error in the AE of CO relative to R12, we now obtain:

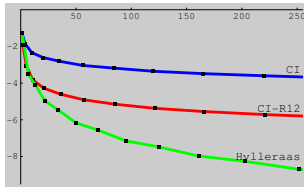
kJ/mol	DZ	TZ	QZ	5Z	6Z
plain	-73.5	-28.3	-11.4	-6.0	-3.5
extr.		-18.5	-0.7	0.0	0.0

- ▶ Chemical accuracy is now achieved with just 168 AOs (QZ), at a fraction of the cost.

Explicitly correlated methods

- ▶ To improve basis-set convergence, we may use *explicitly correlated wave functions* (Hylleraas, 1928).
- ▶ In such wave functions, the interelectronic distances r_{ij} are used as variables.
- ▶ In its simplest manifestation, only terms linear in r_{ij} occur: [the R12 method](#)

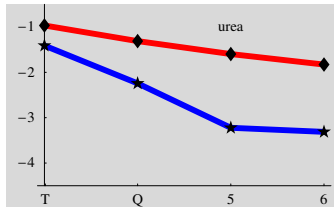
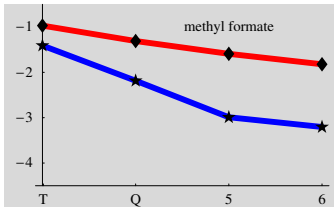
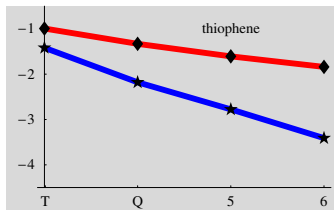
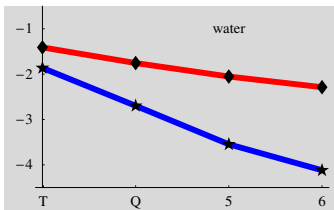
$$\Psi_{R12} = \sum_K C_K \Phi_K + C_{R12} \Phi_0$$



- ▶ Helium-atom convergence of
 - ▶ the [standard CI expansion](#)
 - ▶ the [CI-R12 expansion](#)
 - ▶ the [Hylleraas expansion](#)on a logarithmic energy scale:
- ▶ Here we investigate the convergence of traditional, orbital-based wave functions as parametrized in [coupled-cluster theory](#).
- ▶ However, we often use R12 or its modern variant F12 calculations as benchmarks.

Extrapolation vs. explicit correlation

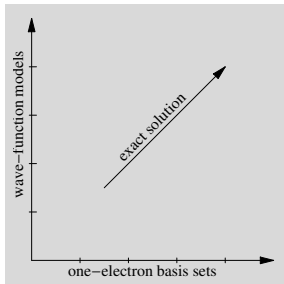
- ▶ For some purposes, extrapolation is competitive with explicit correlation
 - ▶ logarithmic errors in **plain** and **extrapolated** energies relative to R12



Calibration and benchmarking of electronic-structure methods

► We shall now explore the **two-dimensional chart of quantum chemistry**:

- 1 atomization energies
- 2 reaction enthalpies
- 3 molecular bond distances
- 4 vibrational frequencies



► Apart from overall **convergence**, we are interested in

- interplay between excitation-level and basis-set convergence
- sources of error cancellation

Initial comment: changes upon basis-set extension

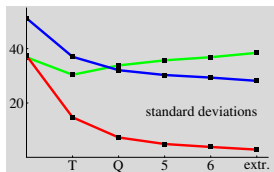
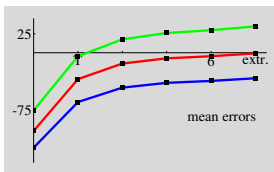
- ▶ When basis set increases, energy is lowered most for systems of low electronic energy
 - ▶ therefore energy differences increase upon basis-set extension
- ▶ This observation helps predict direction of change upon basis-set extension:
 - 1 AEs increase since molecular energies of bound systems are lower than atomic energies
 - 2 exothermic reactions become more exothermic since products have lower energy than reactants
 - 3 bond distances shorten since shorter bonds give lower electronic energy
 - 4 vibrational frequencies increase since shorter bonds gives larger force constants

Atomization energies (AEs)

- ▶ Let us consider the situation for an important molecular property: AEs

$$D_e = \sum_A E^A (2S+1L) - E(R_e)$$

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

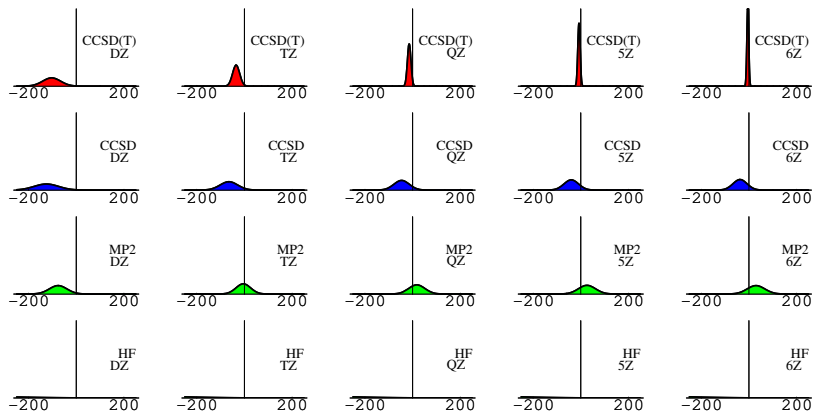


- ▶ AEs increase with cardinal number of basis set
- ▶ AEs increase with excitation level in the coupled-cluster hierarchy:

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

- ▶ MP2 overestimates doubles contribution
 - ▶ benefits from error cancellation at the MP2/TZ level
- ▶ CCSD(T) performs excellently, but large basis sets are needed
 - ▶ DZ and TZ basis are inadequate for CCSD(T)

Atomization energies (kJ/mol)



Comparison of CCSD(T) and experimental AEs (kJ/mol)

	cc-pCVQZ		cc-pCV(TQ)Z		cc-pCV6Z		exp
F ₂	153.4	-9.9	159.5	-3.9	161.1	-2.2	163.4(06)
H ₂	456.6	-1.5	458.4	0.4	458.1	0.1	458.0(00)
HF	586.1	-7.0	592.8	-0.3	593.3	0.1	593.2(09)
O ₃	583.6	-32.6	600.8	-15.4	605.5	-10.7	616.2(17)
HOF	649.5	-25.4	661.7	-13.2	662.9	-12.0	674.9(42)
CH ₂	751.3	-5.7	758.7	1.6	757.9	0.9	757.1(22)
HNO	842.7	-18.8	858.5	-3.0	860.4	-1.1	861.5(03)
N ₂	936.3	-19.9	952.3	-4.0	954.9	-1.3	956.3(02)
H ₂ O	963.5	-11.8	974.7	-0.5	975.5	0.2	975.3(01)
CO	1075.5	-11.2	1086.6	-0.1	1086.9	0.2	1086.7(05)
H ₂ O ₂	1108.0	-18.6	1125.5	-1.1	1126.1	-0.5	1126.6
NH ₃	1232.7	-15.1	1247.3	-0.6	1247.4	-0.5	1247.9(04)
HCN	1294.1	-18.6	1311.3	-1.4	1311.0	-1.7	1312.8(26)
CH ₂ O	1552.4	-14.2	1568.3	1.7	1568.0	1.4	1566.6(07)
CO ₂	1612.3	-20.1	1633.3	0.8	1633.2	0.7	1632.5(05)
C ₂ H ₂	1681.0	-16.8	1698.8	1.0	1697.1	-0.8	1697.8(10)
CH ₄	1749.9	-9.4	1762.0	2.7	1759.4	0.1	1759.3(06)
C ₂ H ₄	2343.6	-16.2	2363.2	3.4	2360.8	1.0	2359.8(10)

- ▶ Only O₃ and HOF have an error larger than 2.2 kJ/mol.
- ▶ Without extrapolation: cc-p(c)V6Z needed; with extrapolation: cc-pCV(TQ)Z sufficient
- ▶ Core correlation important: F₂ - 0.1 kJ/mol; CO - 4.8 kJ/mol; C₂H₂ - 10.8 kJ/mol

The (in)adequacy of CCSD(T) (kJ/mol)

	CCSD(T)		CCSDT		CCSDTQ		experiment	
	cc-pCV(56)Z		cc-pCV(Q5)Z		cc-pVTZ		D_e	D_0
CH ₂	757.9	-0.9	758.9	0.1	759.3	0.5	758.8	714.8±1.8
H ₂ O	975.3	0.1	974.9	-0.3	975.7	0.5	975.2	917.8±0.2
HF	593.2	0.0	593.0	-0.2	593.6	0.4	593.2	566.2±0.7
N ₂	954.7	-1.6	951.3	-5.0	955.2	-1.1	956.3	941.6±0.2
F ₂	161.0	-2.4	159.6	-3.8	162.9	-0.5	163.4	154.6±0.6
CO	1086.7	0.0	1084.4	-2.3	1086.7	0.0	1086.7	1071.8±0.5

- ▶ The excellent performance of CCSD(T) for AEs relies on error cancellation:
 - ▶ relaxation of triples from CCSD(T) to CCSDT reduces the AEs;
 - ▶ inclusion of quadruples from CCSDT to CCSDTQ increases the AEs.
- ▶ The error incurred by treating the connected triples perturbatively is quite large (about 10% of the full triples contribution) but canceled by the neglect of quadruples.
- ▶ The rigorous calculation of AEs to chemical accuracy requires CCSDTQ/cc-pCV6Z!

Reaction enthalpies

- ▶ Closed-shell, isogyric exothermic reactions

$$\Delta_r H_e^\circ(0K) = \sum_P E^P(R_e) - \sum_R E^R(R_e)$$

- ▶ Convergence faster than for AEs but still governed by short-range correlation:

	CCSD(T)	DZ	TZ	QZ	5Z	6Z
REs	plain	33.4	11.3	3.0	0.4	-0.1
	extr.		2.7	-1.1	-0.8	-0.3
AEs	plain	-103.3	-34.9	-14.3	-7.4	-4.7
	extr.		-16.2	-1.1	-0.6	-0.9

- ▶ Errors relative to experimental results corrected for vibrations and relativity:

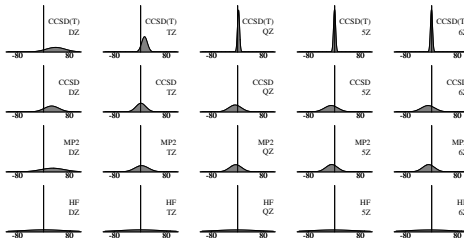
Mean absolute errors:

CCSD(T): 1.8 kJ/mol

CCSD: 14.5 kJ/mol

MP2: 13.5 kJ/mol

HF: 43.0 kJ/mol

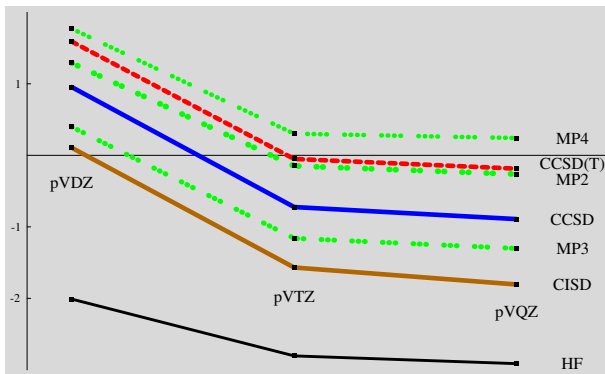


Enthalpies of some reactions (kJ/mol)

	exp.	R12	G2	B3LYP	(56)
$\text{CH}_2 + \text{H}_2 \rightarrow \text{CH}_4$	-544(2)	-542	-534	-543	-543
$\text{C}_2\text{H}_2 + \text{H}_2 \rightarrow \text{C}_2\text{H}_4$	-203(2)	-204	-202	-208	-206
$\text{C}_2\text{H}_2 + 3\text{H}_2 \rightarrow 2\text{CH}_4$	-446(2)	-447	-440	-450	-447
$\text{CO} + \text{H}_2 \rightarrow \text{CH}_2\text{O}$	-21(1)	-22	-17	-34	-23
$\text{N}_2 + 3\text{H}_2 \rightarrow 2\text{NH}_2$	-164(1)	-162	-147	-166	-165
$\text{F}_2 + \text{H}_2 \rightarrow 2\text{HF}$	-563(1)	-562	-564	-540	-564
$\text{O}_3 + 3\text{H}_2 \rightarrow 3\text{H}_2\text{O}$	-933(2)	-943	-912	-909	-946
$\text{CH}_2\text{O} + 2\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$	-251(1)	-250	-235	-234	-250
$\text{H}_2\text{O}_2 + \text{H}_2 \rightarrow 2\text{H}_2\text{O}$	-365(2)	-365	-360	-346	-362
$\text{CO} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{H}_2\text{O}$	-272(1)	-272	-251	-268	-273
$\text{HCN} + 3\text{H}_2 \rightarrow \text{CH}_4 + \text{NH}_2$	-320(3)	-320	-305	-320	-321
$\text{HNO} + 2\text{H}_2 \rightarrow \text{H}_2\text{O} + \text{NH}_2$	-444(1)	-445	-426	-429	-446
$\text{H}_2\text{O} + \text{F}_2 \rightarrow \text{HOF} + \text{HF}$	-129(4)	-118	-123	-119	-118
$\text{CO}_2 + 4\text{H}_2 \rightarrow \text{CH}_4 + 2\text{H}_2\text{O}$	-244(1)	-242	-216	-211	-244
$2\text{CH}_2 \rightarrow \text{C}_2\text{H}_4$	-844(3)	-842	-829	-845	-845

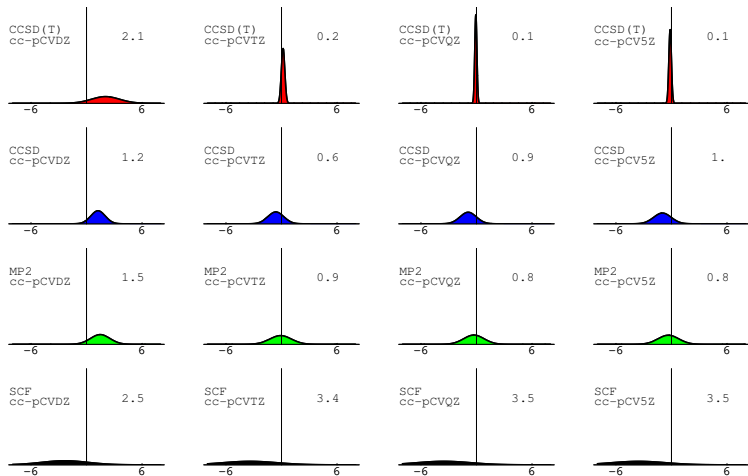
Bond distances

- ▶ Statistics based on 28 bond distances at the all-electron cc-pVXZ level (pm):



- ▶ bonds shorten with increasing basis: $DZ > TZ > QZ$
- ▶ bonds lengthen with increasing excitations: $HF < CCSD < MP2 < CCSD(T)$
- ▶ considerable scope for error cancellation: $CISD/DZ, MP3/DZ$
- ▶ CCSD(T) mean errors: $DZ: 1.68 \text{ pm}; TZ: 0.01 \text{ pm}; QZ: -0.12 \text{ pm}$

Errors in bond distances R_e of BH, CO, N₂, HF, and F₂ (pm)



- ▶ Comparison with experimental and empirical equilibrium bond distances
 - ▶ empirical values from experimental rotation constants and calculated vibration–rotation constants

		HF	MP2	CCSD	CCSD(T)	emp.	exp.
H ₂	<i>R</i> _{HH}	73.4	73.6	74.2	74.2	74.1	74.1
HF	<i>R</i> _{FH}	89.7	91.7	91.3	91.6	91.7	91.7
H ₂ O	<i>R</i> _{OH}	94.0	95.7	95.4	95.7	95.8	95.7
HOF	<i>R</i> _{OH}	94.5	96.6	96.2	96.6	96.9	96.6
HNC	<i>R</i> _{NH}	98.2	99.5	99.3	99.5	99.5	99.4
NH ₃	<i>R</i> _{NH}	99.8	100.8	100.9	101.1	101.1	101.1
N ₂ H ₂	<i>R</i> _{NH}	101.1	102.6	102.5	102.8	102.9	102.9
C ₂ H ₂	<i>R</i> _{CH}	105.4	106.0	106.0	106.2	106.2	106.2
HCN	<i>R</i> _{CH}	105.7	106.3	106.3	106.6	106.5	106.5
C ₂ H ₄	<i>R</i> _{CH}	107.4	107.8	107.9	108.1	108.1	108.1
CH ₄	<i>R</i> _{CH}	108.2	108.3	108.5	108.6	108.6	108.6
N ₂	<i>R</i> _{NN}	106.6	110.8	109.1	109.8	109.8	109.8
CH ₂ O	<i>R</i> _{CH}	109.3	109.8	109.9	110.1	110.1	110.1
CH ₂	<i>R</i> _{CH}	109.5	110.1	110.5	110.7	110.6	110.7
CO	<i>R</i> _{CO}	110.2	113.2	112.2	112.9	112.8	112.8
HCN	<i>R</i> _{CN}	112.3	116.0	114.6	115.4	115.3	115.3
CO ₂	<i>R</i> _{CO}	113.4	116.4	115.3	116.0	116.0	116.0
HNC	<i>R</i> _{CN}	114.4	117.0	116.2	116.9	116.9	116.9
C ₂ H ₂	<i>R</i> _{CC}	117.9	120.5	119.7	120.4	120.4	120.3
CH ₂ O	<i>R</i> _{CO}	117.6	120.6	119.7	120.4	120.5	120.3
N ₂ H ₂	<i>R</i> _{NN}	120.8	124.9	123.6	124.7	124.6	124.7
C ₂ H ₄	<i>R</i> _{CC}	131.3	132.6	132.5	133.1	133.1	133.1
F ₂	<i>R</i> _{FF}	132.7	139.5	138.8	141.1	141.3	141.2
HOF	<i>R</i> _{OF}	136.2	142.0	141.2	143.3	143.4	143.4

Bond distances: CCSD(T) error cancellation

- ▶ The high accuracy of CCSD(T) arises partly because of **error cancellation**
- ▶ CCSD(T) distances compared with experiment:

pm	DZ	TZ	QZ
$\overline{\Delta}$	1.68	0.01	-0.12
$ \overline{\Delta} $	1.68	0.20	0.16

- ▶ Bond distances are further reduced by
 - ▶ basis-set extension QZ \rightarrow 6Z: ≈ -0.10 pm
 - ▶ full triples treatment CCSD(T) \rightarrow CCSDT: ≈ -0.04 pm
- ▶ Intrinsic error of the CCSDT model: ≈ -0.2 pm
- ▶ Connected quadruples increase bond lengths by about 0.1–0.2 pm

Error cancellation: some general considerations

- ▶ In a **complete basis**, the CC hierarchy converges from below to BDs and AEs:
 - ▶ $\text{HF} < \text{CCSD} < \text{CCSDT} < \text{CCSDTQ} < \dots < \text{FCI} = \text{exp}$ (BDs and AEs)
 - ▶ no error cancellation may occur (ignoring relativity and adiabatic corrections)
- ▶ In a **finite basis**, BDs become longer and AEs smaller
 - ▶ $\text{HF} < \text{CCSD} \approx \text{exp} < \text{CCSDT} < \text{CCSDTQ} < \dots < \text{FCI}$ (BDs shifted up)
 - ▶ $\text{HF} < \text{CCSD} < \text{CCSDT} < \text{CCSDTQ} < \dots < \text{FCI} < \text{exp}$ (AEs shifted down)
 - ▶ many opportunities for error cancellation of BDs
- ▶ Perturbative treatments of connected excitations tend to overestimate their effect:
 - ▶ $\text{HF} < \text{CCSD} < \text{MP2}$
 - ▶ $\text{CCSD} < \text{CCSDT} < \text{CCSD(T)}$
- ▶ For BDs and AEs, we now observe
 - ▶ $\text{HF} < \text{CCSD} < \text{MP2} \approx \text{CCSD(T)} \approx \text{exp}$ (BDs)
 - ▶ $\text{HF} < \text{CCSD} < \text{CCSD(T)} \approx \text{exp} < \text{MP2}$ (AEs)
 - ▶ error cancellation may now occur also for AEs
- ▶ Vibrational frequencies behave like BDs since basis-set extension and increased excitation levels work in opposite directions

Vibrational frequencies of diatoms

- ▶ The frequency of a diatomic fundamental transition is given by

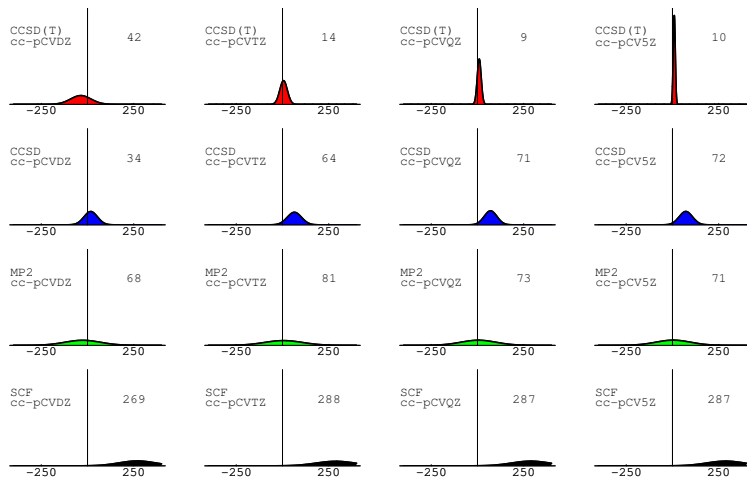
$$\nu = \omega_e - 2\omega_e x_e$$

- ▶ the harmonic constant ω_e requires 2nd derivatives of PES
- ▶ the anharmonic constant $\omega_e x_e$ requires 4th derivatives of PES
- ▶ It is easier to converge harmonic than anharmonic constants:

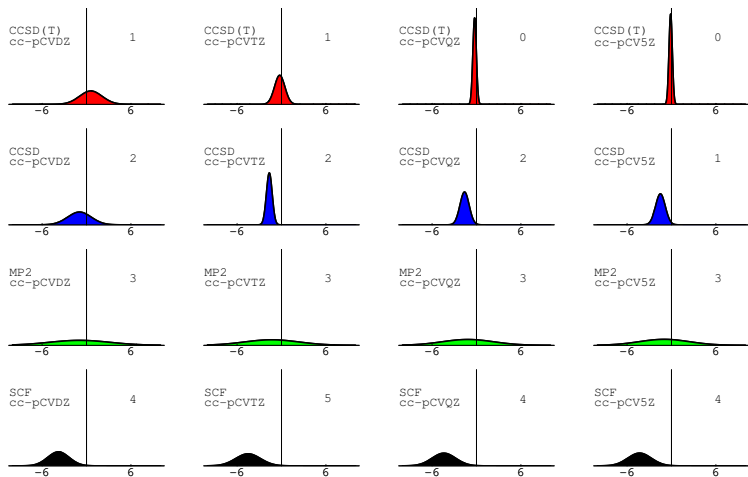
$ \Delta $ (cm ⁻¹)	RHF	MP2	CCSD	CCSD(T)	judgment
ω_e	287	71	72	10	DIFFICULT!
$\omega_e x_e$	4	3	1	0	EASY!

- ▶ Historical note:
 - ▶ With the advent of CCSD(T) in the 1990s, it was soon agreed that this model is capable of highly accurate vibrational constants (to within a few wavenumbers at the TZ level)
 - ▶ However, with the development of codes capable of handling very large basis sets and high excitation levels, it has slowly transpired that things are perhaps not so simple

Harmonic constants ω_e of BH, CO, N₂, HF, and F₂ (cm⁻¹)

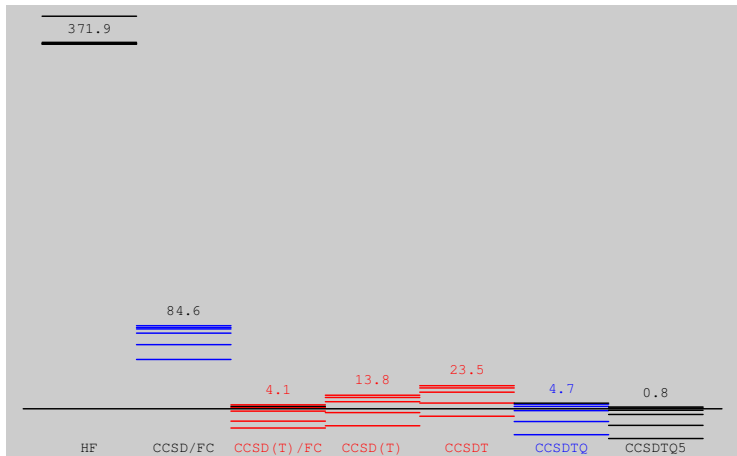


Anharmonic constants $\omega_e x_e$ of BH, CO, N₂, HF, and F₂ (cm⁻¹)



Higher-order connected contributions to ω_e in N_2

- ▶ There are substantial higher-order corrections:



- ▶ connected triples relaxation contributes 9.7 cm^{-1} (total triples -70.5 cm^{-1})
- ▶ connected quadruples contribute -18.8 cm^{-1}
- ▶ connected quintuples contribute -3.9 cm^{-1}

Conclusions

- ▶ **Two-dimensional chart** of nonrelativistic quantum chemistry
 - ▶ 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, perturbation theory
 - ▶ small basis sets
 - ▶ errors may cancel
- ▶ 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
 - ▶ knowledge of methods and awareness of pitfalls needed
- ▶ Discipline is always needed!